# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

**Title**
Many-Body Quantum Information Dynamics

**Permalink**
https://escholarship.org/uc/item/2694w4p8

**Author**
SCHUSTER, THOMAS

**Publication Date**
2022

Peer reviewed|Thesis/dissertation

Many-Body Quantum Information Dynamics

by

Thomas Schuster

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Norman Y. Yao, Chair
Professor Joel E. Moore
Professor Umesh Vazirani

Fall 2022

Many-Body Quantum Information Dynamics

Abstract

Many-Body Quantum Information Dynamics

by

Thomas Schuster

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Norman Y. Yao, Chair

Quantum information has emerged as a unifying force in our understanding of large-scale quantum mechanical systems. In quantum technologies, the precise dispersal of quantum information enables computational advantages beyond those of any classical computer. In quantum materials, spatial patterns in the quantum information of electrons give rise to topological phases, with exotic properties when these patterns are "cut" by a material boundary. In more remarkable systems still, quantum information appears central to the unification of quantum mechanics and gravity, one of the frontier open questions in our understanding of the universe. At the same time, the emergence of quantum technologies and accompanying theoretical developments have motivated a widespread interest in *dynamical* quantum phenomena. In this thesis, I explore several questions at the interface of quantum information and many-body quantum dynamics. Can dynamical many-body quantum systems realize stable coherent quantum phenomena—for example, topological phases—beyond those of static systems? How does quantum information move under many-body dynamics? How can we characterize this movement, and what are its physical implications? And finally, what are the best strategies to learn properties of unknown quantum dynamics from measurements involving them? These inquiries are united by their pursuit to realize and understand large-scale quantum mechanical phenomena in the world around us.

# Contents

# List of Figures

# List of Tables

# Acknowledgments

Many people have contributed to my growth as physicist throughout graduate school.

First and foremost, of course, I thank my advisor Norman Yao. Norm has been the most singular factor in a rewarding Ph.D. experience. His boundless enthusiasm for physics never ceases to amaze me, and I am fortunate to have been able to work with him as long as I have. I am particularly grateful for his confidence in me in the early years of my Ph.D., and his tireless support and advice in more recent years. I have learned many lifelong skills from Norm, including how to formulate interesting scientific questions, how to present engaging and honest scientific stories, and the value of pursuing science as a social endeavor. More than all, it has simply been fun.

I am also grateful to the Yao group for providing a vibrant and supportive research environment. It has been a joy to watch the group flourish from its first years to today. I especially thank my fellow graduate students Francisco Machado, Kamphol Akkaravarawong, and Bingtian Ye for innumerable conversations, scientific and otherwise. Above all, I thank Bryce Kobrin for our many and ongoing collaborations, which constitute some of my most productive and joyful scientific experiences.

I have also received invaluable mentorship from many others. In particular, I would like to thank Joel Moore for his guidance in my early days of graduate school and our collaborations on the Hopf insulator. I am also lucky to have collaborated with Beni Yoshida, whose work sparked my initial interest in quantum information dynamics and has formed the basis of much of my understanding since. I also thank many other friends and collaborators, in particular Snir Gazit, Quntao Zhuang, Felix Flicker, Soonwon Choi, Masoud Mohseni, Murphy Niu, and Tom O'Brien, for their advice and guidance throughout my career thus far. Finally, I sincerely thank Claudio Chamon and Tom Iadecola for their mentorship during the year between my undergraduate and graduate studies. Their guidance in this critical early period of my research career allowed this Ph.D. experience to be possible.

I will conclude with a few people from my personal life. I thank David Keating for our close friendship over the last decade. I also thank my parents, Mary and Steve, and my sister, Claire, for their love and support. Lastly, I thank my partner, Haixing. Our life together forms the foundation on which this work has been worthwhile.

# Citations to previously published work

The work presented in this thesis is adapted from previously published work. In order of appearance:

Chapter 3 is adapted from:
Thomas Schuster et al. "Realizing Hopf insulators in dipolar spin systems". In: Physical Review Letters 127.1 (2021), p. 015301.

Chapter 4 is adapted from:
Thomas Schuster et al. "Floquet engineering ultracold polar molecules to simulate topological insulators". In: Physical Review A 103.6 (2021), p. 063322.

Chapter 5 is adapted from:
Thomas Schuster et al. "Floquet hopf insulators". In: Physical Review Letters 123.26 (2019), p. 266803.

Chapter 6 is adapted from:
Jiho Noh et al. "Braiding photonic topological zero modes". In: Nature Physics 16.9 (2020), pp. 989-993.

Chapter 7 is adapted from:
Kevin A Landsman et al. "Verified quantum information scrambling". In: Nature 567.7746 (2019), pp. 61-65.

Chapter 8 is adapted from:
Machiel S Blok et al. "Quantum information scrambling on a superconducting qutrit processor". In: Physical Review X 11.2 (2021), p. 021010.

Chapter 9 and parts of Chapter 2.3 are adapted from:
Thomas Schuster et al. "Many-body quantum teleportation via operator spreading in the traversable wormhole protocol". In: Physical Review X 12.3 (2022), p. 031013.

Chapter 10 is adapted from:
Thomas Schuster and Norman Y Yao. "Operator Growth in Open Quantum Systems". In: arXiv preprint arXiv:2208.12272 (2022).

Chapter 11 is adapted from:
Quntao Zhuang et al. "Scrambling and complexity in phase space". In: Physical Review A 99.6 (2019), p. 062334.

Chapter 12 is adapted from:
Thomas Schuster et al. "Learning quantum systems via out-of-time-order correlators". In: arXiv preprint arXiv:2208.02254 (2022).

Chapter 13 and parts of Chapter 2.3 are adapted from:
Jordan Cotler, Thomas Schuster, and Masoud Mohseni. "Information-theoretic Hardness of Out-of-time-order Correlators". In: arXiv preprint arXiv:2208.02256 (2022).

# Part I

# Introduction

# Chapter 1

# Introduction to Many-body quantum information dynamics

Physics' central concern is the movement of objects through space and time. In the world around us, the objects we interact with are in fact collections of enormous numbers of more elementary components. For example, a cup of water is composed of $\sim 10^{23}$ individual $H_2O$ molecules. Many-body physics seeks to understand how collective properties of the larger object arise from the object's elementary components and the interactions between them.

The field of many-body physics is built on two twin themes. First, the collective behavior of the larger object is often much simpler than the sum of its components. For instance, to describe the sloshing of water in the cup, one does not need to keep track of the motion of each individual $H_2O$ molecule. This allows many-body physics to be tractable. Second, the collective behavior is not easily predicted from the components. One cannot predict whether water in the cup will be a liquid, rather than frozen or a vapor, from the properties of a single $H_2O$ molecule. It is necessary to study large numbers of molecules interacting together. This allows many-body physics to be interesting.

Traditionally, the study of many-body physics has centered on condensed matter systems, corresponding to solids or liquids composed of innumerable atoms or molecules. For example, a common question might be how the conductivity of a material changes as a function of the temperature or the density of some impurity. Surprisingly, many of the phenomena uncovered are intrinsically *quantum mechanical*. For instance, the reason that some materials do not conduct electricity is because their electrons cannot move unless they each absorb a minimum "quanta" of energy. More exotic phases of matter, such as superconductors, are enabled by correlated quantum behavior between large numbers of electrons. The fact that quantum mechanics—a theory to describe nature at the smallest of scales—underlies such large-scale phenomena is fascinating. Theoretical frameworks such as quantum field theory and the renormalization group allow us to understand these phenomena, and are amongst the seminal achievements of physics in the latter twentieth century.

Nevertheless, the range of quantum phenomena accessible in condensed matter systems is limited. To state it roughly, we can only realize certain quantum states and can only

experiment on them in certain ways. The first restriction arises because condensed matter systems are usually in *equilibrium* with their environment. This means that they only realize quantum states that do not change substantially in time. Even dynamical features of these systems, such as their conductance, can often be understood via properties of a time-independent equilibrium state. The second restriction arises because condensed matter experiments mostly involve "larger objects". The experimentalist typically has no ability to probe the "elementary components" of the system, such as the individual electrons or atoms or molecules.

The emergence of quantum technologies has broken both of these paradigms. The defining feature of quantum technologies is their precise control over microscopic quantum components. This precise control in turn allows them to realize a vast range of quantum states and phenomena. The microscopic quantum components in quantum technologies can take many forms. They might be atoms themselves [60], or the nuclear [76] or electronic [141] states of the atoms, or junctions between superconducting materials [30], or electrons trapped at material defects [424], and so on [23]. In all cases, the components are made to be almost completely isolated from the rest of the world. The only (intended) exception is the control of the experimenter themselves, which is usually provided through precise lasers or microwave drives.

This technological breakthrough has paralleled a shift in the way we think about many-body quantum systems. Conventionally, many-body physics has focused on the behavior of so-called "local" quantities of a system, for example the electron density or energy density. These are the properties that are directly accessible in everyday life as well as condensed matter experiments. Over the last several decades, a confluence of scientific developments and fundamental questions have led to a more abstract viewpoint. Instead of local quantities, this viewpoint centers on the behavior of information—specifically, *quantum information*—itself [376]. I list a few prominent contributors to this shift below. The work in this thesis will touch on each of these topics.

- *Topological phases of matter*—Attempts to understand the experimental observations of high-temperature superconductivity [27] and the fractional quantum Hall effect [467] eventually led to the notion of topological phases of matter [488]. Conventional phases of matter such as solids or liquids or gases are distinguished by how their local quantities (in these cases, the atom or molecule density) behave in the system's equilibrium state. Topological phases of matter are distinguished by how quantum information behaves in the equilibrium state. Topological phases display a number of remarkable and bizzare phenomena that have re-drawn the boundaries of what was thought possible in many-body quantum mechanics [488].

- *Classical simulation of quantum systems*—The study of physics is greatly assisted by numerical simulations. However, simulating the behavior of a many-body quantum system is in general exponentially difficult. Exact simulation methods can only incorporate small numbers ($\sim$20-30) of elementary quantum components. To move beyond

this, simulation algorithms must decide what information about the system to keep track of and what to neglect. To decide this intelligently, we need to understand the physics of information in many-body quantum systems, and how it connects to system properties of interest. This line of research began with the development of tensor network algorithms for equilibrium states in one-dimensional systems [490]. Extending these methods to different physical settings [358, 521] and quantum dynamics [286, 489] are topics of vigorous interest.

- *Equilibration in quantum systems*—A standard lore in thermodynamics is that a system loses memory of its initial state as information is exchanged with an environment. This explains why many different initial states can evolve to a common equilibrium. However, this story is more subtle in the absence of an environment. In classical systems, chaos serves to create an *effective* loss of memory. In quantum systems, the unitary nature of time-evolution precludes such an understanding. Over the last decade, we have learned that the equilibration of many-body quantum systems is instead intimately entwined with the dynamics of quantum information. In particular, equilibration can be understood as the dispersal of initially local quantum information into complex non-local correlations [240, 384], which are effectively invisible to experimenters with access to local quantities [395].

- *Quantum gravity*—Despite a half century of dedicated pursuit, the unification of quantum mechanics and gravity remains elusive. In the 1990s, the AdS/CFT correspondence provided a path forward: a precise mapping between certain theories of quantum gravity and certain many-body quantum systems [313]. Over the last decade, it has become clear that quantum information lies at the heart of this correspondence [196]. The structure of spacetime is encoded in the correlations of quantum information within the many-body quantum system [409], and the motion of particles in spacetime is governed by quantum information dynamics [249, 436].

- *Quantum technologies and quantum computing*—The precise manipulation of quantum information is central to the future promise of quantum computing and other quantum technologies. These endeavors intersect with many-body physics in eclectic ways. For instance, one of the leading architectures for fault-tolerant quantum computing, the surface code [69, 150], is directly inspired by a topological phase of matter [248]. The quest for quantum supremacy [30] has also accelerated interest in the classical simulability of many-body quantum dynamics [11]. Finally, a primary near-term application of quantum technologies is to simulate many-body quantum systems. Quantum simulators can allow unprecedented experimental access to quantum information properties [163, 224], which in turn feeds in to each of the research directions above.

In most of these cases, the tools of quantum information theory used are quite rudimentary. The interest is in how these tools intersect with many-body physics.

In this thesis, we will be specifically focused on the *dynamics* of many-body quantum information. Indeed, many of the most exciting connections above—to quantum thermalization, classical simulation, and quantum gravity—occur in a dynamical setting. Dynamics are also natural for current quantum technologies, in which imperfect isolation and control prevent easy access to many equilibrium quantum states [375]. Below, I list three broad directions at the intersection of quantum information and many-body quantum dynamics that motivate the works that follow. Of course, in real life, we accept insight and interest where we find them, and the results are not quite so neatly categorized.

Our first direction concerns stable quantum phenomena in many-body dynamics. Specifically, we will show that dynamics can provide both simpler realizations of static topological phases, as well as allow the existence of new topological phenomena altogether. Our second direction concerns the dynamics, or "scrambling", of quantum information itself. We will largely focus on how information at one time is stored in a system at later times. This notion underlies wide-ranging connections across quantum chaos, thermodynamics, quantum communication, the classical simulability of quantum systems, and quantum gravity. Our third and final direction concerns the process of learning information about an unknown quantum system from measurements involving it. We will show that there exist exponential separations in the learnability of many-body quantum dynamics depending on the classes of measurements that can be performed.

In more detail, the structure of the thesis is as follows.

- We first continue in the current Part with background material related to the results of this thesis. In Ch. 2, we provide a rapid introduction to the basic elements of quantum information theory, which will be instrumental to each of the Parts in the remainder of the work. The remaining sections of this Chapter, Ch. 2.2, 2.3, and 2.4, contain introductory material for Parts II, III, and IV, respectively. These introductions can be read in sequence (i.e. in section order), or independently preceding their corresponding Part. Ch. 2.2 contains a brief introduction to topological insulators; Ch. 2.3 to quantum information dynamics; and Ch. 2.4 to aspects of quantum learning.

- In Part II, we present results on topological phenomena in quantum dynamics. We primarily focus on topological insulators, which are some of the most widespread examples of topological phases. In Ch. 3, we present a scheme to realize the hitherto unobserved Hopf topological insulator using long-range dipolar interactions and strong periodic driving. In Ch. 4, we provide extensive details on how this scheme can be implemented in optical lattices of polar molecules, including novel polar molecule quantum simulation techniques. In Ch. 5, we present an intrinsically dynamical extension of the Hopf insulator—the "Floquet Hopf insulator"—which possesses a few particularly bizarre properties that differ from all other Floquet topological insulators to date. In Ch. 6, we turn away from topological insulators and present the experimental observation of an analogue of non-Abelian braiding statistics in topological defect modes of dynamical but non-interacting photons.

- In Part III, we present results on quantum information dynamics in many-body systems. In Chs. 7 and 8, we present experimental measurements of quantum information dynamics in trapped ion and superconducting quantum processors, respectively. For the sake of this thesis, these experiments are notable for two reasons: (i) they represented a proof-of-principle demonstration of a novel "many-body" generalization of quantum teleportation, and (ii) through teleportation, they achieved the first experimental detection of quantum information dynamics that was robust to experimental noise. In Ch. 9, we expand upon point (i) and present a generalized mechanism for many-body teleportation in a protocol arising from quantum gravity. In Ch. 10, we expand upon point (ii) and provide a full account of how quantum information dynamics are modified in the presence of noise. Finally, in Ch. 11, we extend the framework of quantum information dynamics to continuous variable systems, and forge connections between quantum chaos and information scrambling in this context.

- In Part IV, we present results on quantum learning. Specifically, we focus on learning properties of dynamical quantum processes. In Ch. 12, we show that various properties of local strongly-interacting dynamics can be learned much more efficiently with "out-of-time-order" experiments (introduced in Ch. 2.3), compared to conventional time-ordered experiments. In Ch. 13, we present a theoretical framework that encapsulates this distinction, and prove that there exist tasks for which out-of-time-order experiments provide an exponential advantage in the system size over time-ordered experiments. As a corollary, this result rigorously establishes that quantum information dynamics—i.e. the entire content of Part III—can in general only be efficiently measured in experiments that involve reversed time-evolution.

# Chapter 2

# Background on quantum information and many-body physics

In this Chapter we give several brief introductions to concepts in quantum information theory and many-body physics that will be essential to the remainder of the thesis. We begin by introducing a small handful of basic concepts in quantum information theory, then proceed to introduce topological phases, quantum information dynamics, and quantum learning theory.

## 2.1   Quantum information theory

In this section, we quickly introduce the notions of entropy and mutual information, quantum teleportation, and quantum complexity. For a more complete introduction to quantum information we refer to the resources [353, 483, 494].

### From classical to quantum information

We begin by defining two key concepts within classical information theory, and then quickly progress to the quantum case. These concepts are the entropy and the mutual information.

Consider a probability distribution $p(x)$ defined over a discrete random variable $x \in \mathcal{X}$. The entropy seeks to quantify the amount of "uncertainty" or "surprise" in $p(x)$. The Shannon entropy is defined as:

$$S(p) = -\sum_{x \in \mathcal{X}} p(x) \log(p(x)). \tag{2.1}$$

We can interpret the quantity $-\log(p(x))$ as a measure of the "surprise" associated with the outcome $x$. Symbols with lower values of $p(x)$ are less likely to occur, and thus associated with a greater surprise when they do occur. The entropy can then be viewed as the expectation value of the surprise.

The choice of the functional form $-\log(p(x))$ for the surprise initially appears quite arbitrary. Could one have equally chosen any monotonically decreasing function of $p(x)$, for instance $-p(x)$? To some extent, the answer is yes. Choices that are monomials in $p(x)$ define the so-called Renyi entropies. Specifically, the $\alpha$-order Renyi entropy is defined as:

$$S^{(\alpha)}(p) = \frac{1}{1-\alpha} \log\left(\sum_{x \in \mathcal{X}} p(x)^{\alpha}\right). \tag{2.2}$$

In many cases, the Renyi entropies behave qualitatively similar to the Shannon entropy, although for specific $p(x)$ they can differ sharply. The Shannon entropy also obeys certain properties (the "chain rule of conditional probability") that the Renyi entropies lack, which are important in certain situations. This will come up shortly when defining the mutual information.

We now turn to probability distributions $p(x,y)$ over multiple variables (in this case, two) $x \in \mathcal{X}, y \in \mathcal{Y}$. We can define the entropy over both variables $S(p(x,y))$ as before. We can also define the entropy of a single variable by ignoring the values of the other variable, e.g. defining the probability distribution $p(x) = \sum_{y \in \mathcal{Y}} p(x,y)$ gives $S(p(x))$.

The mutual information seeks to quantify the amount of mutual dependence between the two random variables. The mutual information is defined as:

$$I(\mathcal{X}, \mathcal{Y}) = S(p(x)) + S(p(y)) - S(p(x,y)). \tag{2.3}$$

In keeping with its goal, the mutual information is zero when the probability distribution is separable, $p(x,y) = p_1(x)p_2(y)$. For two binary variables, the mutual information is maximal for the probability distribution $p(0,0) = p(1,1) = 1/2$, $p(0,1) = p(0,1) = 0$, with value $I = \log(2)$. A Renyi mutual information can be defined similarly (replacing all Shannon entropies with Renyi entropies), however it is not guaranteed to be semi-positive. This poses a serious conceptual problem, since a probability distribution could have "less" Renyi mutual information than a separable distribution.

We now turn to quantum systems. The natural quantum analog of a random variable with probability distribution $p(x)$ is a density matrix $\rho$ defined on a Hilbert space $\mathcal{H}$. The von Neumann entropy of $\rho$ is defined as:

$$S(\rho) = -\text{tr}(\rho \log(\rho)). \tag{2.4}$$

We can similarly define the Renyi entropies:

$$S^{(\alpha)}(\rho) = -\frac{1}{1-\alpha} \text{tr}(\rho^{\alpha}). \tag{2.5}$$

These are simply the classical entropies associated with the probability distribution defined by the eigenvalues of $\rho$.

Now consider a density matrix $\rho$ defined on a tensor product Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. The quantum mutual information between subsystems $A$ and $B$ is defined as:

$$I(A, B) = S_A + S_B - S_{AB}, \tag{2.6}$$

where we abbreviate $S_{AB} = S(\rho)$, $S_A = S(\text{tr}_B(\rho))$, $S_A = S(\text{tr}_A(\rho))$. Similar to the classical case, the mutual information is zero for separable states $\rho = \rho_A \otimes \rho_B$. Interestingly, the maximal value of the quantum mutual information is $2\log(2)$, twice that of the classical mutual information. This is achieved by the EPR state $\rho = |\text{EPR}\rangle\langle\text{EPR}|$, with $|\text{EPR}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. The key difference here is that the entire quantum system can be pure (entropy zero) even when each subsystem is maximally mixed (entropy maximal). This cannot occur in classical systems.

## Quantum teleportation

We have already seen one way in which quantum information differs from classical information, in that a quantum system can be pure even when its subsystems are maximally mixed. We will now outline a much more stark example of the difference between quantum and classical information: quantum teleportation [370].

To set the stage, consider the process of transferring classical information from some location A to another location B. Under the laws of classical physics, the only way to perform such a transfer is to physically "march" the information contained at location A to location B. Here by "march" we mean that the information must physically traverse some path between the two locations (i.e. to recover the information at some intermediate time within the transfer, one would need access to physical locations along the path).

It turns out that to transfer quantum information from one location to another, one only has to physically march *classical* information between them. The process by which this occurs is known as quantum teleportation. It in fact requires one additional ingredient: an entanglement resource shared between location A and location B before the transfer. In its simplest form, this entanglement resource can be an EPR pair.

We now outline the prototypical quantum teleportation protocol for transferring a single qubit [353]. We suppose that location A and B share an initial EPR pair $|\text{EPR}\rangle$, and that the qubit is initially at location A in the state $|\psi\rangle$. The total system is thus in the state $|\psi\rangle \otimes |\text{EPR}\rangle$, with the first two qubits located at A and the third qubit at B. To teleport the first qubit from A to B, we first perform a Bell measurement on the two qubits at location A. A Bell measurement has four possible outcomes $\{P_l |\text{EPR}\rangle\}$, where $P \in \{\mathbb{1}, X, Y, Z\}$ runs over the single-qubit Pauli matrices and the subscript $l$ denotes that $P$ acts on the left qubit of the EPR pair (i.e. the first qubit of the system). After performing the Bell measurement, the measurement result $P$ is communicated from location A to B, requiring two classical bits. The state $|\psi\rangle$ is obtained on the third qubit upon applying the Pauli operation corresponding to $P$. The reason this protocol succeeds is most easily understood using the diagrammatic notation introduced in Ch. 9.

## Quantum complexity

We conclude the Chapter with a brief discussion of quantum complexity (in reality, a subject of its own). We will jump straight to quantum complexity and neglect the classical

case.

Consider a quantum computer with a gate set $\mathcal{G}$. For instance, the gate set might consist of every possible single-qubit and two-qubit gate. If the system of interest has some notion of locality, we might restrict these gates to be geometrically local. We could also restrict that the gates respect a particular symmetry operation on the system.

We define the relative complexity [70, 353] $\mathcal{C}(|\psi\rangle, |\phi\rangle)$ of a quantum state $|\psi\rangle$ with respect to a reference state $|\phi\rangle$ as the minimal length of a gate sequence $g_1, \ldots, g_{\mathcal{C}} \in \mathcal{G}^{\otimes \mathcal{C}}$ such that:

$$g_1 \cdot \ldots \cdot g_{\mathcal{C}} |\phi\rangle = |\psi\rangle . \tag{2.7}$$

The relative complexity defines a metric on the set of quantum states as long as the gate set $\mathcal{G}$ is closed under gate inversion. This metric will be useful in characterizing topological phases of many-body quantum systems in Ch. 2.2.

We can also define the quantum complexity of a computation. We consider a quantum computer with $n$ qubits, and we suppose that the qubits are always initialized in the state $|0\rangle^{\otimes n}$ and read out in the computational basis. In between initialization and read-out we apply some unitary gate sequence drawn from $\mathcal{G}$.

To define the computational complexity, suppose we would like to use the quantum computer to evaluate a function $f : \mathcal{X} \to \mathcal{Y}$ for discrete sets $\mathcal{X}, \mathcal{Y}$. That is, for each value of $x \in \mathcal{X}$ we would like to construct a set of quantum gates $g_1(x), \ldots, g_{\mathcal{C}(x)}(x)$ such that the final measurement result provides the value of $y$. The quantum computational complexity of an algorithm accomplishing this task is the maximum of the number of gates $\mathcal{C}(x)$ required over the input value $x$. Since quantum measurements are inherently probabilistic, one often generalizes this setup to only require the quantum computation to succeed with high probability.

In addition to the relative complexity and the computational complexity, in certain setups we can also define a notion of query complexity. This is introduced in Ch. 13. Intuitively, it counts the number of times that one needs to query some unknown state or process in order to complete a given task.

## 2.2 Topological insulators

In this section we provide a rapid introduction to topological insulators [51]. We begin by introducing two essential concepts in condensed matter physics, band theory and tight-binding models. We then show that gapped band structures can possess topological invariants, which we understand through the mathematical concept of homotopy. Finally, we introduce the notion of stability to our classification, which leads to the seminal tenfold way classification of topological insulators and superconductors.

## Band theory and tight-binding models

For the sake of brevity, we will skip past a continuum approach to band theory and begin directly with tight-binding models. The extension of the below results to continuous real space is underlied by Bloch's theorem, we refer to Ref. [51] for a full treatment.

Consider a system composed of non-interacting fermions in a translation-invariant lattice of potential wells. Suppose the potential wells are deep and the energy of the system is low enough that the fermions can occupy only a few low-lying eigenstates (or *orbitals*) of each potential well. This is known as the tight-binding limit. When the potential wells are infinitely far away from one other, the single-particle Hamiltonian is simply a sum of the individual well Hamiltonians:

$$\hat{H} = \sum_{\boldsymbol{r}} E_\alpha c^\dagger_{\boldsymbol{r},\alpha} c_{\boldsymbol{r},\alpha}, \tag{2.8}$$

where $c^\dagger_{\boldsymbol{r},\alpha}, c_{\boldsymbol{r},\alpha}$ are creation and annihilation operators for a fermion in orbital $\alpha$ of well $\boldsymbol{r}$. Here we index the orbitals by $\alpha = 1, \ldots, M$, and the wells (i.e. lattice sites) by $\boldsymbol{r} = (r_{x_1}, \ldots, r_{x_d}) \in \mathbb{Z}^d$. As the potential wells approach one another, the fermions are increasingly able to tunnel between nearby wells. This introduces matrix elements of the form:

$$t^{\alpha\beta}_{\Delta\boldsymbol{r}} c^\dagger_{\boldsymbol{r}+\Delta\boldsymbol{r},\alpha} c_{\boldsymbol{r},\beta}, \tag{2.9}$$

which allow a fermion in orbital $\beta$ of well $\boldsymbol{r}$ to "hop" to orbital $\alpha$ of well $\boldsymbol{r} + \Delta\boldsymbol{r}$. Combining the two, we find the tight-binding Hamiltonian:

$$\hat{H} = \sum_{\boldsymbol{r},\Delta\boldsymbol{r}} \sum_{\alpha,\beta} t^{\alpha\beta}_{\Delta\boldsymbol{r}} c^\dagger_{\boldsymbol{r}+\Delta\boldsymbol{r},\alpha} c_{\boldsymbol{r},\beta}, \tag{2.10}$$

where we include the individual well Hamiltonians using $t^{\alpha\beta}_{\boldsymbol{0}} = E_\alpha \delta_{\alpha,\beta}$.

We can better analyze the tight-binding Hamiltonian by exploiting translation invariance and Fourier transforming the fermion operators. Namely, we define the fermionic creation and annihilation operators of momentum $\boldsymbol{k}$ via:

$$c^\dagger_{\boldsymbol{k},\alpha} = \sum_{\boldsymbol{r}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} c^\dagger_{\boldsymbol{r},\alpha}, \qquad c_{\boldsymbol{k},\alpha} = \sum_{\boldsymbol{r}} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} c_{\boldsymbol{r},\alpha}. \tag{2.11}$$

Note that the momentum is periodic on a $d$-torus with radius $2\pi$ in each direction, since $\boldsymbol{r}$ is a vector of integers. Inserting these definitions into the tight-binding Hamiltonian gives:

$$\hat{H} = \sum_{\boldsymbol{k}} \sum_{\alpha,\beta} H_{\alpha\beta}(\boldsymbol{k}) c^\dagger_{\boldsymbol{k},\alpha} c_{\boldsymbol{k},\beta}, \tag{2.12}$$

where we have:

$$H_{\alpha\beta}(\boldsymbol{k}) = \sum_{\Delta\boldsymbol{r}} e^{i\boldsymbol{k}\cdot\Delta\boldsymbol{r}} t^{\alpha\beta}_{\Delta\boldsymbol{r}}. \tag{2.13}$$

Diagonalizing the $M \times M$ Hamiltonian matrix $H_{\alpha\beta}(\boldsymbol{k})$,

$$\sum_{\beta} H_{\alpha\beta}(\boldsymbol{k})u_{\beta}^m(\boldsymbol{k}) = \varepsilon_m(\boldsymbol{k})u_{\alpha}^m(\boldsymbol{k}), \tag{2.14}$$

gives $M$ energy eigenvalues $\varepsilon_m(\boldsymbol{k})$ that are continuous functions of the momentum $\boldsymbol{k}$. We refer to the $m^{\text{th}}$ eigenvalue as a "band" of the system. A *band gap* of size $\Delta \geq 0$ between the $m^{\text{th}}$ and $(m+1)^{\text{th}}$ band occurs when:

$$\min_{\boldsymbol{k}}(\varepsilon_{m+1}(\boldsymbol{k})) - \max_{\boldsymbol{k}}(\varepsilon_m(\boldsymbol{k})) = \Delta. \tag{2.15}$$

An insulating state occurs when the chemical potential lies within such a band gap. In this case, at temperatures that are small compared to the gap, the bands below the gap will be fully filled while the bands above the gap will be fully empty. There is thus a finite energy penalty $\sim \Delta$ associated with any excitation to the system. A conducting state occurs when there is no such energy penalty; a small perturbation to the system can thus excite the electrons into a non-zero total momentum state with only a infinitesimal energy penalty. It is remarkable that this simple picture, involving only non-interacting fermions in an idealized discrete lattice, already has the power to explain such a prominent property of materials in the world around us.

## Homotopy invariant of gapped band structures

The above discussion left one aspect of the band structure conspicuously absent: the eigenvectors $u_{\beta}^m(\boldsymbol{k})$. These eigenvectors encode how the orbital index of a fermion lying in the $m^{\text{th}}$ changes with the momentum. (Oftentimes this index is instead formed by a spin or sublattice of this system and goes by the alternate names "spin" or "pseudospin".) If one is in a perfect insulating state with bands either fully filled or fully empty, the eigenvectors contain *all* of the information about the quantum state of the system.

For instance, suppose the chemical potential is zero and there are $m$ bands with energies below zero and $M - m$ bands with energies above. The quantum state of the system corresponds to filling the $m$ bands and leaving the $M - m$ bands empty. We would find the *same* quantum state if we instead considered the "flattened" Hamiltonian:

$$\hat{H}^{\text{flat}}(\boldsymbol{k}) = \sum_{j=m+1}^{M} |u^m\rangle\langle u^m| - \sum_{j=1}^{m} |u^m\rangle\langle u^m| \tag{2.16}$$

which has the same eigenvectors but a flat dispersion of $-1$ for the $m$ filled bands and $+1$ for the $M - m$ empty bands. Here we have changed notation to write the Hamiltonian in first-quantized form in an $M$-dimensional Hilbert space.

Note that the above Hamiltonian is invariant under $U(m)$ unitary rotations within the filled bands and $U(M - m)$ unitary rotations within the empty bands. The quantum state of the topological insulator is therefore characterized by the map:

$$H^{\text{flat}} : T^d \to U(M)/(U(m) \otimes U(M - m)). \tag{2.17}$$

The latter group is known as a Grassmanian manifold. In two-band systems with one filled band, the Grassmanian manifold simply corresponds to the Bloch sphere of the filled band, $U(2)/(U(1) \otimes U(1)) \cong S^2$.

We are interested in universal characteristics of the above quantum state. The notion of homotopy from mathematics provides precisely such a characteristic. Viewing the quantum state as a map from the $d$-torus to the Grassmanian manifold as above, we say that two quantum states are homotopy equivalent if their respective maps $H_1^{\text{flat}}(\boldsymbol{k})$ and $H_2^{\text{flat}}(\boldsymbol{k})$ can be continuously deformed into one another, i.e. one can construct a continuous function $H^{\text{flat}}(\boldsymbol{k}, \lambda)$ such that $H^{\text{flat}}(\boldsymbol{k}, 0) = H_1^{\text{flat}}(\boldsymbol{k})$ and $H^{\text{flat}}(\boldsymbol{k}, 1) = H_2^{\text{flat}}(\boldsymbol{k})$. For a given dimensionality, the homotopy equivalence relation defines a set of equivalence classes. The equivalence class of a given insulator quantum state is a universal property of the state, in the sense that it is robust to any continuous deformation of the Hamiltonian that does not close a bulk band gap (since under such a deformation, the corresponding $H^{\text{flat}}(\boldsymbol{k})$ transforms continuously).

The equivalence classes above define distinct topological insulator phases of matter. In particular, the equivalence class of a product state defines the trivial insulator phase. Any distinct equivalence class defines a non-trivial topological insulator. The number of equivalence classes that we get will depend on the dimensionality and symmetries of the system. (Symmetries can also be taken into account in the above definition by requiring that $H^{\text{flat}}(\boldsymbol{k})$ respect the symmetry.)

To illustrate these ideas, let us introduce the prototypical example of a topological insulator, the so-called Chern insulator. The Chern insulator exists in two dimensions in the absence of any symmetry. We will work in two-band systems for simplicity. In this setting, the quantum state with one band filled is defined by the map $|u\rangle : T^2 \to S^2$, where $|u\rangle\,(\boldsymbol{k})$ is the lower eigenvector of $H(\boldsymbol{k})$. The homotopy classes of this map are characterized by the *Chern invariant*, $C \in \mathbb{Z}$, which is an integer characterizing the "winding number" of the 2-torus about the 2-sphere. The simplest Chern insulator Hamiltonian is the Qi-Wu-Zhang model [380],

$$H(k_x, k_y) = \sin(k_x)\sigma_x + \sin(k_y)\sigma_y + (u + \cos(k_x) + \cos(k_y))\sigma_z. \tag{2.18}$$

The Chern invariant of an arbitrary Hamiltonian can be calculated via $C = \int d\boldsymbol{k}\,\mathcal{F}(\boldsymbol{k})$, where $\mathcal{F}$ is known as the Berry curvature,

$$\mathcal{F}(\boldsymbol{k}) = \langle \partial_{k_x} u(\boldsymbol{k}) | \partial_{k_y} u(\boldsymbol{k}) \rangle - \langle \partial_{k_y} u(\boldsymbol{k}) | \partial_{k_x} u(\boldsymbol{k}) \rangle. \tag{2.19}$$

Before proceeding, we comment on two crucial aspects of topological insulators. First, we note that the chief physical signature of a non-trivial topological insulator is the presence of gapless modes at the boundary between the topological insulator and a trivial state. These can be intuitively understood when the boundary is smooth, i.e. when the Hamiltonian interpolates slowly in real space from the topological insulator state to the trivial state. In this case, the non-trivial homotopy classification indicates that the band gap of the Hamiltonian must close at some point within the interpolation. This gap closing corresponds to a gapless mode on the boundary. The fact that the gapless boundary modes persist at sharp

boundaries can be more rigorously understood from the viewpoint of topological quantum
field theory and mathematical index theorems [343].

Second, we note that this classification is robust to adding interactions as long as the
interaction strength $V$ is weak compared to the band gap $\Delta$. When this is the case, local
perturbations to the quantum state by the interactions are suppressed by $\sim V/\Delta$ and thus
small.

## Tenfold way classification of topological insulators and superconductors

In the homotopy treatment above, the classification of possible topological invariants
implicitly depends on the number of bands in the system. Is this physical? In fact it is
not, at least under the most general assumptions about the robustness of the topological
invariant. For example, consider a two-band insulator with one band filled and one band
empty. Now consider adding a small perturbation to the system that doubles the size of the
unit cell. The Brillouin zone is thus halved, and the original bands fold on top of themselves
to give two bands filled and two bands empty. If we wish for our topological invariants to be
robust to such small perturbations, they clearly cannot depend arbitrarily on the number of
bands.

A stronger condition holds if we allow for taking tensor products of the system of interest
with a trivial topological insulator (e.g. a single band $H_{\mathrm{triv}} = E \sum_i c_i^\dagger c_i^i$ which we can take
to be either filled or empty). Under this procedure one can clearly extend the number of
filled and empty bands of the original system to be any number greater than the original.
If we desire robustness to this procedure, we should seek a topological invariant that is
independent of the number of bands.

The result of this pursuit is the seminal "tenfold way" classification of topological insu-
lators and superconductors [251, 425]. The full classification due to Kitaev utilizes a heavy
dose of mathematical $K$-theory and Clifford algebras [251]. The mathematical setup is a
rigorization of the previous paragraph: we consider an equivalence relation on the set of all
gapped band structures, in which two band structures are equivalent if they can be continu-
ously deformed into one another (without closing the band gap) after allowing for arbitrary
tensor products with trivial bands. The resulting equivalence classes each represent a dis-
tinct topological insulator, and the set of topological insulators possesses a group structure
via the tensor product. The classification depends only on the symmetry and dimensionality
of the system.

Owing to the stringent robustness requirements of the tenfold way classification, we
emphasize that there exist many classes of topological insulators that lie beyond its remit.
Notable examples include topological insulators that are protected by crystalline symmetries
of the system [154], and higher-order topological insulators where gapless surface states are
protected only at "hinges" of the surface [423]. The tenfold way classification also misses
topological insulators that are unstable to the addition of higher bands in the system, but

which are robust in few-band systems. The most notable example of this latter case is the Hopf insulator [335], which we will encounter in Section II. Finally, the introduction of periodic driving to the system leads to the notion of Floquet topological insulators, which we introduce in Ch. 5.

## Quantum information perspective

From our description above, it is far from obvious how topological insulators are connected to quantum information. The most direct connection relates to the notion of relative complexity of quantum states, introduced in the previous chapter. Specifically, consider introducing an equivalence relation on states where two states are equivalent if they have relative complexity $\sim O(1)$ i.e. not scaling with the system size $n$. (To precisely define this, we should properly consider sequences of states as a function of $n$ [524], but we will neglect this subtlety here.) Consider a gate set $\mathcal{G}_{\text{local}}$ consisting of all geometrically local gates. Roughly, we say that the ground state of a local Hamiltonian possesses *topological order* if it is in a different equivalence class than product states, under the gate set $\mathcal{G}_{\text{local}}$. It turns out that (most) topological insulators are trivial under this definition. The reason is that the gate set $\mathcal{G}_{\text{local}}$ does not need to conserve fermion number. Now consider instead a gate set $\mathcal{G}_{\text{local},U(1)}$ consisting of all geometrically local gates that conserve fermion number. It turns out that the topological insulators are each non-trivial under this definition. Since their non-triviality relies on the presence of the fermion-conserving $U(1)$ symmetry, we say that topological insulators are an example of *symmetry-protected topological phase*.

## 2.3 Quantum information dynamics

In this section, we provide an introduction to a few of the developments in quantum information dynamics that have occupied the past decade. The developments that we discuss largely originated in the study of quantum gravity, and center around quantities known as out-of-time-order correlators (OTOCs). The OTOC had been studied previously in a few different contexts (semiclassical physics, single-particle quantum chaos, and nuclear magnetic resonance spectroscopy), yet its behavior and applications in many-body systems were only understood recently. These applications span diverse sectors of physics, including quantum chaos (Ch. 11), thermodynamics, classical simulation of quantum dynamics, quantum communication (Chs. 7, 8, 9), error propagation in quantum simulators (Ch. 10) and quantum gravity.

We begin in this section by motivating the OTOC through the context of classical chaos [277]. We then provide an information-theoretic point of view of the OTOC in the context of many-body systems [506]. From this viewpoint, we introduce the notion of operator size distributions [397], which provide a distinct probe of quantum information dynamics that will be central to the results of Chs. 9 and 10.

## From classical chaos to out-of-time-order correlators

We begin in the setting of classical mechanics, and for concreteness consider a phase space $(x_1, x_2, x_3, p_1, p_2, p_3)$ equipped with Hamiltonian dynamics. If the system starts at position $x_0 = (x_1(0), x_2(0), x_3(0))$ and is evolved by a time $t$, then we denote its new position by $x(t, x_0)$, although we will often suppress the dependence on the initial condition. To ascertain the sensitivity to initial conditions, we can compute the derivative of the first coordinate $x_1(t)$ with respect to, say, the first coordinate of the initial condition $x_1(0)$, giving $\frac{\partial x_1(t)}{\partial x_1(0)}$. For chaotic systems this quantity can exhibit exponential growth in $t$, with a growth rate characterized by a so-called Lyapunov exponent. This is the classical chaos: the behavior of the system at later times is exponentially sensitive to the choice of initial conditions.

To motivate a quantum generalization, we can write $\frac{\partial x_1(t)}{\partial x_1(0)} = \{x_1(t), p_1(0)\}_{\text{PB}}$ where the right-hand side is the Poisson bracket. The quantum version of this quantity is naturally $\frac{1}{i\hbar}[\hat{x}_1(t), \hat{p}_1(0)]$, as was suggested in the seminal work of Larkin and Ovchinnikov [277]. Notice that this object is an operator; since it is convenient to have a single number which captures the exponential growth of chaos, it is natural to take the expectation value of the operator with respect to a state $\rho$ as $\frac{1}{i\hbar}\text{tr}(\rho\,[\hat{x}_1(t), \hat{p}_1(0)])$. Often $\rho$ is chosen to be a thermal state, although in this case the expectation value may fluctuate around zero. To ameliorate this, the expectation value of the *square* of the commutator can be considered, namely $-\frac{1}{\hbar^2}\text{tr}(\rho\,[\hat{x}_1(t), \hat{p}_1(0)]^2)$. Expanding this out, there are terms of the form $-\frac{1}{\hbar^2}\text{tr}(\rho\,\hat{x}_1(t)\,\hat{p}_1(0)\,\hat{x}_1(t)\,\hat{p}_1(0))$, which are indeed out-of-time-order: we start at time zero, evolve to time $t$, evolve back to time zero, and evolve back to time $t$. It is this OTOC term that gives the quantum analog of exponential growth.

More generally, in quantum many-body systems the preferred quantity is often $\text{tr}(\rho\,[W(t), V(0)]^2)$ where $\rho$ is a thermal state and $W(0), V(0)$ are (initially) spatially local operators [250, 315]. Expanding out the commutator, we obtain the OTOC:

$$\text{OTOC}(t, \rho; V, W) = \text{tr}(\rho W(t) V(0) W(t) V(0)). \qquad (2.20)$$

The decay of the OTOC thus quantifies how much $W(t)$ and $V(0)$ fail to commute in the Heisenberg picture. We should note that various re-definitions of the OTOC are possible depending on how one "distributes" the density matrix $\rho$. For instance, in Ch. 9 we will be naturally led to OTOCs of the form $\text{tr}(\rho^{1/2} W(t) V(0) \rho^{1/2} W(t) V(0))$ where $\rho$ is equally distributed at two locations in the trace. More general OTOCs are possible by taking the times in $W$ and $V$ to have imaginary components. At most points in this thesis, we will focus on OTOCs at *infinite temperature*, which corresponds to the density matrix $\rho = \mathbb{1}/d$ where $d$ is the total Hilbert space dimension.

## Out-of-time-order correlators in many-body systems

We begin with a technical aside to introduce the qudit Pauli operators. Consider a collection of $n$ qudits of local Hilbert space dimension $q$ with an on-site Hilbert space $\{|j\rangle \,|\, i = 0, \ldots, q-1\}$. The qudit Pauli operators are unitary operators defined as follows. The "clock"

Pauli operator $Z$ is defined as $Z\,|j\rangle = \exp(i2\pi j/q)\,|j\rangle$. The "shift" Pauli operator $X$ is defined
as $X\,|j\rangle = |j+1\rangle$. The clock and shift operators commute up to a phase, $ZX = e^{i2\pi/q}XZ$.
The total set of $q^2$ Pauli operators are obtained as multiples of the clock and shift operators
of the form $Z^a X^b$ with $a, b = 0, \ldots, q-1$. These form a complete operator basis for a single
qudit.

In systems of many qudits, it is natural to consider Pauli string operators $\hat{R} = \bigotimes_{i=1}^{n} \hat{R}_i$,
corresponding to tensor products of Pauli operators $\hat{R}_i$ on each qudit. The set of Pauli
strings forms a complete operator basis of the many-qudit system.

Consider the OTOC between two (initial) Pauli operators $W, V$. We can decompose the
time-evolved operator as a sum of Pauli strings,

$$W(t) = \sum_R c_R \hat{R}, \tag{2.21}$$

where the coefficients squared are normalized to unity, $1 = tr(W^\dagger W)/d = \sum_R |c_R|^2$. Now
insert this decomposition into the infinite temperature OTOC:

$$
\begin{aligned}
\mathrm{OTOC} &= \frac{1}{d}\mathrm{tr}(W(t)VW^\dagger(t)V^\dagger) \\
&= \frac{1}{d}\sum_{R,R'} c_{R'} c_R^* \mathrm{tr}(R'VR^\dagger V^\dagger) \\
&= \frac{1}{d}\sum_R |c_R|^2 \mathrm{tr}(RVR^\dagger V^\dagger) \\
&= \sum_R |c_R|^2 \cdot e^{i\omega_{RV}},
\end{aligned}
\tag{2.22}
$$

where $\omega_{RV}$ quantifies how much the Pauli operators $V$ and $R$ do not commute, $RV = e^{i\omega_{RV}}VR$. The OTOC is equal to the expectation value of $e^{i\omega_{RV}}$ over the distribution of
Pauli strings in $W(t)$.

Now consider averaging the OTOC over a complete basis of Pauli operators $V$ on a qudit
$j$. We use the identity:

$$\frac{1}{q}\sum_V \hat{V}_j \hat{A} \hat{V}_j^\dagger = \mathrm{tr}_j(\hat{A}) \otimes \mathbb{1}_j \tag{2.23}$$

where $\hat{A}$ is a many-qudit operator. This identity holds whenever $V$ is summed over a complete
single-qudit basis of operators (i.e. a "one-design"). We obtain:

$$
\begin{aligned}
\frac{1}{q^2}\sum_V \mathrm{OTOC}(V,W) &= \frac{1}{q^2}\frac{1}{d}\sum_R |c_R|^2 \sum_V \mathrm{tr}(RV_j R^\dagger V_j^\dagger) \\
&= \sum_R |c_R|^2 \delta_{\hat{R}_j, \mathbb{1}_j}.
\end{aligned}
\tag{2.24}
$$

We thus arrive at the most intuitive interpretation of the OTOC, as quantifying the probabil-
ity that a Pauli string drawn from $W(t)$ does not have support on qudit $j$. This interpretation

only holds precisely when the OTOC is averaged over a complete basis of operators $V$ on qudit $j$. For fixed $V$, it will hold approximately if the Pauli strings in $W(t)$ that have support on $j$ have roughly equal probability to contain each non-identity qudit-$j$ Pauli operator (up to some constant shift due to the fact that $V$ is non-identity).

There is also a nice interpretation of the OTOC as a mutual information between two time-slices of the quantum dynamics [518]. To define this mutual information, we consider an initial state where every qudit is in an EPR pair with a reference system. We then act unitary time dynamics on the first half of these EPR pairs. After the time dynamics are applied, the first half of qudits correspond to the "final state" while the second half of qudits correspond to the "initial state". (To understand this, note that if we project the second half into a many-body state $|\psi^*\rangle$, the first half are in the state $U|\psi\rangle$. Hence our terminology. The complex conjugation is a subtlety related to EPR pairs, see Ch. 9.)

We would like to understand how local information in the initial state is stored in the final state. Specifically, consider partitioning the initial state into a local subsystem A of $n_A$ qudits and its complement B, and partitioning the final state into a local subsystem D of $n_D$ qudits and its complement C. Note that B and C are extensive. We will now show that the average of the OTOC over Pauli operators $W$ on subsystem A and Pauli operators $V$ on subsystem D is equal (up to a constant) to the second order Renyi mutual information between subsystems A and C.

The second order Renyi mutual information is defined as:

$$2^{I^{(2)}(A,C)} = 2^{-S^{(2)}(AC)+S^{(2)}(A)+S^{(2)}(C)} = \frac{\text{tr}(\rho_{AC}^2)}{\text{tr}(\rho_A^2)\text{tr}(\rho_C^2)} = d_A d_C \text{tr}(\rho_{AC}^2). \quad (2.25)$$

We can re-express the trace in terms of an average OTOC by using the identity Eq. (2.23), generalized to arbitrary subsystems. This is most easily seen in the diagrammatic notation that we introduce in Ch. 9. Referring therein for an introduction, here we simply state the result:



$$\text{tr}(\rho_{AC}^2) = \quad = \frac{1}{d^2\, d_A d_D}\sum_{V,W} \quad = \frac{1}{d^2 d_A d_D}\sum_{V,W}\text{tr}\left(\hat{W}(t)\,\hat{V}\,\hat{W}(t)\,\hat{V}\right)$$

$$(2.26)$$

We thus arrive at:

$$I^{(2)}(A, C) = 2n_A + \log\left[\frac{1}{d_A^2 d_D^2}\sum_{W,V}\text{OTOC}(W, V)\right], \tag{2.27}$$

which is our final result.

We can check the limits of this equation as follows. For trivial dynamics the average OTOC is one, thus the logarithm is zero and the mutual information is maximal. This keeps with our expectation: if the system has not scrambled, the initial information contained in A lies entirely in C at later times (since A is in C). On the other hand, consider fully scrambling time dynamics in which every non-identity time-evolved operator $W(t)$ has probability one to have support on every site. In this case, the OTOC averaged over $V$ is zero when $W \neq \mathbb{1}$ and one when $W = \mathbb{1}$. Since $W$ has probability $1/d_A^2$ to be the identity, the average OTOC is $1/d_A^2$. The logarithm thus takes value $-2n_A = \log(1/d_A^2)$ and the mutual information is zero. This again keeps with our expectation: if the system has fully scrambled, the initial information cannot be recovered unless we have access to every site of the system.

## Operator size distributions

In bosonic qudit systems, we define the size of a Pauli string as its number of non-identity elements [397]. For instance, the Pauli string $\mathbb{1} \otimes X \otimes \mathbb{1} \otimes \mathbb{1} \otimes Z \otimes X \otimes \mathbb{1}$ has size 3. A more general operator can be written as a sum of Pauli strings, $R$:

$$Q(t) = \sum_R c_R(t)R, \tag{2.28}$$

and possesses a corresponding *size distribution* [379, 397]:

$$P(\mathcal{S}) = \sum_{\mathcal{S}[R]=\mathcal{S}} |c_R(t)|^2. \tag{2.29}$$

The distribution is normalized to 1 if $Q$ is unitary,

$$\sum_{\mathcal{S}} P(\mathcal{S}) = \sum_R |c_R(t)|^2 = \frac{1}{d}\sum_R |c_R(t)|^2 \text{tr}(R^\dagger R) = \frac{1}{d}\text{tr}(Q^\dagger Q) = 1. \tag{2.30}$$

One can naturally characterize the size distribution via its moments—for instance, the average size, $\overline{\mathcal{S}} \equiv \sum_{\mathcal{S}} P(\mathcal{S})\mathcal{S}$, and the size width, $\delta\mathcal{S} = \overline{\mathcal{S}^2} - (\overline{\mathcal{S}})^2$.

We note that the operator size distribution probes fundamentally distinct properties of quantum information dynamics compared to the OTOC. As a simple proof-of-point example, consider a two-qubit system and the operators $W_1(t) = X \otimes \mathbb{1} + \mathbb{1} \otimes X$ and $W_2(t) = \mathbb{1} \otimes \mathbb{1} + X \otimes X$. The operators $W_1$ and $W_2$ clearly have the same OTOCs as long as $V$ is a single-qudit operator. The operators $W_1$ and $W_2$ also clearly have different size

distributions, $P_1(\mathcal{S}) = \delta_{\mathcal{S},1}$ and $P_2(\mathcal{S}) = (\delta_{\mathcal{S},0} + \delta_{\mathcal{S},2})/2$. Roughly speaking, the OTOC probes the probability that an operator has support on the qudit that $V$ acts on. The size distribution probes correlations in the operators' support across all qudits.

Finally, we note that we can introduce a corresponding size superoperator $\mathcal{S}\{\cdot\}$ defined by its action of Pauli strings $\mathcal{S}\{\hat{R}\} = \mathcal{S}[\hat{R}]\hat{R}$. The size superoperator can be neatly expressed as a sum of single-qudit Pauli operations as follows:

$$\mathcal{S}\{\cdot\} = N(\cdot) - \frac{1}{q^2} \sum_{i=1}^{n} \sum_{P} \hat{P}_i(\cdot)\hat{P}_i^\dagger, \tag{2.31}$$

where $P$ is summed over the single-qudit Pauli operators. This can be verified using the identity $\frac{1}{q^2} \sum_P \hat{P}\hat{Q}\hat{P}^\dagger = \delta_{Q,\mathbb{1}}\hat{Q}$ for Pauli operator $Q$. Writing $\hat{R} = \bigotimes_i \hat{R}_i$ and applying this identity for each $i$, we obtain a sum of $N - \mathcal{S}[\hat{R}]$ copies of $\hat{R}$ from the right term in the above equation. The subtraction from $N$ then gives the size.

## 2.4   Quantum learning theory

The fundamental goal of all physical experiments is to learn properties of a system from measurements on it. Most generally, quantum learning is the study of how best to achieve this goal when some aspect of the process is quantum.

In this thesis, we will focus on the scenario where the unknown system itself is quantum, e.g. either an unknown quantum state or quantum process. We suppose that there is some property of the quantum state that we wish to learn, and we would like to know how well various strategies to learn this property succeed. Here, a strategy consists of both the choice of what measurements to perform on the quantum system, as well as how to process those measurement outcomes. The former aspect is particularly interesting in quantum systems owing to the non-commutativity of physical observables: we cannot learn everything about a quantum system from a single measurement on it. The latter aspect opens up the possibility of quantum advantage in learning, if a quantum computer can assist our ability to process the experimental outcomes. Whether this occurs has yet to be seen, outside of the case of "quantum memories" which we discuss at the end of this Chapter.

The above definition is quite broad, and spans a number of distinct physical scenarios. A common theme is to minimize the experimental resources (e.g. the total time or number of experiments) required for a given learning task. For instance, the field of quantum metrology represents a particular example of quantum learning in which one wants to learn the value of some continuous parameter (e.g. a magnetic field) as precisely as possible, often in as little time as possible [128]. As another example, given a quantum state $\rho$ and a set of observables $\{\hat{O}_i\}$, one can wonder how many experiments are necessary in order to estimate the expectation value of every $\hat{O}_i$ in the state [217]. As a third example, suppose we have a quantum state $\rho(\mathbf{x})$ that depends on some continuous vector of parameters $\mathbf{x}$. We can now ask how many experiments are required to estimate the expectation value of an observable

$\hat{O}$ for all values of $\mathbf{x}$ [218]. Tasks such as these are already crucial in quantum technologies such as quantum sensing and quantum device characterization, and one can imagine they will become more widespread as quantum technology continues to progress. Quantum learning theory seeks to uncover common patterns and themes in the difficulty of, and solutions to, such problems.

In this thesis, we will be predominantly occupied with the class of questions: How much information do experiments of a type A contain about a property B? Clearly, certain types of experiments are much better for learning certain properties than others. For instance, if one wished to distinguish between the two single-qubit quantum states $|0\rangle$ and $|1\rangle$, an experiment that measures the qubit in the $Z$-basis will succeed while an experiment that measures in the $X$-basis will not. Thus, we will be particularly interested in these questions in the context of broad classes of measurement strategies, where each class encompasses many choices of, for example, the measurement basis. We will also focus on learning tasks in *many-body* systems, and analyze the scaling of the number of experiments required with the system size. Specifically, in Chs. 12 and 13, we will establish that certain features of many-body quantum dynamics can be efficiently learned in experiments that involve reversed time-evolution, but are exponentially difficult to learn with experiments that do not.

With this endpoint in mind, in this section we provide a brief introduction to a few relevant tools in quantum learning theory. We begin with an introduction to the Fisher information, which quantifies the amount of information that an entity $B$ contains about a continuous entity $A$ [242]. The Fisher information is ubiquitous throughout (classical) statistics as well as quantum learning theory, and will be utilized in Ch. 12. We then turn to a quite different context, and introduce the setting of a few notable no-go theorems recently proven in learning with a quantum memory [10, 95, 96, 219]. The results in Ch. 13 will borrow much of these techniques, albeit for a quite different learning task.

## Quantifying learnability via the Fisher information

Suppose we have access to a random variable $x$ drawn from a probability distribution $p(x|J)$. Here $J$ is a fixed unknown parameter of the probability distribution. For example, we could image that $x$ is given by some deterministic function of $J$ plus a small amount of Gaussian noise, $p(x|J) = f(J) + \mathcal{N}(0, \sigma)$, where $\mathcal{N}(\mu, \sigma)$ is a Gaussian distribution with mean $\mu$ and standard deviation $\sigma$. In the context of quantum mechanics, we could also imagine that we have a quantum state $\rho(J)$ parameterized by $J$, and would like to learn the value of $J$ from measurements on $\rho(J)$. A measurement of an observable $\hat{X}$ gives eigenvalue $x$ with probability:

$$p(x) = \langle x| \rho(J) |x\rangle. \tag{2.32}$$

The classical Fisher information quantifies how information a single draw of $x$ contains about the parameter $J$. The classical Fisher information is defined as [242]:

$$\mathcal{I}(J) = \mathbb{E}\left[(\partial_J \log(p(x|J)))^2\right] = \sum_x \frac{(\partial_J p(x|J))^2}{p(x|J)}. \tag{2.33}$$

When the probability distribution depends strongly on $J$, the squared derivative will be larger and the Fisher information will be greater. The classical Fisher information derives its power from the associated Cramer-Rao bound, which lower bounds the variance in *any* unbiased estimate $J_{\text{est}}$ of $J$ based on a single draw of $x$. We have:

$$\text{var}(J_{\text{est}}) \geq \frac{1}{\mathcal{I}(J)}. \tag{2.34}$$

A higher Fisher information allows a lower variance in $J_{\text{est}}$, and hence a more precise estimate.

In multi-parameter estimation scenarios, the relevant quantity is the Fisher information matrix [242]:

$$\mathcal{I}_{ij}(J) = \mathbb{E}\left[\left(\partial_{J_i} \log(p(x|J))\right)\left(\partial_{J_j} \log(p(x|J))\right)\right] = \sum_x \frac{\partial_{J_i} p(x|J)\partial_{J_j} p(x|J)}{p(x|J)}. \tag{2.35}$$

The associated Cramer-Rao bound involves the matrix inverse:

$$\text{var}(J_{i,\text{est}}) \geq \left(\mathcal{I}(J)^{-1}\right)_{ii}, \tag{2.36}$$

Notably, the quantity on the RHS can be much greater than the reciprocal single-variable Fisher information $1/\mathcal{I}_{ii}(J)$. This is because uncertainty in the parameters $j \neq i$ can inhibit our ability to estimate the parameter $i$.

Although it will not be used in the results of this thesis, there is also a quantum generalization of the Fisher information, defined for a density matrix $\rho(J)$. The quantum Fisher information matrix takes the form [300]:

$$\mathcal{I}_{ij}(J) = 2 \sum_{\lambda,\lambda'} \frac{\text{Re}\left\{\langle\lambda|\partial_{J_i}\rho|\lambda'\rangle\langle\lambda'|\partial_{J_j}\rho|\lambda\rangle\right\}}{\lambda + \lambda'}. \tag{2.37}$$

Restricted to a single variable, the quantum Fisher information is equal to the maximum of the classical Fisher information over the choice of measurement basis. There is also an associated Cramer-Rao bound for the quantum Fisher information much like the classical Fisher information. In the multi-parameter case, there are subtleties surrounding the achievability of the quantum Cramer-Rao bound due to the non-commutativity of quantum observables.

## Exponential separations in quantum learning

An exciting recent line of progress in quantum learning theory is the development of rigorous proofs regarding when learning is and is not possible. Many recent results in this direction have focused on separation between learning with and without a quantum memory [10, 95, 96, 219]. In particular, it has been shown that there exist properties of quantum systems that are exponentially difficult to learn given only a single copy of the system. However, these same properties become easy to learn (i.e. learnable in either constant or

polynomial time) if one has the ability to perform joint measurements on multiple copies of the system. The additional copies are viewed as being stored from previous experiments in a "quantum memory".

In Ch. 13, we establish a new separation in quantum learning theory, between learning a quantum process with or without access to its time-reverse. In so doing, we will utilize (and indeed, substantially extend) some of the proof techniques developed for learning with quantum memories. We therefore provide here a brief introduction and intuition for learning with a quantum memory. For brevity, we focus on a single, particularly relevant learning task [10, 96].

The learning task we consider is as follows. Suppose one is given oracle access to either (i) an unknown state $\rho = |\psi\rangle\langle\psi|$ drawn uniformly randomly from the set of $n$-qubit states, or (ii) a maximally mixed state $\rho = \mathbb{1}/d$, where $d = 2^n$. In case (i), the same state $|\psi\rangle$ is provided in each subsequent query to the oracle. The goal is to distinguish whether the quantum state is drawn from (i) or (ii), i.e. whether the state is pure or maximally mixed.

This question is easily answered if one can perform joint measurements on two copies of the unknown state. One simply measures the SWAP operator $S$ between the two copies. In case (i), the state $|\psi\rangle \otimes |\psi\rangle$ is an eigenstate of $S$:

$$S(|\psi\rangle \otimes |\psi\rangle) = +(|\psi\rangle \otimes |\psi\rangle), \tag{2.38}$$

and thus returns measurement outcome $+1$ in every instance. In case (ii), the maximally mixed state returns $+1$ and $-1$ with near equal probability in the large $n$ limit (for $n = 1$, it returns $-1$ one quarter of the time, corresponding to a single singlet state and three triplet states; at higher $n$ the ratio of symmetric to asymmetric states approaches 1). Cases (i) and (ii) can thus be distinguished in $O(1)$ queries.

We would like to know whether it is still possible to distinguish the two states if one has access to only a single copy of the unknown quantum state at a time. The result, that it is exponentially difficult in $n$ to distinguish between case (i) and (ii), is established in Refs. [10, 96]. Here we provide intuition for this result by considering a simple setup where one can only measure a single experimental observable in each repetition of the experiment. We then introduce the significantly more general setup considered in Refs. [10, 96].

How can one distinguish a single copy of a pure state from the maximally mixed state? This task would be easy if we knew what basis the pure state was in. For example, if the pure state lied in the computational basis, e.g. $|\psi\rangle = |0011010001\rangle$, then we could simply measure any qubit in the $Z$-basis. If the measurement provided the same outcome every time (i.e. the expectation value obeyed $|\langle Z_i \rangle| = 1$) then we have case (i); if not, then case (ii). The difficulty of distinguishing (i) and (ii) thus lies in our inability to correctly guess the basis of the unknown pure state.

To illustrate this a bit more concretely, suppose we choose to measure some non-identity Pauli observable $\hat{M}$. What is the expectation value of $\hat{M}$ in the given random pure state? Well, the average expectation value is zero, using properties of state one-design:

$$\int d\psi \, \langle\psi| \hat{M} |\psi\rangle = \frac{1}{d}\text{tr}(\hat{M}) = 0, \tag{2.39}$$

since the Pauli operators have zero trace. But this is not too helpful, in principle the average expectation value could be small but its absolute value could be large. We would like to know what is the probability $P_\varepsilon$ that the absolute value of $\langle \psi | \hat{M} | \psi \rangle$ is greater than some constant $\varepsilon$. Let's consider the second moment, evaluated using properties of a state two-design:

$$\int d\psi \, \langle \psi | \hat{M} | \psi \rangle^2 = \frac{1}{d(d+1)} \left[ \text{tr}(\hat{M} \otimes \hat{M}) + \text{tr}((\hat{M} \otimes \hat{M})\hat{S}) \right] = \frac{1}{d(d+1)} \text{tr}(\hat{M}^2) = \frac{1}{d+1}.$$
(2.40)

The second moment is suppressed by Hilbert space dimension $d = 2^n$, and so is exponentially small in the number of qubits. Further, we can use the second moment to upper bound the probability $P_\varepsilon$ via Markov's inequality: $P_\varepsilon \leq 1/(\varepsilon^2 d)$. Therefore our probability of choosing a successful observable $\hat{M}$ is exponentially small, and so (in this restricted setup) it is exponentially difficult to distinguish case (i) and (ii).

The setups considered in Refs. [10, 96] are much broader than this. There are two important generalizations. First, the measurements are allowed to be any POVM. For example, one could imagine measuring all $n$ qubits in the $Z$-basis and obtaining a full bit-string from each measurement result. Second, the entire protocol is allowed to be *adaptive*. That is, the measurement basis in a given round of the protocol is allowed to depend on the measurement results of all previous rounds. Remarkably, even within these near-maximally generous assumptions, it is possible to prove that cases (i) and (ii) cannot be distinguished without exponentially many queries to the unknown quantum state. The proof requires a substantial amount of technical machinery, relating to unitary designs, Weingarten functions, and properties of the permutation group.

In Ch. 13, we will prove an exponential separation in learning quantum processes for the general setup above, i.e. allowing adaptive POVMs. There is actually one additional generalization that we allow: *partial* measurements, in between successive applications of the quantum process of interest. This ingredient was disallowed in previous works; our expanded proof techniques allow us to include it.

# Part II

# Topology in Dynamical Quantum Phenomena

# Chapter 3

# Realizing Hopf insulators in dipolar spin systems

In this Part, we address topological phenomena in *dynamical* quantum systems. As introduced in Ch. 2.2, static topological phases are characterized by ground states that cannot be transformed into product states under any finite-depth unitary circuit (perhaps restricted to obey a symmetry). This leads to a host of interesting static quantum phenomena, such as robust gapless edge modes and excitations with exotic braiding statistics.

Dynamics will play two different roles in this Part. In Chs. 3, 4, and 6, we will use dynamics to enable novel realizations of essentially static quantum phenomena. In Ch. 5, we discuss instead how dynamics (specifically, in the form of periodic driving) can induce entirely new, intrinsically-dynamical phases of matter.

We begin in this Chapter with our first example of how dynamical phenomena—specifically, high-frequency periodic driving—can enable simpler realizations of static topological phenomena. Our specific topic of interest will be the so-called Hopf topological insulator [132, 243, 299, 335]. We recall from Ch. 2.2 that the interplay between a system's symmetries and dimensionality leads to a rich landscape of topological insulators [156, 191, 237, 260, 334, 403, 460, 525], captured by the so-called 'Tenfold Way' classification [251, 425]. The Hopf insulator (HI) represents a topological insulator phase of matter beyond this classification. Existing in three dimensions in the absence of any time-reversal or particle-hole symmetries, the Tenfold Way predicts that no topological state can exist, yet the Hopf Insulator features an insulating bulk and conducting edges protected by an integer-valued $\mathbb{Z}$ topological invariant. The apparent contradiction is avoided in two ways. First, HIs are weak TIs, meaning they only exist in two-band models, and the addition of further non-interacting bands can destroy the topology. Weak TIs are not captured by the Tenfold Way. Second, the $\mathbb{Z}$ topological invariant is not the usual Chern number, but is instead a *linking number* familiar from knot theory and deriving from a relation to the Hopf map [212, 335].

Despite a recent resurgence of interest in HIs [7, 19, 130, 131, 132, 206, 207, 215, 243, 244, 299, 335, 428, 456, 469, 479, 509], fundamental difficulties have led to only a few proposals for their physical implementation [132]. There are three main barriers to implementing HIs
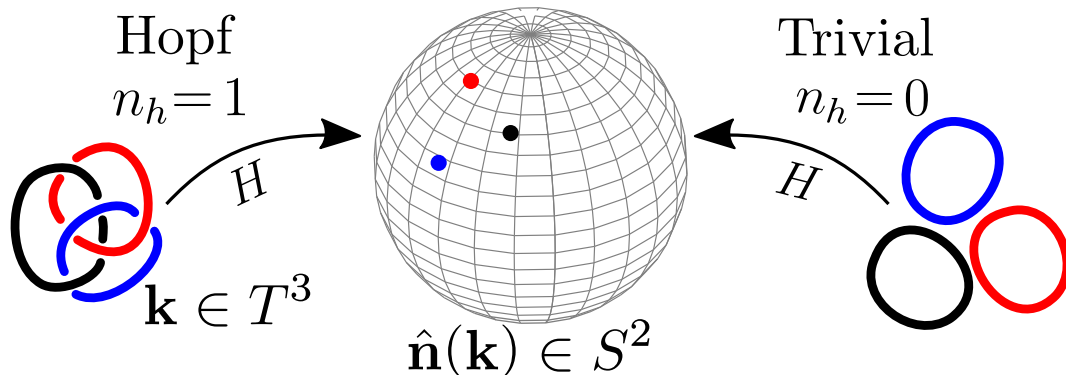
Figure 3.1:   Three-dimensional two-band systems implement maps from the Brillouin zone $T^3$ to the Bloch sphere $S^2$. The pre-image of any point in $S^2$ is a closed loop in $T^3$. There exist topologically non-trivial states, Hopf Insulators (HIs), in which the pre-images of any two points on $S^2$ are *linked* in $T^3$. HIs are characterized by a non-zero Hopf invariant $h$ equaling the linking number of the loops; pictured schematically are three points on $S^2$ and their pre-images in $T^3$ for both a HI with $h = 1$ (left) and a trivial insulator with $h = 0$ (right).

in any tight-binding model of (say) electrons hopping on a lattice. First, the necessity of having precisely two bands rules out many material implementations. Second, as we will see, the nature of the non-zero linking number invariant requires a delicate structure in reciprocal space, meaning real-space interactions must be specified to large distances. Third, a strong spin-orbit coupling is required between the two bands.

In this Chapter, we demonstrate that these barriers may be overcome by implementing HIs in lattices of driven, dipolar-interacting spins. Electron 'hopping' can be replaced by transitions between rotational eigenstates, which are much easier to create and control; the long range of the dipole-dipole interaction then naturally realizes the long-distance hoppings. The two bands can be created from two sub-lattices, and different spin orientations can lead the hoppings to have a complex phase structure able to simulate a strong spin-orbit coupling. Further, we demonstrate that the key experimental signatures of HIs, gapless edge states, are present at any smooth (adiabatic) termination of our model and are robust to all smooth perturbations. This is in support of previous theoretical arguments for topologically-protected gapless modes at smooth boundaries, where translational-invariance and, as a consequence, the two-band picture are preserved. Nevertheless, we show that gapless edge states may persist at judiciously-chosen sharp (non-adiabatic) edges, owing to a crystalline symmetry that stabilizes the Hopf insulator to higher bands. This connects recent work predicting this 'crystalline-symmetric Hopf insulator' with past numerical findings, which similarly observed gapless modes at sharp edges. In Ch. 4, we will provide a specific experimental blueprint for implementing our proposal using ultracold polar molecules of $^{40}K^{87}Rb$. Due to the close

relation between this Chapter and Ch. 4, we postpone a discussion and outlook regarding our results until the end of Ch. 4.

## 3.1  Model Hamiltonian

The Hopf Insulator (HI) can be understood by considering the three-dimensional two-band system at half filling described by the Hamiltonian

$$\hat{H}(\mathbf{k}) = \sum_{i=1}^{3} n_i(\mathbf{k})\hat{\sigma}^i \tag{3.1}$$

with Pauli matrices $\hat{\sigma}^i$. The bulk of the system is assumed gapped, requiring $|\mathbf{n}(\mathbf{k})| > 0$. Eq. (3.1) defines a map from the three-torus $T^3$ — the Brillouin zone in which the wavevectors $\mathbf{k}$ reside—to the Bloch sphere $S^2$ describing the possible states of the two-band system. The pre-image of any point on $S^2$ is then a closed loop in $T^3$. Restricting attention to the case in which the system has zero Chern number across any two-dimensional slice through the three-torus, the linking number of any two of these loops (necessarily an integer) is equal to the Hopf invariant of the map [212]:

$$h = -\frac{1}{4\pi} \int_{\mathrm{BZ}} \mathrm{d}^3\mathbf{k} \sum_{ijk} \epsilon^{ijk} A_i \partial_j A_k \tag{3.2}$$

with $\epsilon_{ijk}$ the Levi-Civita symbol, $\partial_i = \partial/\partial k^i$, and $A_i(\mathbf{k}) = -i\langle u_{\mathbf{k}}|\partial_i|u_{\mathbf{k}}\rangle$ the Berry connection for eigenstate $|u_{\mathbf{k}}\rangle$. Changing $h$ requires the gap to close. The $h = 0$ state, in which all loops are unlinked, is a trivial insulator, and so for $n_i(\mathbf{k})$ such that $h > 0$ the system is in a topologically non-trivial state: the Hopf insulator [334, 335]. This $\mathbb{Z}$ topological invariant is fundamentally distinct from the Chern number appearing in the Tenfold Way. The situation is shown schematically in Fig. 3.1. The HI is a weak TI, meaning that mixing with further non-interacting bands can destroy the topology.

Recently [299], it was realized that if the system obeys a certain crystalline symmetry,

$$\mathcal{J}^{-1}\hat{H}(\mathbf{k})\,\mathcal{J} = -\hat{H}(\mathbf{k})^*, \tag{3.3}$$

where $\mathcal{J}\mathcal{J}^* = -1$, the HI is promoted to a *strong* TI characterized by a $\mathbb{Z}_2$ invariant (an example of a topological crystalline insulator, TCI) [299]. In two-band models at half-filling this symmetry is always present, with any half-filled two-band Hamiltonian obeying Eq. (3.3) with $\mathcal{J} = \hat{\sigma}_y$. In systems with more than two bands it can be viewed as the composition of inversion and particle-hole symmetries.

Our implementation of the Hopf Insulator is based on the following Hamiltonian:

$$\hat{H}_{\mathrm{eff}} = \frac{1}{2} \sum_{\mathbf{v},\mathbf{r}\neq\mathbf{0}} \sum_{\alpha\beta} t_{\mathbf{r}}^{\alpha\beta} \hat{a}_{\mathbf{v}+\mathbf{r},\alpha}^{\dagger} \hat{a}_{\mathbf{v},\beta} + H.c.$$
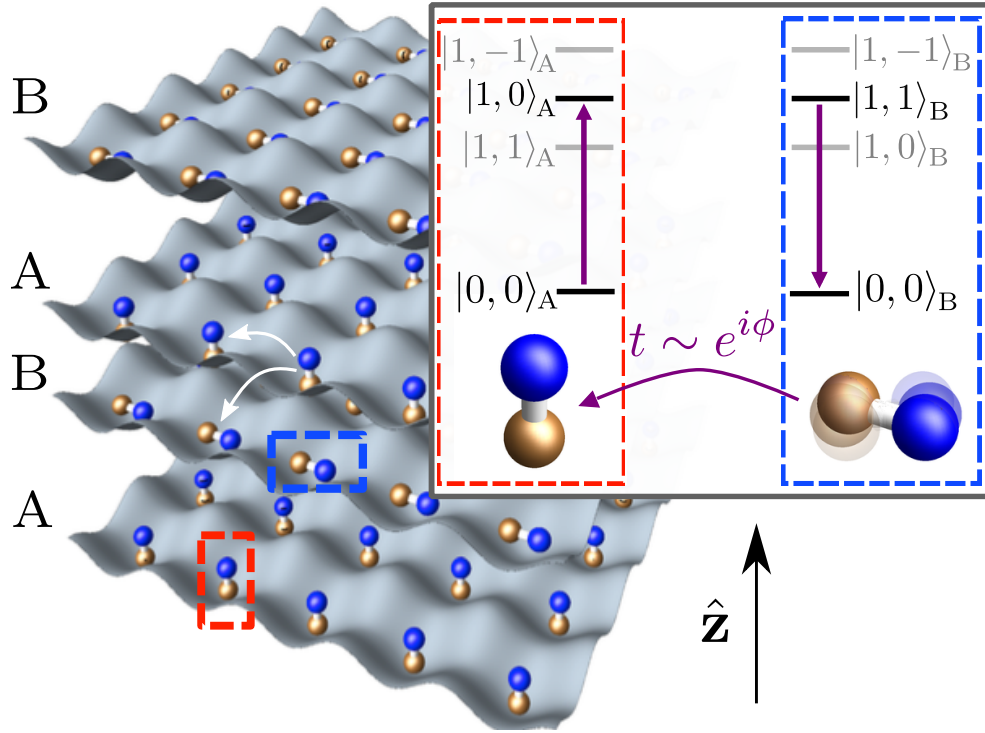
Figure 3.2:   The proposed experimental set-up consists of dipolar molecules confined in a three-dimensional optical lattice, with two sub-lattices $A$ and $B$ separated in the $z$-direction. A combination of applied electric and magnetic fields and the intensities of the lattice beams themselves set the molecules' rotational axes along the $z$-direction, and are tuned so that $|J = 1, m = 0\rangle$ excitations (depicted as $z$-oriented molecules) on the $A$ sub-lattice can 'hop' to $|J = 1, m = 1\rangle$ excitations (depicted as molecules spinning in the $xy$-plane) on the $B$ sub-lattice via the dipolar interaction, while conserving energy. Adding space- and time-dependence to these parameters leads to Floquet modulations $\mu_{\mathbf{v}}^{\alpha}(t)$ of the on-site energies, allowing further control over the hopping magnitudes.

$$+ \sum_{\mathbf{v}} \sum_{\alpha} \mu^{\alpha} \hat{a}^{\dagger}_{\mathbf{v},\alpha} \hat{a}_{\mathbf{v},\alpha} \tag{3.4}$$

where $a^{\dagger}_{\mathbf{v},\alpha}$ creates a hard-core boson at lattice site $\mathbf{v}$ and sub-lattice $\alpha \in \{A, B\}$. We adopt a bosonic rather than fermionic description, permitted by the single-particle nature of the HI, as we will model the hopping of electrons between sites with the exchange of angular momentum eigenstates. The sum over positions $\mathbf{r}$ indicates the presence of long-range hoppings necessary to realize the delicate $k$-space structure needed for all loop pre-images to link. The model has two sub-lattices, which will form the two bands. Both intra- and inter-sub-lattice hoppings are present, $t^{\alpha\beta}_{\mathbf{r}}$, as well as a sub-lattice-dependent chemical potential $\mu^{\alpha}$.

## 3.2  Realizing the Hopf insulator

We propose an implementation of the Hamiltonian of Eq. (3.4), which can naturally be realized in three dimensional lattices of dipolar interacting spins. Our proposal can be realized in a number of experimental platforms, ranging from highly-magnetic neutral atoms such Erbium and Dysprosium [13, 149, 305, 465] to strongly-coupled solid-state spin defects [41, 85, 270] to Rydberg-dressed atom tweezer arrays [52, 285, 523]. Here, we focus on ultra-cold polar molecules trapped in a three-dimensional optical lattice (Fig. 3.2), where tunable strong dipolar interactions have already been experimentally demonstrated [205, 507]. Recent progress has led to the development of numerous molecular species for such set-ups [103, 129, 185, 331, 351, 364, 413, 455, 468]. To demonstrate that this proposal is accessible in near-term experiments, we provide a detailed, quantitative blueprint for its implementation in the specific case of $^{40}K^{87}Rb$ [15, 16, 337, 351, 359, 507] in Ch. 4.

The basic geometry of the setup we envision is a three-dimensional optical lattice generated using four pairs of counter-propagating beams: two pairs forming the $xy$-lattice and two pairs forming the $A$ and $B$ sub-lattices in the $z$-direction (Fig. 3.2). We assume the molecules completely fill the lattice, and each molecule is well-localized to its site by a deep confining potential. Rather than having molecules physically hop between lattice sites, we instead utilize the molecules' rotational degrees of freedom to simulate hard-core bosonic excitations. At lowest order, these rotational states are governed by the Hamiltonian $\hat{H}_{\mathrm{rot}} = \Delta \hat{J}^2$, where $\hat{J}$ is the total angular momentum operator with eigenstates $|J, m_J\rangle$. The energies of these eigenstates are lifted by intrinsic hyperfine interactions, as well as tunable extrinsic effects resulting from applied electric and magnetic fields and incident laser light. These extrinsic effects set the molecules' quantization axes and enable a direct modulation of the rotational energy levels, and hence the two sub-lattices.

Focusing on the four lowest-energy rotational eigenstates, we define two distinct hard-core bosonic degrees of freedom. On the $A$-sub-lattice we utilize $\{|0_A\rangle = |0, 0\rangle_A, |1_A\rangle = |1, 0\rangle_A\}$, while on the $B$-sub-lattice we utilize $\{|0_B\rangle = |0, 0\rangle_B, |1_B\rangle = |1, 1\rangle_B\}$, as illustrated in Fig. 3.2. These hard-core bosons interact with one another via a dipolar interaction, which gives rise to the effective hoppings:

$$t_{\mathbf{r}}^{AA} = -C^{AA} \frac{3\cos^2(\theta) - 1}{R^3}$$

$$t_{\mathbf{r}}^{BB} = C^{BB} \frac{3\cos^2(\theta) - 1}{R^3}$$

$$t_{\mathbf{r}}^{AB} = \left(t_{-\mathbf{r}}^{BA}\right)^* = -C^{AB} \frac{\cos(\theta)\sin(\theta)}{R^3} e^{i\phi} \tag{3.5}$$

where $\{R, \theta, \phi\}$ defines the separation of molecules in spherical polar co-ordinates, and $C^{AA}, C^{BB}$, and $C^{AB}$ are positive constants. Details are provided in Ch. 4. This particular choice of rotational states ensures that the inter-sub-lattice hopping $t_{\mathbf{r}}^{AB}$ is induced solely by the $\Delta m_J = -1$ term, which immediately gives rise to a hopping phase $\propto e^{i\phi}$[1]. This choice, motivated by the model of Ref. [334], locks the intra-sub-lattice components of the Hamiltonian $n_{x,y}(\mathbf{k})$ to the momenta $k_x, k_y$. As illustrated in Ch. 4, this locking naturally achieves the Hopf requirement that all Bloch sphere pre-images link.

We further enhance the relative strength of next-nearest neighbor hopping with a simple Floquet engineering strategy. The basic principle is that, by periodically modulating the on-site chemical potentials $\mu_{\mathbf{v}}^{\alpha}(t)$ inhomogeneously at frequencies, $\hbar\Omega$, significantly higher than the energy of the dipolar interaction, the time-averaged behavior emulates that of a different time-independent Hamiltonian. In this effective Hamiltonian, sites that oscillate out-of-phase with one other will have the hopping between them suppressed, while hoppings between sites oscillating in-phase remain unaffected. Although the Floquet modulation $\mu_{\mathbf{v}}^{\alpha}(t)$ necessarily varies with the lattice site $\mathbf{v}$, we choose it such that effective hoppings remain translationally-invariant. Specifically, we take the Floquet modulation to be a checkerboard pattern in the $xy$-plane, such that next-nearest-neighbor hoppings (even $r_x + r_y$, in-phase) are enhanced relative to nearest-neighbour hoppings (odd $r_x + r_y$, out-of-phase).

Additionally, although the slow decay of the $1/R^3$ dipole-dipole interaction is helpful in establishing the next-nearest neighbor interactions in the $xy$-plane that are necessary to realize the HI, our numerical studies indicate that the same interactions cause unnecessary long-range couplings in the $z$-direction. To address this, we utilize an additional, second patterning of the previous Floquet engineering strategy, which truncates the dipolar interaction to effectively nearest-neighbor in the $z$-direction [281]. This patterning is guaranteed to operate independently of the previous Floquet engineering patterning if their modulation frequencies are well-separated in scale; we verify this quantitatively in Ch. 4.

By this process we are able to identify parameters in Eq. (3.4) that realize the Hopf insulating phase, $h = 1$, with band gaps as large as $E_g \gtrsim 0.26\, t_{\text{nn}}$ (in units of the nearest-neighbor hopping), as well as gapless transitions between the Hopf and trivial insulating phases (see Ch. 4). Utilizing the $\Delta m_J = +1$ component of the dipolar interaction (as opposed to the $\Delta m_J = -1$ component) leads instead to the phase $h = -1$; higher linking numbers are in principle possible, but require an even more delicate structure in $k$-space.

---

[1]Intriguingly, we note that this exact same hopping in *two*-dimensions can give rise to Chern insulating physics, both in the prototypical Qi-Wu-Zhang model for the Chern insulator [380] as well as in positionally disordered systems [9].
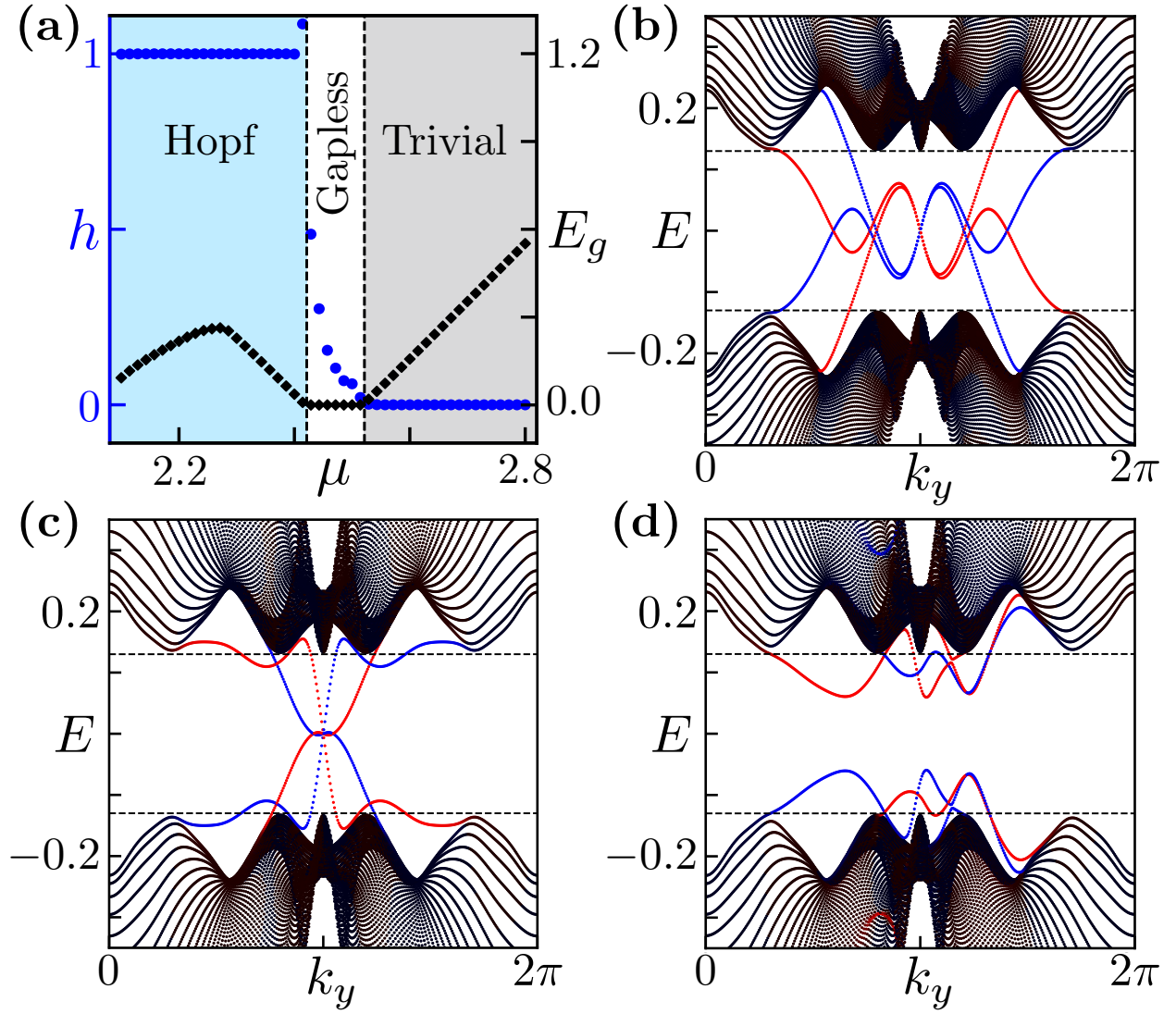
Figure 3.3:   **(a)** Hopf invariant $h$ (left axis) and band gap $E_g$ (right axis) as a function of the staggered chemical potential $\mu = (\mu^A - \mu^B)/2$, found by discretizing the Floquet engineered dipolar Hamiltonian using $70 \times 70 \times 70$ $k$-points, periodic boundary conditions, and setting the nearest-neighbor inter-sublattice hopping in the $xy$-plane to 1. The remaining plots show the energy spectrum with (100) edge terminations. Black states indicate the bulk, red/blue indicate states localized to the left/right edge respectively, and dashed lines the bulk band gap. **(b)** Adiabatic edge termination over 20 sites. The conducting edge states are protected by the $h = 1$ topological invariant. **(c)** Sharp edge termination respecting the $\mathcal{J}$ crystalline symmetry [Eq. (3.3)]. The edge states are now protected by the symmetry. **(d)** Introducing terms that break the $\mathcal{J}$ symmetry gaps the edge states.

Before concluding, we detail the separation of scales required for Eqs. (3.4) and (3.5) to govern the low-energy dynamics of the polar molecular system. First, we work in the natural experimental regime where the dipolar interaction strength is significantly smaller than the energy splittings between the rotational states within the $J = 1$ manifold. The external fields should be tuned such that the splitting between the $|0_A\rangle$ and $|1_A\rangle$ states is resonant with the $|0_B\rangle$ and $|1_B\rangle$ states, and far detuned from all other rotational transitions. Conservation of energy then dictates that the dipolar interaction can only induce transitions within our prescribed hard-core bosonic doublets. Details on how this level scheme can be precisely realized in the specific case of polar molecular quantum simulation based upon $^{40}K^{87}Rb$ can be found in Ch. 4. Here, we note only that the orientation of the spins is fixed via applied fields oriented in the $z$-direction, and that the degeneracy between the $|1, 0\rangle$ and $|1, 1\rangle$ states, as well as the sub-lattice symmetry between the $A$ and $B$ planes, is broken by using *different* intensities of light to form each sub-lattice. Our scheme naturally leads to a separation of energy scales $t \ll \delta \ll \Delta$, where $t$ is the dipolar interaction strength ($\sim 100$ Hz), $\delta$ is the splitting within the $J = 1$ manifold ($\sim 5$ kHz), and $\Delta$ is the splitting between the $J = 0$ and $J = 1$ sectors ($\sim 2$ GHz).

## 3.3 Band structure and edge physics

In Fig. 3.3 we show the band structure found by exact diagonalization of the dipolar Hamiltonian [Eqs. (3.4) and (3.5)] after applying our Floquet engineering strategy (see Ch. 4). We assume periodic boundary conditions in the $y$- and $z$-directions (crystal momenta $k_y$ and $k_z$ are therefore good quantum numbers) but a finite length in the $x$-direction. We also truncate the hopping range to $|\mathbf{r}| \leq 8$ sites for numerical feasibility; increasing the truncation range does not qualitatively affect the results. Fig. 3.3(b) shows the result of a smooth adiabatic termination over twenty lattice sites. The bulk (black) is gapped, but the edges (red and blue) host conducting states, which we found to be stable for any sizable bulk gap. This is the Hopf insulator: the adiabatic termination approximately preserves translational-invariance, leading to the survival of the two-band picture. Since the Hopf invariant is trivial outside the system and unity inside, gapless edge states result at the interface.

Fig. 3.3(c) shows the result of an abrupt termination of the edge. Lacking adiabaticity, the band picture is expected to break down; since the HI exists only for two-band models, we would then not expect topologically-protected edge states. Remarkably, however, edge states are again present. In fact, a serendipitous choice of edge-termination plane (100) has lead to the bulk $\mathcal{J}$ symmetry surviving at the edge, and these edge states are a manifestation of the resulting strong $\mathbb{Z}_2$ invariant (which does not require a two-band model). To see this 'accidental' symmetry, note that open boundary conditions are equivalent to an infinite potential barrier $\hat{H}_{\text{edge}} = \rho\hat{\sigma}^z\delta_x$, $\rho \to \infty$ at the system's edge, where $\hat{\sigma}_z$ acts on the sub-lattice degrees of freedom. In momentum space, this corresponds to real couplings between different $k_x$, $\hat{H}_{\text{edge}}^{\mathbf{k},\mathbf{k}'} = \rho\hat{\sigma}^z\delta_{k_y,k_y'}\delta_{k_z,k_z'}$, which is easily seen to obey Eq. (3.3). Nearly any perturbation to

naive open boundary conditions – for instance a small potential $\gamma\,\hat{\sigma}^z\delta_{x-1}$ on the site nearest the edge – breaks the $\mathcal{J}$ symmetry and gaps the edge states [Fig. 3.3(d)]. We predict that this same mechanism is responsible for stabilizing the edge states at sharp boundaries observed in previous numerical studies of the HI [131, 335]. All of these edge mode structures can be probed experimentally via molecular gas microscopy [117, 321] by exciting individual edge spins and observing the extent to which the excitation remains localized on the edge (see Ch. 4 for further details).

# Chapter 4

# Floquet engineering ultracold polar molecules to simulate topological insulators

In Ch. 3, we demonstrated that the band structure of the Hopf topological insulator could be naturally realized in lattice of driven dipolar-interacting spins. In this Chapter, we present a quantitative near-term experimental blueprint for realizing this proposal in lattice-trapped ultracold polar molecules. Our proposal includes novel methods to (i) realize maximally strong spin-orbit coupling, and (ii) precisely modulate the interactions' spatial profile, which we hope may be of broader use in polar molecule quantum simulation.

The rich internal structure of ultracold polar molecules has led to intense interest for their use in a wide range of applications, ranging from quantum simulation and computation, to ultracold chemistry and precision measurement [34, 42, 103, 129, 185, 221, 267, 331, 351, 352, 364, 412, 413, 421, 455, 468]. Understanding and controlling this structure has led to the development of a host of techniques enabling the preparation and manipulation of rovibrational states in polar molecules [15, 16, 26, 125, 337, 338, 349, 351, 359, 507]. From the perspective of quantum simulation, polar molecules enjoy a unique advantage compared to their neutral atom cousins, owing to the presence of strong, anisotropic, long-range dipolar interactions; these interactions have proven useful for theoretical proposals aiming to realize a number of exotic phases, including disordered quantum magnets [170, 204, 511], Weyl semimetals [452] and fractional Chern insulators [511, 513]. Motivated, in part, by these prospects, the last decade has seen tremendous experimental progress, advancing from rovibrational ground state cooling [351] to the recent realization of a Fermi degenerate molecular gas [125]. Moreover, from a geometric perspective, molecules can either be loaded into optical lattices [337] or optical tweezer arrays [26]. As in other quantum simulation platforms, Floquet engineering [79, 281]—high-frequency, periodic time-modulation—can further sculpt the molecules' interaction, broadening the scope of accessible phases [280, 327].

In this Chapter, we leverage these ingredients to provide an explicit experimental blueprint

for realizing the Hopf insulator in polar molecules. As described in Ch. 3, realizing the Hopf insulator requires two essential ingredients: 1) the presence of long-range hoppings and 2) strong spin-orbit coupling, manifested in hoppings whose phase is spatially anisotropic. In Ch. 3, we predict that combining the dipolar interaction with Floquet engineering [79, 281] can naturally give rise to the Hopf insulator in interacting spin systems. In the present Chapter, we build upon this result by providing a quantitative blueprint using lattice-trapped ultracold polar molecules, focusing for concreteness on $^{40}$K$^{87}$Rb [15, 16, 337, 351, 359, 507].

Our approach takes advantage of the full toolset of controls developed for polar molecular systems. In particular, we envision a deep, three-dimensional optical lattice, so that the molecules' rotational motion constitutes the fundamental degrees of freedom in the system. Rotational excitations are exchanged between lattice sites via the dipolar interaction, which simulates the hopping of hardcore bosons on the lattice. The two band, or 'spin', degrees of freedom of the Hopf insulator are formed from two sublattices, distinguished from each other by the lattice light itself—different intensity light forming the two sublattices induces different level structures in the trapped molecules, according to the molecules' polarizability [349]. In contrast to prior studies [368, 511, 513], we utilize this polarizability to isolate the $\Delta m = \pm 1$ angular-momentum-changing component of the dipolar interaction, which *precisely* induces the requisite spin-orbit coupling of the Hopf insulator [335]. To complete our construction, we demonstrate that Floquet engineering can be implemented using amplitudes of applied laser light and DC electric fields which are easily accessible in current generation experiments; moreover, we show that this engineering can tune the system's hoppings into the Hopf insulating phase with large band gaps $\gtrsim 0.26\, t_{\mathrm{nn}}$ (in units of the nearest-neighbor hopping, $t_{\mathrm{nn}}$), enabling easier experimental observation. Finally, a particularly simple way to achieve the requisite rotational level structure (Fig. 4.1) is to utilize circularly-polarized optical radiation in conjunction of the molecule's AC polarizability. To this end, in order to demonstrate quantitative feasibility, we provide the first detailed calculations of the relevant circular polarizabilities for $^{40}$K$^{87}$Rb.

Direct experimental verification of the Hopf insulator is most simply achieved through spectroscopy of its gapless edge modes. In Ch. 3, we demonstrate that these edge modes are robust at any smooth boundary of the Hopf insulating phase, while for sharp boundaries their presence or absence signifies the existence of an underlying crystalline symmetry [299]. We will show that all three of these qualitatively distinct boundary spectra can be manufactured and probed in ultracold polar molecule simulations. Since the Hopf insulator's edge behavior is a direct result of it being outside the conventional tenfold way, this serves as a direct experimental probe of the Hopf insulator's unique topological classification.

The Chapter is organized as follows. We begin with an overview of the Hopf insulator, with a specific focus on the requirements—a two band system, and long-range, spin-orbit coupled hoppings. We then turn to the setting of our proposal, outlining precisely how the rotational excitations of polar molecules can simulate spin-orbit coupled particles hopping on a lattice. Next, we demonstrate how particular patterns of Floquet driving can provide tremendous control over these hoppings, and numerically verify that these can be used to tune the system into a large band-gap, Hopf insulator phase. We present the edge modes

of the polar molecular Hopf Hamiltonian, and show that they display three qualitatively
distinct spectra dependent on the lattice termination. Finally, we conclude by providing a
detailed description of all aspects of the proposal's implementation in a three dimensional
optical lattice of $^{40}\text{K}^{87}\text{Rb}$.

## 4.1 Physical requirements for realizing Hopf insulators

We begin with an introduction to the Hopf insulator, seeking to motivate the connection
between the linking number interpretation of the Hopf invariant and the long-range spin-orbit
coupling required for its physical realization.

The Hopf insulator is a particular type of topological insulator [156, 191, 237, 260, 334,
403, 460, 525], a class of phases of matter most notable for exhibiting conducting surface
states despite an insulating bulk. They are differentiated from conventional insulators by
a non-zero topological invariant associated with their underlying spin-orbit-coupled band
structure; moreover, their surface states are unusually robust to the detrimental effects of
impurities. Their organization was first captured via the so-called ten-fold way classifica-
tion [251, 425], and consists of a wide landscape of phases dependent on a system's dimen-
sionality and symmetries. Nevertheless, more recent work has exposed topological insulators
that exist beyond this classification framework; notable examples include topological crys-
talline insulators [154], higher-order topological insulators [423], and our insulator of interest,
the Hopf insulator [130, 131, 132, 243, 299, 335, 520].

The Hopf insulator exists in three-dimensions in the *absence* of any symmetries, for
which the ten-fold way classification [251, 425] would nominally predict only an ordinary
insulator. In our context, it will consist of hard-core boson degrees of freedom hopping on a
three-dimensional lattice (although one is accustomed to thinking of topological insulators in
terms of fermions, their single-particle nature also enables a hard-core bosonic realization).
The bosons come in two 'pseudospins', $A$ and $B$, which will form the two bands of the system.
These may be formed from physical spins, but are not required to be—in our realization,
they will correspond to two sublattices of the three-dimensional lattice. In real space, the
Hopf insulator Hamiltonian takes the generic form

$$H_{\text{eff}} = \frac{1}{2} \sum_{\substack{\mathbf{v}, \mathbf{r} \neq \mathbf{0}, \\ \alpha, \beta}} \left[ t_{\mathbf{r}}^{\alpha\beta} a_{\mathbf{v}+\mathbf{r},\alpha}^{\dagger} a_{\mathbf{v},\beta} + h.c. \right] + \sum_{\mathbf{v},\alpha} \mu^{\alpha} a_{\mathbf{v},\alpha}^{\dagger} a_{\mathbf{v},\alpha}, \tag{4.1}$$

where $a_{\mathbf{v},\alpha}^{\dagger}$ is the creation operator for a hard-core boson at lattice site $\mathbf{v}$ of pseudospin
$\alpha \in \{A, B\}$. The Hamiltonian consists of both pseudospin-preserving ($t_{\mathbf{r}}^{AA}$ and $t_{\mathbf{r}}^{BB}$) and
pseudospin-flipping ($t_{\mathbf{r}}^{AB}$ and $t_{\mathbf{r}}^{BA}$) hoppings, as well as a pseudospin-dependent chemical
potential $\mu^{\alpha}$.

The topology of the Hopf insulator is most easily seen in its momentum-space representa-
tion, governed by the two-by-two matrix $H^{\alpha\beta}(\mathbf{k}) = \sum_{\mathbf{r}} \tilde{t}_{\mathbf{r}}^{\alpha\beta} e^{i\mathbf{k}\cdot\mathbf{r}} + \mu^{\alpha} \delta^{\alpha\beta}$. This is conveniently
decomposed as $H(\mathbf{k}) = n_0(\mathbf{k})\mathbb{1} + \mathbf{n}(\mathbf{k}) \cdot \boldsymbol{\sigma}$, where the Pauli matrices $\boldsymbol{\sigma}$ act on the pseudospin
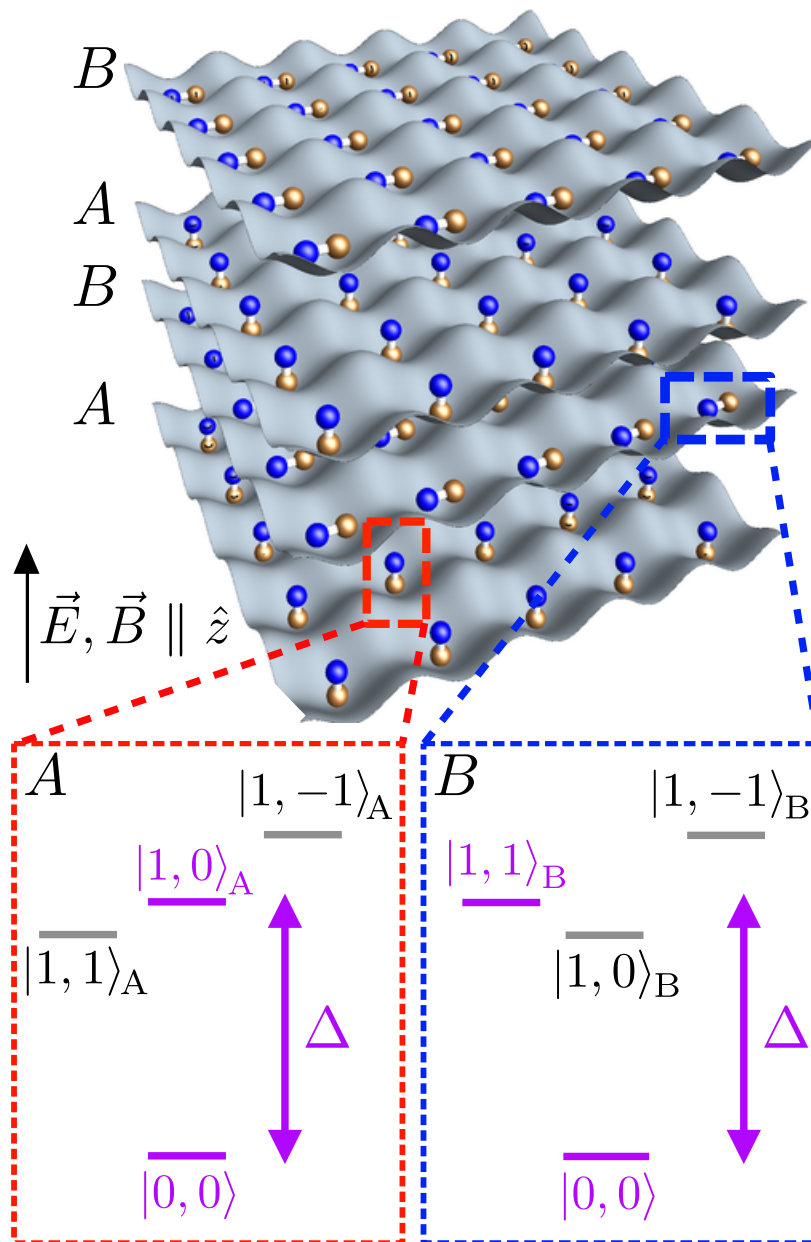
Figure 4.1: Schematic geometry depicting a 3D optical lattice of polar molecules with two layered sublattices $A$ and $B$. Orbital motion of the molecules is frozen by the optical lattice. The level structure of the $J = 0, 1$ rotational states on the $A$ (left) and $B$ (right) sublattices. The purple highlighted states form the hard-core bosonic doublet for each sublattice, and their energy splitting $\Delta$ is tuned by external fields to be degenerate between sublattices.
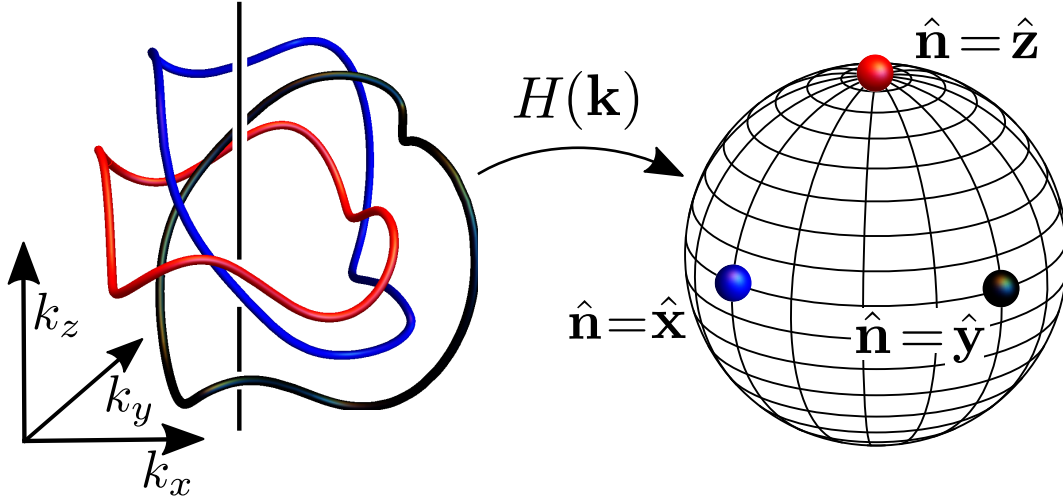
Figure 4.2: The Hamiltonian of the Hopf insulator maps closed loops in the Brillouin zone
to points on the Bloch sphere, with the unique property that any two loops have linking
number equal to the Hopf invariant. The above loops are solved for using the proposed
experimental dipolar spin Hamiltonian specified in the text – their linking provides a visual
verification of the Hopf insulating phase. The $\hat{\mathbf{n}} = \hat{\mathbf{x}}, \hat{\mathbf{y}}$ pre-images (blue, black tubes) are 90°
rotations of each other about the $k_z$-axis (vertical black line) due to the spin-orbit coupled
hopping $t_{\mathbf{r}}^{AB} \sim e^{i\phi}$.

degrees of freedom, which form the two bands of the Hopf insulator, and the condition that
the bands are gapped requires $|\mathbf{n}(\mathbf{k})| > 0$. We can view this Hamiltonian as a map that
takes vectors $\mathbf{k}$ in the Brillouin zone to points $\hat{\mathbf{n}} \equiv \mathbf{n}/|\mathbf{n}|$ on the Bloch sphere. To see the
Hopf insulator's topology, consider the *pre-images* of two different Bloch sphere points $\hat{\mathbf{n}}, \hat{\mathbf{n}}'$
in the Brillouin zone, i.e. the set of momenta $\mathbf{k}$ such that $\hat{\mathbf{n}}(\mathbf{k}) = \hat{\mathbf{n}}$, or $\hat{\mathbf{n}}(\mathbf{k}) = \hat{\mathbf{n}}'$. Since
the Brillouin zone is three-dimensional—one dimension higher than the Bloch sphere—these
pre-images are generically 1D loops in the Brillouin zone. The Hopf invariant $h$ of the
Hamiltonian $H(\mathbf{k})$ is precisely equal to the *linking number* of these two loops, for any choice
of $\hat{\mathbf{n}}, \hat{\mathbf{n}}'$ [Fig. 4.2(a)]. The invariant can be calculated from the Bloch Hamiltonian via the
Chern-Simons form [335]:

$$h = \int_{BZ} d^3\mathbf{k}\ j^\mu(\mathbf{k})A_\mu(\mathbf{k}), \tag{4.2}$$

where $j^\mu(\mathbf{k}) = \frac{1}{8\pi}\epsilon^{\mu\nu\lambda}\hat{\mathbf{n}} \cdot (\partial_{k_\nu}\hat{\mathbf{n}} \times \partial_{k_\lambda}\hat{\mathbf{n}})$ is the Berry curvature and $A_\mu(\mathbf{k})$ its associated vector
potential.

The linking number interpretation leads to two observations, one which explains the
need for long-range hoppings and the other which justifies the required form of spin-orbit
coupling. First, the rapid variation in $\mathbf{n}(\mathbf{k})$ required for pre-image linking necessitates the
presence of strong long-range hoppings, which contribute oscillations $\sim e^{i\mathbf{k}\cdot\mathbf{r}}$ to $\mathbf{n}(\mathbf{k})$, at a

frequency proportional to their range $\mathbf{r}$. Specifically, no nearest neighbor Hamiltonian is
known for the Hopf insulator; the prototypical Hopf insulator Hamiltonian [335] features as
far as next-next-nearest neighbor hoppings. Second, pre-image linking by definition requires
a strong coupling between the pseudospin degree of freedom and the momentum, much as
is true for other topological insulators. Inspired by the model of Ref. [335], in this work we
realize a specific form of this spin-orbit coupling, generated via pseudospin-flipping hoppings
with a direction-dependent phase $t_{\mathbf{r}}^{AB} \sim e^{i\phi}$, where $\phi$ is the azimuthal angle of the hopping
displacement $\mathbf{r}$ (Fig. 4.3). This form of hopping locks the $n_x, n_y$ components of pseudospin to
the $k_x, k_y$ components of the momentum, such that the pre-image of, e.g. $\hat{\mathbf{n}} = \hat{\mathbf{x}}$, is exactly a
90 degree rotation about the $k_z$-axis of the pre-image of $\hat{\mathbf{n}} = \hat{\mathbf{y}}$. As illustrated in Fig. 4.2, this
simple correspondence leads naturally to linking of the two pre-images. While this simple
argument applies only to pre-images related by 90 or 180 degree rotations about the $z$-axis
(due to the cubic lattice symmetry), this is in fact sufficient: in a gapped model, the linking
number is *constant* for all pairs of pre-images. We note that this same phase profile of the
hoppings is also present in two-dimensional realizations of Chern insulating physics, both in
the prototypical Qi-Wu-Zhang model [380] as well as in positionally disordered systems [9]

In the following two sections, we demonstrate that systems of dipolar interacting spins
provide a natural ground to realize both of these key ingredients. We begin by describing
how a particular configuration of the spins' level structures leads to the effective hard-core
boson Hamiltonian of Eq. (4.1), including the desired spin-orbit coupling $t_{\mathbf{r}}^{AB} \sim e^{i\phi}$. We
then augment the bare dipolar hoppings with a Floquet engineering scheme, which serves to
decrease the relative strength of nearest-neighbor hoppings and provides useful experimental
parameters for tuning into the Hopf insulating phase.

## 4.2   The Dipolar Hamiltonian

We now turn to the setting of our proposal. We envision a three-dimensional optical
lattice filled with ultracold polar molecules. We work in the deep lattice limit, so that the
molecules themselves do not hop between lattice sites, and the molecules' rotational states
form the fundamental degrees of freedom of our system [508]. As shown in Fig. 4.1, the
lattice is formed by alternating planes of two-dimensional square lattices, stacked in the
$z$-direction. These form two sublattices, $A$ and $B$, which will play the role of the pseudospin
in the Hopf insulator.

The molecules are most strongly governed by the rotational Hamiltonian $H_{\text{rot}} = \Delta \mathbf{J}^2$,
with eigenstates $|J, m_J\rangle$ indexed by their orbital ($J$) and magnetic ($m_J$) angular momentum
quantum numbers, which have energies $E = \Delta J(J + 1)$ and wavefunctions described by
the spherical harmonic functions[1]. While naturally organized into degenerate manifolds of
each $J$, the $m_J$ eigenstates are split by both intrinsic hyperfine interactions and *tunable*
extrinsic effects resulting from electric fields, magnetic fields and incident laser light. These

---

[1]In the presence of external fields, which slightly mix the rotational eigenstates, we use $|J, m_J\rangle$ to refer
to the state adiabatically connected to the associated rotational eigenstate.

extrinsic effects (which set the molecules' quantization axis, i.e. $\hat{\mathbf{z}}$ in Fig. 4.1) enable a direct
modulation of the rotational states' energies in both *space* (to distinguish between the $A$
and $B$ sublattices) and *time* (to implement Floquet engineering).

We now aim to use these rotational states to realize an effective Hamiltonian of hard-core
bosons, as in Eq. (4.1). We focus on the lowest four rotational eigenstates (i.e. the $J = 0, 1$
manifolds), and use these to define two distinct hard-core bosonic degrees of freedom. On
the $A$-sublattice we form a hard-core boson from the doublet $\{|0_A\rangle = |0, 0\rangle_A, |1_A\rangle = |1, 0\rangle_A\}$,
while on the $B$-sublattice we utilize $\{|0_B\rangle = |0, 0\rangle_B, |1_B\rangle = |1, 1\rangle_B\}$, as illustrated in Fig. 4.1.
The hard-core bosons interact with each other through the dipolar interaction [72]:

$$H_{\mathrm{dd}}^{ij} = \frac{-\sqrt{6}}{4\pi\epsilon_0 r^3} \sum_{\Delta m_J = -2}^{2} C_{-\Delta m_J}^2(\theta, \phi) T_{\Delta m_J}^2(\mathbf{d}^{(i)}, \mathbf{d}^{(j)}), \tag{4.3}$$

where $(r, \theta, \phi)$ parameterizes the separation of the interacting molecules $i$ and $j$ in spheri-
cal coordinates, and we compress unit and sublattice indices into a single index $i = \mathbf{v}, \alpha$.
The dipole moment operator $\mathbf{d}^{(i)} = (d_-^{(i)}, d_z^{(i)} d_+^{(i)})$ is a rank-1 spherical tensor acting on the
rotational states of the molecule $i$, whose three components change the molecule's mag-
netic quantum number by $(-1, 0, +1)$ respectively. The spherical harmonics $C_{-\Delta m}^2(\theta, \phi)$
capture the spatial dependence of the interaction, and are accompanied by the correspond-
ing component of $T_{\Delta m}^2$, the unique rank-2 spherical tensor generated from the dipole op-
erators $\mathbf{d}^{(i)}$, $\mathbf{d}^{(j)}$. Explicitly, we have $T_{\pm 2}^2 = d_\pm^{(i)} d_\pm^{(j)}$, $T_{\pm 1}^2 = (d_\pm^{(i)} d_z^{(j)} + d_z^{(i)} d_\pm^{(j)})/\sqrt{2}$, $T_0^2 = (d_\pm^{(i)} d_\mp^{(j)} + 2d_z^{(i)} d_z^{(j)} + d_\mp^{(i)} d_\pm^{(j)})/\sqrt{6}$.

A few remarks are in order. First, we will assume that the dipolar interaction strength
is significantly weaker than the energy splittings within the $J = 1$ manifold. Second, we
will tune the splitting between the $|0_A\rangle$ and $|1_A\rangle$ states to be resonant with that of the $|0_B\rangle$
and $|1_B\rangle$ states (Fig. 4.1). Conservation of energy then dictates that the dipolar interaction
can only induce transitions within our prescribed hard-core bosonic doublets, i.e. those that
preserve boson number. These transitions take the form of hoppings in the bosonic Hamil-
tonian, $t_{ij} = \langle 0_i, 1_j | H_{\mathrm{dd}}^{ij} | 1_i, 0_j \rangle$. These hoppings may occur either within a sublattice ($t_{\mathbf{r}}^{AA}$
and $t_{\mathbf{r}}^{BB}$) or across sublattices ($t_{\mathbf{r}}^{AB}$). With the prescribed geometry and level structure, we
have:

$$\begin{aligned}
t_{\mathbf{r}}^{AA} &= -\frac{d_{00}^2}{4\pi\epsilon_0} \frac{3\cos^2(\theta) - 1}{r^3} \\
t_{\mathbf{r}}^{BB} &= \frac{d_{01}^2}{8\pi\epsilon_0} \frac{3\cos^2(\theta) - 1}{r^3} \\
t_{\mathbf{r}}^{AB} = (t_{-\mathbf{r}}^{BA})^* &= -\frac{3\,d_{00}d_{01}}{4\sqrt{2}\pi\epsilon_0} \frac{\cos(\theta)\sin(\theta)}{r^3} e^{i\phi},
\end{aligned} \tag{4.4}$$

where $(r, \theta, \phi)$ parameterizes the displacement between sites in spherical coordinates, equal to
$\mathbf{r}$ for intra-sublattice hoppings and $\mathbf{r} + b\hat{\mathbf{z}}$ for inter-sublattice hoppings (where $b$ is the vertical
distance between $A$ and $B$ planes), and $d_{00}$, $d_{01}$ are the dipole moments $d_{00} = \langle 1, 0 | d_z | 0, 0 \rangle$
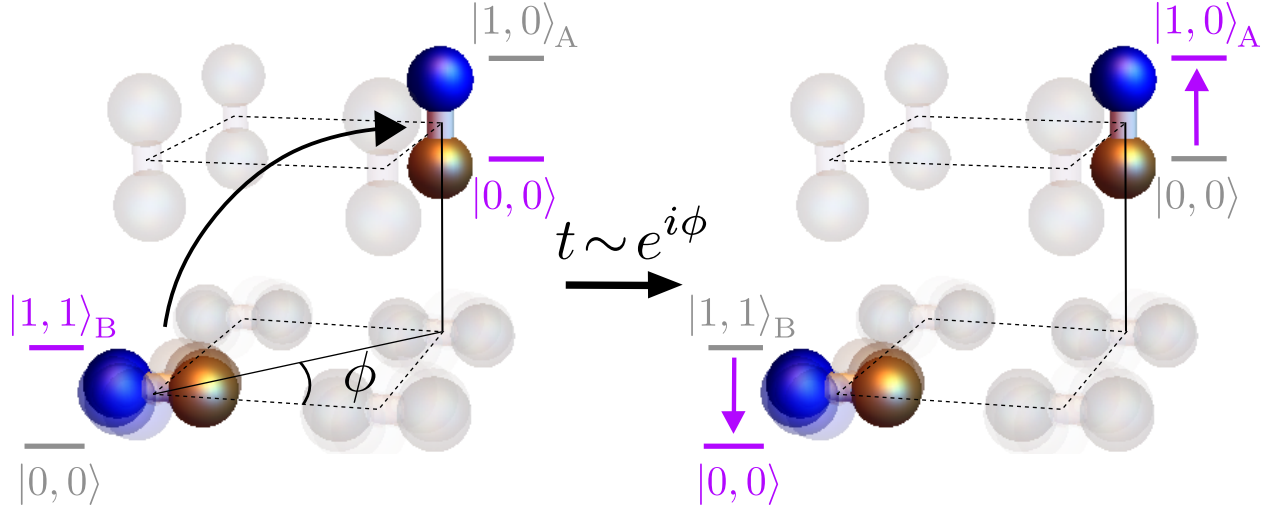
Figure 4.3:   Depiction of the inter-sublattice 'hopping' $|0,0\rangle_A |1,1\rangle_B \to |1,0\rangle_A |0,0\rangle_B$, in which a hard-core bosonic excitation on sublattice $B$ hops to sublattice $A$. This is induced by the dipolar interaction, and occurs with a hopping matrix element $t_{\mathbf{r}}^{AB} \sim e^{i\phi}$ with phase equal to the azimuthal angle $\phi$ between the dipoles. This phase profile arises from the spherical harmonic $C^2_{-\Delta m=1}(\theta,\phi)$, since the hopping changes the total angular momentum of the system by $\Delta m = -1$. Sublattice B molecules are depicted as spinning to indicate their non-zero $z$-angular momentum in the excited state.

and $d_{01} = \langle 1,\pm 1| \, d_{\pm} \, |0,0\rangle$. Our choice of rotational states guarantees that the inter-sublattice hopping, $t_{\mathbf{r}}^{AB}$, arises solely from the $\Delta m_J = +1$ term in $H_{\mathrm{dd}}$, which gives rise to a hopping phase $t_{\mathbf{r}}^{AB} \sim e^{i\phi}$ via the $C^2_{-1}$ spherical harmonic. As illustrated in Fig. 4.2, this naturally leads to linking between the Bloch sphere pre-images. Finally, variations in the energy splitting between sublattices naturally appear as effective chemical potentials $\mu^{\alpha}$, completing the realization of the Hamiltonian Eq. (4.1).

## 4.3   Floquet engineering

While the dipolar interaction elegantly realizes the requisite spin-orbit coupling, relatively strong nearest neighbor hopping as well as the slow asymptotic decay of the $1/R^3$ power-law preclude numerical observation of Hopf insulating behavior. To this end, we utilize Floquet engineering to two effects: first, to decrease 'odd' hoppings in the $xy$-plane (those with odd $r_x + r_y$) and second, to truncate the dipolar power-law in the $z$-direction [281]. We achieve each effect by adding spatio-temporal dependence to the chemical potential $\mu_{\mathbf{v}}^{\alpha}(t)$, and oscillating each $\mu_{\mathbf{v}}^{\alpha}(t)$ at timescales much faster than the hopping. Under certain conditions (specified below), this leads to an effective time-$independent$ Hamiltonian of the

same form as Eq. (4.1), but with modified hoppings

$$t_{\mathbf{r}}^{\alpha\beta} \to \beta_{\mathbf{r}}^{\alpha\beta} \, t_{\mathbf{r}}^{\alpha\beta}, \tag{4.5}$$

where the damping coefficients, $\beta_{\mathbf{r}}^{\alpha\beta}$, are determined by the specific profiles of the oscillated chemical potentials, $\mu_{\mathbf{v}}^{\alpha}(t)$. In what follows, we first derive this relation explicitly [Eq. (4.12)], and then introduce two Floquet engineering schemes [i.e. explicit profiles for the spatio-temporal dependence of $\mu_{\mathbf{v}}^{\alpha}(t)$] that achieve the hopping modifications described above.

## Overview of Floquet engineering

We begin with a broad introduction to Floquet engineering using a time-dependent chemical potential, following Ref. [281] but modified to include sublattices and complex hoppings. We consider a time-dependent Hamiltonian of the form Eq. (4.1) where the chemical potential $\mu_{\mathbf{v}}^{\alpha}(t)$ now varies with the lattice site $\mathbf{v}$ as well as periodically in time $t$, with a period $T$. To calculate the effect of the driving, we move into a rotating frame, defining the unitary

$$U(t) = \exp\left[ -i \int_0^t dt' \frac{1}{2} \sum_{\mathbf{v},\alpha} \mu_{\mathbf{v}}^{\alpha}(t) \sigma_{\mathbf{r}}^z \right], \tag{4.6}$$

and the rotated wavefunction

$$|\Psi'(t)\rangle = U^{\dagger}(t) |\Psi(t)\rangle, \tag{4.7}$$

whose time-evolution is governed by the Hamiltonian

$$\begin{aligned}
H'(t) =& U^{\dagger}(t)H(t)U(t) - iU^{\dagger}(t)\dot{U}(t) \\
=& \frac{1}{2} \sum_{\substack{\mathbf{v},\mathbf{r}\neq\mathbf{0}, \\ \alpha,\beta}} \left( \exp\left[ -i \int_0^t dt' \, (\mu_{\mathbf{v}+\mathbf{r}}^{\alpha}(t') - \mu_{\mathbf{v}}^{\beta}(t')) \right] \times \right. \\
& \left. t_{\mathbf{r}}^{\alpha\beta} \, a_{\mathbf{v}+\mathbf{r},\alpha}^{\dagger} a_{\mathbf{v},\beta} + h.c. \right).
\end{aligned} \tag{4.8}$$

At high-frequencies, $1/T \gg |t_{\mathbf{r}}^{\alpha\beta}|$, the rotated Hamiltonian is well-approximated by replacing all quantities by their average over a single period. This gives an effective time-independent Hamiltonian

$$\begin{aligned}
H^{\text{eff}} =& \frac{1}{2} \sum_{\substack{\mathbf{v},\mathbf{r}\neq\mathbf{0}, \\ \alpha,\beta}} \left[ \beta_{\mathbf{v}+\mathbf{r},\mathbf{v}}^{\alpha\beta} \, t_{\mathbf{r}}^{\alpha\beta} \, a_{\mathbf{v}+\mathbf{r},\alpha}^{\dagger} a_{\mathbf{v},\beta} + h.c. \right] \\
& + \sum_{\mathbf{v},\alpha} \mu_{\mathbf{v}}^{\alpha} \, a_{\mathbf{v},\alpha}^{\dagger} a_{\mathbf{v},\alpha},
\end{aligned} \tag{4.9}$$

with a static chemical potential

$$\mu_{\mathbf{v}}^{\alpha} = \frac{1}{T} \int_0^T dt \, \mu_{\mathbf{v}}^{\alpha}(t), \tag{4.10}$$

and hoppings suppressed by the dampings

$$
\begin{aligned}
\beta^{\alpha\beta}_{\mathbf{v+r,v}} =& \frac{1}{T} \int_0^T dt \, \exp\left[ -i \int_0^t dt' \, (\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\beta_{\mathbf{v}}(t')) \right] \\
=& \frac{1}{T} \int_0^T dt \left( \cos\left[ \int_0^t dt' \, (\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\beta_{\mathbf{v}}(t')) \right] \right. \\
& \left. - i \sin\left[ \int_0^t dt' \, (\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\beta_{\mathbf{v}}(t')) \right] \right).
\end{aligned}
\tag{4.11}
$$

For convenience, we will always choose $\mu^\alpha_{\mathbf{v}}(t)$ to be an even function of $t$, in which case the imaginary part of the damping vanishes and we have

$$
\beta^{\alpha\beta}_{\mathbf{v+r,v}} = \frac{1}{T} \int_0^T dt \, \cos\left[ \int_0^t dt' \, (\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\beta_{\mathbf{v}}(t')) \right].
\tag{4.12}
$$

In this case, the dampings modulate only the hoppings' magnitudes, and not their phase.

Since the modulation is generically inhomogeneous, care must be taken to ensure that the dampings are in fact translation invariant, $\beta^{\alpha\beta}_{\mathbf{v+r,v}} = \beta^{\alpha\beta}_{\mathbf{r}}$, if one desires translation invariance in the effective Hamiltonian. This constraint requires that $\cos\left[ \int_0^t dt' \, (\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\beta_{\mathbf{v}}(t')) \right]$ be independent of $\mathbf{v}$. For intra-sublattice hoppings ($\alpha = \beta$), there are two ways to achieve this: 1) with a 'gradient' modulation, where $\mu^\alpha_{\mathbf{v}}(t)$ is linear in $\mathbf{v}$, and 2) with an 'even-odd' modulation $\mu^\alpha_{\mathbf{v}}(t) = \mu^\alpha(-1)^{s_x v_x + s_y v_y + s_z v_z}$, $s_i \in \{0,1\}$. (The latter is possible because we restrict to the cosine term of Eq. (4.11), which is even in $\mu$ and thus requires only the absolute value of $\mu^\alpha_{\mathbf{v+r}}(t') - \mu^\alpha_{\mathbf{v}}(t')$ to be independent of $\mathbf{v}$.) For inter-sublattice hoppings ($\alpha \neq \beta$), this constraint additionally requires that the sublattices' modulations differ only by a *position-independent* function of time, namely

$$
\begin{aligned}
\mu^A_{\mathbf{v}}(t) &= \mu_{\mathbf{v}}(t) \\
\mu^B_{\mathbf{v}}(t) &= \mu_{\mathbf{v}}(t) + \mu_{SL}(t).
\end{aligned}
\tag{4.13}
$$

These lead to damping coefficients

$$
\begin{aligned}
\beta^{AA}_{\mathbf{r}} = \beta^{BB}_{\mathbf{r}} &= \frac{1}{T} \int_0^T dt \, \cos\left[ \int_0^t dt' \, (\mu_{\mathbf{v+r}}(t') - \mu_{\mathbf{v}}(t')) \right] \\
\beta^{AB}_{\mathbf{r}} &= \frac{1}{T} \int_0^T dt \, \cos\left[ \int_0^t dt' \, (\mu_{\mathbf{v+r}}(t') - \mu_{\mathbf{v}}(t') - \mu^{SL}(t')) \right]
\end{aligned}
\tag{4.14}
$$

for the intra- and inter-sublattice hoppings. We must also ensure that $\mu^\alpha_{\mathbf{v}}$ is translation invariant, which requires only that the average modulation is the same in each unit cell $\mathbf{v}$.

## Even-odd modulation in $xy$-plane

The first scheme for Floquet engineering serves to suppress the strength of nearest neighbor hoppings relative to next nearest neighbor hoppings in the $xy$-plane. The modulation

takes the form of the even-odd modulation previously mentioned, with $s_x = s_y = 1, s_z = 0$.
Specifically, we take

$$\mu_{\mathbf{v}}^{xy}(t) = \frac{1}{2}(-1)^{v_x+v_y}\Omega^{xy}g^{xy}\cos(\Omega^{xy}t)$$
$$\mu_{SL}^{xy}(t) = \Omega^{xy}g_{SL}^{xy}\cos(\Omega^{xy}t),$$

(4.15)

where frequency $\Omega^{xy}$ is $2\pi$ times the inverse period, and $g^{xy}$ and $g_{SL}^{xy}$ are parameters to
be tuned. Performing the integral inside Eq. (4.14) and using $\frac{1}{T}\int_0^T dt \cos\left[g\sin(2\pi t/T)\right] = \frac{1}{2\pi}\int_0^{2\pi} dx \cos\left[g\sin(x)\right] = J_0(g)$ gives damping coefficients

$$\beta_{\mathbf{r}}^{xy,AA} = \begin{cases} J_0(g^{xy}) & r_x + r_y = \text{odd} \\ 1 & r_x + r_y = \text{even} \end{cases}$$
$$\beta_{\mathbf{r}}^{xy,AB} = \begin{cases} J_0(g^{xy} + g_{SL}^{xy}) & r_x + r_y = \text{odd} \\ J_0(g_{SL}^{xy}) & r_x + r_y = \text{even}, \end{cases}$$

(4.16)

where $J_0(g)$ is a Bessel function of the first kind. We see that 'odd' distance hoppings
(including nearest neighbor, $r_x + r_y = 1$) are reduced relative to 'even' hoppings (including
next-nearest $[r_x = r_y = 1]$ and next-next-nearest neighbor $[r_x = 2, r_y = 0$, and vice versa.]
hoppings, both with $r_x + r_y = 2$). The parameters $g^{xy}$ and $g_{SL}^{xy}$ give independent control over
the ratio of even to odd hoppings for both inter- and intra-sublattice hoppings.

## Truncation in the $z$-direction

The second use of Floquet modulation is to truncate hoppings from power law to short
ranged in the $z$-direction [281]. Unlike the previous $xy$-modulation, we do not have an intu-
itive explanation for why one needs such a truncation. Nevertheless, we observe numerically
that it is necessary for realizing the Hopf insulator phase. We take $\mu_{\mathbf{v}}^z(t)$ to be a gradient in
the $z$-direction,

$$\mu_{\mathbf{v}}^z(t) = v_z\Omega^z g^z(\Omega^z t)$$
$$\mu_{SL}^z(t) = \Omega^z g_{SL}^z(\Omega^z t)$$

(4.17)

with frequency $\Omega^z$ in time. This gives dampings

$$\beta_{\mathbf{r}}^{z,AA} = \frac{1}{2\pi}\int_0^{2\pi} dx \cos\left[r_z \int_0^x dx' \, g^z(x')\right]$$
$$\beta_{\mathbf{r}}^{z,AB} = \frac{1}{2\pi}\int_0^{2\pi} dx \cos\left[\int_0^x dx' \left[r_z g^z(x') - g_{SL}^z(x')\right]\right].$$

(4.18)

These can be evaluated numerically once the functions $g^z(\Omega_z t), g_{SL}^z(\Omega_z t)$ are chosen. Ref. [281]
showed that the modulation can be tuned to give hoppings that are effectively nearest-
neighbor in the $z$-direction, at the cost of some loss of magnitude of the nearest-neighbor
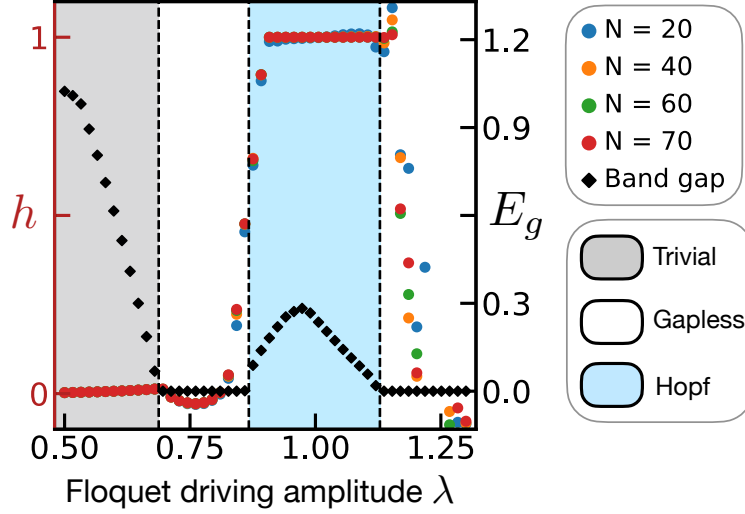
Figure 4.4:   Numerical evaluation of the Hopf invariant $h$ for $N \times N \times N$ discretizations of momentum space (colored circles) and the band gap $E_g$ (black diamonds; in units of the nearest-neighbor hopping $t_{\mathrm{nn}}$) of the specified dipolar spin system as a function of the strength $\lambda$ of the Floquet driving (defined in the main text), calculated with hopping range $R = 8$. The Hopf insulating phase (blue, right shaded) is observed across a range of $\lambda$; outside this range the system transitions to gapless (white) and trivial insulating (gray, left shaded) phases.

($|r_z| = 1$) hopping.  For experimental simplicity, we take the modulations to be piecewise constant in time:

$$
\begin{aligned}
\frac{\mu^z_{A,\mathbf{v}}(\Omega_z t)}{\Omega_z} &=
\begin{cases}
(2g_1 + 2)v_z & 0 < \Omega_z t \leq \phi_1/2 \\
(2g_2 + 2)v_z & \phi_1 < \Omega_z t \leq \phi_2/2 \\
(2g_3 + 2)v_z & \phi_2 < \Omega_z t \leq \pi/2 \\
-\mu^z_{A,\mathbf{v}}(\pi - \Omega_z t) + 4v_z & \pi/2 < \Omega_z t \leq \pi \\
-\mu^z_{A,\mathbf{v}}(2\pi - \Omega_z t) & \pi < \Omega_z t \leq 2\pi
\end{cases} \\[2mm]
\frac{\mu^z_{B,\mathbf{v}}(\Omega_z t)}{\Omega_z} &=
\begin{cases}
2g_1 v_z + 2g_1^{SL} & 0 < \Omega_z t \leq \phi_1/2 \\
2g_2 v_z + 2g_2^{SL} & \phi_1 < \Omega_z t \leq \phi_2/2 \\
2g_3 v_z + 2g_3^{SL} & \phi_2 < \Omega_z t \leq \pi/2 \\
-\mu^z_{B,\mathbf{v}}(\pi - \Omega_z t) & \pi/2 < \Omega_z t \leq \pi \\
-\mu^z_{B,\mathbf{v}}(2\pi - \Omega_z t) & \pi < \Omega_z t \leq 2\pi.
\end{cases}
\end{aligned}
\tag{4.19}
$$

Note that $\mu^z_{A,\mathbf{v}}(\Omega_z t)$ is even about $\pi$, guaranteeing that damping coefficients are real-valued [see Eq. (4.12)]. The parameters $g_i, g_i^{SL}$ can be tuned to achieve the desired hopping truncation.

## Combining the two modulations

We now show that both of the above schemes can be implemented simultaneously, by choosing the frequencies of each to be well-separated. Specifically, we take the modulation to be the sum of two components,

$$\mu_{\mathbf{r}}^{\alpha}(t) = \mu_{\mathbf{r}}^{xy,\alpha}(t) + \mu_{\mathbf{r}}^{z,\alpha}(t), \tag{4.20}$$

where $\mu_{\mathbf{r}}^{xy,\alpha}(t)$ is periodic with frequency $\Omega^{xy}$ and $\mu_{\mathbf{r}}^{z,\alpha}(t)$ with frequency $\Omega^z$, and the frequencies satisfy either $\Omega^{xy} \gg \Omega^z$ or $\Omega^{xy} \ll \Omega^z$. Under this assumption, the dampings $\beta_{\mathbf{v}}^{\alpha\beta}$ factorize into a product of the two individual damping coefficients defined in Eqs. (4.16) and (4.18),

$$\beta_{\mathbf{v}}^{\alpha\beta} = \beta_{\mathbf{v}}^{xy,\alpha\beta}\, \beta_{\mathbf{v}}^{z,\alpha\beta}, \tag{4.21}$$

as desired. We verify that this assumption holds quantitatively in Fig. 4.6.

# 4.4   Numerical verification of the Hopf insulating phase

We now turn to a numerical exploration of the single particle bandstructures supported in our dipolar Floquet system. By tuning the geometric and Floquet engineering parameters, we find transitions from topologically trivial bandstructures to the Hopf insulator and identify parameter regimes where the Hopf insulator's band gap can be as large as $E_g \gtrsim 0.26 t_{\mathrm{nn}}$ (see Figs. 4.4, 4.5). This occurs with a spacing $a = 0.99$ between adjacent planes of the same sublattice in the $z$-direction (in units of the nearest-neighbor spacing in the $xy$-plane), a spacing $b = 0.66$ between adjacent planes of the opposite sublattice, a staggered chemical potential $\mu^A - \mu^B = 2.28$ (in units of the nearest-neighbor hopping in the $xy$-plane), and Floquet engineering parameters $g^{xy} = 1.2$, $g_{SL}^{xy} = 0.1$, $g_1 = -0.6$, $g_2 = 0.1$, $g_3 = 1.1$, $g_1^{SL} = 0.7$, $g_2^{SL} = -0.4$, $g_3^{SL} = 1.6$, $\phi_1 = 0.2$, $\phi_2 = 1.8$. These optimal parameters were found to optimize the Hopf insulating band gap via the basin-hopping optimization algorithm, a stochastic optimization algorithm that works well in rugged, high-dimensional optimization landscapes. [474, 477]. It consists of alternating steps of $i$) performing local optimization to find a nearby local minima in the nearby energy landscape (i.e. the 'basin'), and $ii$) randomized 'hopping' to more distant basins, whose local minima are then computed by repeating the first step. The Floquet engineering amplitudes are quite robust and can be varied together (replacing $g \to \lambda g$ for all amplitudes defined above) by $\sim 25\%$ about their optimal values while preserving Hopf insulating behavior (Fig. 4.4). The staggered chemical potential can be varied by $\sim 20\%$ (see Ch. 3). Performing similar calculations for the lattice parameters, we find that the intra-sublattice distance is also relatively robust and can be varied between $0.5 - 0.9$ (Fig. 4.5), while the $z$-lattice spacing is slightly more sensitive, and should be kept within $0.92 - 1.08$ in units of the $x/y$-lattice spacing (note that the most natural value, 1, lies well-within this range).

We compute the momentum-space Bloch Hamiltonian by summing the Floquet engineered dipolar hoppings defined in Eqs. (4.4, 4.14, 4.18). To truncate the infinite sum over
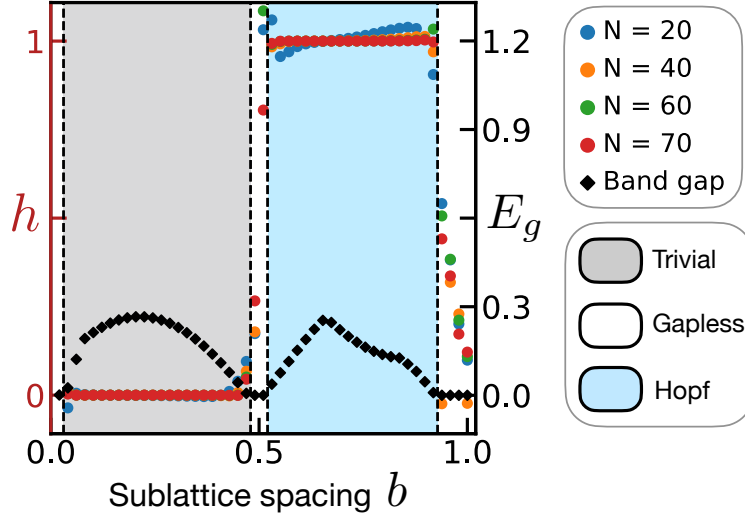
Figure 4.5:    Numerical evaluation of the Hopf invariant $h$ for $N \times N \times N$ discretizations
of momentum space (colored circles) and the band gap $E_g$ (black diamonds; in units of
the nearest-neighbor hopping $t_{nn}$) of the specified dipolar spin system as a function of the
vertical spacing $b$ between sublattices (in units of the nearest-neighbor spacing in the $xy$-
plane), calculated with hopping range $R = 8$. The Hopf insulating phase (blue, right shaded)
is observed across a large range of $b$; outside this range the system transitions to gapless
(white) and trivial insulating (gray, left shaded) phases. Note that the Floquet modulation
breaks the geometric symmetry $b \to 1 - b$, and hence the spectrum is not symmetric about
$b = 0.5$.

hopping distance, we only including hoppings to sites at most $R$ unit cells away in each
direction, i.e. $|r_\mu| \leq R$ for each $\mu = \{x, y, z\}$. The Hopf invariant is computed by evaluating
the integral Eq. (4.2) on an $N \times N \times N$ grid in momentum space, solving $\nabla \times \mathbf{A(k)} = \mathbf{j(k)}$
in the inverse Fourier domain to obtain the Berry connection [335]. The computed invari-
ant converges quickly to 1 as the discretization $N$ becomes large, e.g. $h - 1 \approx 10^{-6}$ at
$N = 70, R = 4$ (see also Figs. 4.5, 4.4). We also see quick convergence of the band gap when
increasing $R$, observing quantitative agreement within 10% for all $4 \leq R \leq 32$ and within
1% for all $8 \leq R \leq 32$.

## 4.5   Edge modes of the dipolar Floquet Hopf insulator

In addition to its linking number invariant, the Hopf insulator's edge modes represent
one of its key signatures, and crucially, one which can be experimentally probed. Up to now,
these edge modes are only expected to appear at boundaries that are smooth at the scale
of the lattice length, which act as a continuous variation of the two-band momentum-space

Hamiltonian $H(\mathbf{k})$ across the boundary region. In this case, the Hopf insulator's nontrivial
homotopy classification *requires* a gap closing in any edge between the Hopf insulator and
the trivial insulator. Nevertheless, gapless edge modes have been observed numerically for
'sharp' boundaries (i.e. open boundary conditions) [335] and moreover, for the (001)-edge,
were even shown to be robust to certain perturbations [131].

Meanwhile, recent work [299] has shown that the Hopf insulator's classification can be
stabilized to higher bands by a certain crystalline symmetry,

$$JH(\mathbf{k})^*J^{-1} = -H(\mathbf{k}), \qquad (4.22)$$

where $JJ^* = -1$, although its classification is reduced to a $\mathbb{Z}_2$ invariant for band number
greater than 2. This symmetry is in fact automatically satisfied in translationally-invariant
two band systems (taking $J = \sigma_y$), and can generally be viewed as the composition of
inversion and particle-hole symmetries.

Interestingly, we observe that—despite involving inversion symmetry—this crystalline
symmetry is also obeyed at the *edge* of a two-band system, in the specific case of a sharp
boundary (open boundary conditions). To see this, note that open boundary conditions
are equivalent to an infinite delta function potential barrier at the edge of the system,
$H_{\text{edge}} = \rho\sigma_z\delta_z$, $\rho \to \infty$, where $\sigma_z$ acts on the sublattice degrees of freedom. In momentum
space, this potential induces real all-to-all couplings between different values of $k_z$, $H_{\text{edge}}^{\mathbf{k},\mathbf{k}'} =
\rho\sigma_z\delta_{k_x,k_x'}\delta_{k_y,k_y'}$. This is now easily seen to obey Eq. (4.22) with $J = \sigma_y$.

This observation suggests that the edge modes previously observed at sharp boundaries of
the Hopf insulator are in fact protected by this 'unintentional' crystalline symmetry, and are
therefore not robust to perturbations that break the symmetry. To test this, we solve for the
(100)-edge modes of the dipolar Hopf insulator via exact diagonalization for three different
edge terminations: sharp, sharp with a symmetry-breaking perturbation, and adiabatic. We
observe three qualitatively distinct spectra [Fig. 4.6(a-c)]. The sharp edge hosts a linear
energy degeneracy, consistent with previous studies [131, 335]. To break the crystalline
symmetry, we add a site-dependent chemical potential $\mu_{\mathbf{v}}\mathbb{1}$ localized on the two unit cells $\mathbf{v}$
nearest the edge. This perturbation gaps the edge mode, supporting our conjecture that the
sharp edge modes of the Hopf insulator are in fact crystalline-symmetry-protected[2].

Finally, we consider smooth boundaries between the Hopf insulator and the trivial in-
sulator. To construct smooth boundaries, we take the hoppings to be constant throughout
the lattice, while an $x$-dependent staggered chemical potential $\mu_x\sigma_z$ tunes the Hamiltonian
between the trivial phase at each end of the lattice and the Hopf insulating phase in the
center. This interpolation occurs smoothly over two 'edge regions' on either side of the Hopf
insulating phase, consisting of $\sim 20$ lattice sites each. Shown in Fig. 4.6(c), these smooth
edges also feature gapless edge modes. Importantly, the gaplessness of these edge modes
is robust to *any* smooth perturbation to the lattice, including a 'smoothed' version of the
site-dependent chemical potential that was observed to gap the sharp edge mode [Fig. 4.6(d)].

---

[2]We hypothesize that the robustness of the (001)-edge modes [131] are artifacts of a similar stabilizing
crystalline symmetry, perhaps related to the discrete rotation symmetry $k_x \to k_y, k_y \to -k_x, n_x \to n_y, n_y \to
-n_x$ of the model studied, but do not study this here.
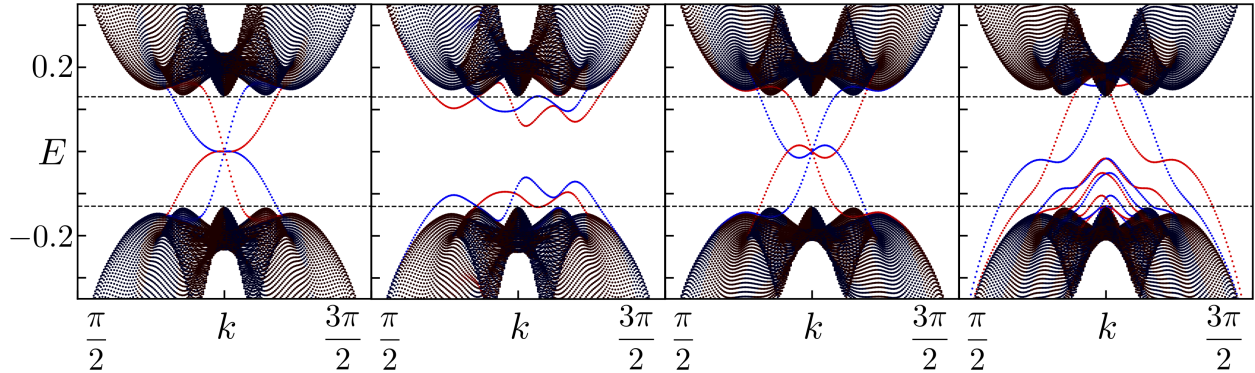
Figure 4.6:    Spectra for the (100)-edge of the Hopf insulator along the diagonal $k_y = k, k_z = \pi + k$, calculated from the effective Hamiltonian in Eq. (4.9). Color indicates a mode's mean $x$-position, from red/light gray (localized at left edge), to black (bulk), to blue/dark gray (localized at right edge). Dashed lines mark the bulk band gap. A sharp edge (open boundary conditions) respects the symmetry Eq. (4.22) and leads to a gapless Dirac cone spectrum (far left). Adding a symmetry-breaking perturbation – in this case, a chemical potential on the two sites nearest the edge – gaps the Dirac cone (left center), demonstrating the non-adiabatic edge modes' lack of protection. In the adiabatic limit with edge termination smoothed over $\sim 20$ lattice sites, the edge spectrum is again gapless (right center). However, the Hopf invariant now protects the edge modes against *all* smooth perturbations to edge, including a smoothed bump in the chemical potential in the edge region (far right). For given transverse momenta $k_y, k_z$, the spectrum is calculated by first Fourier transforming the Hamiltonian along the $y$- and $z$-directions, and then performing exact diagonalization on the remaining 1D Hamiltonian. Sharp/smooth edge spectra are calculated for a lattice with 80/160 unit cells in the $x$-direction, and hoppings are truncated at a range $R = 8$. As a check on the high-frequency approximation leading to the effective Hamiltonian, Eq. (4.9), we also perform the same computation for the *exact* Floquet Hamiltonian, $H_F = i \log \left( \mathcal{T} \exp \left( -i \int_0^{2\pi/\Omega_{xy}} H(t) \, dt \right) \right)$, at driving frequencies $\Omega_{xy} = 25 \, t_{\mathrm{nn}}$, $\Omega_z = 600 \, t_{\mathrm{nn}}$, and observe qualitatively identical edge spectra.

## 4.6   Experimental proposal

We now turn to our central result: a detailed blueprint for realizing the dipolar Hopf insulator using ultracold polar molecules. An explosion of recent experimental progress has led to the development of numerous possible molecular species [185, 331, 351, 364, 455], but for concreteness (and to demonstrate that the requisite separation of energy scales can be quantitatively realized), here we focus on $^{40}$K$^{87}$Rb [15, 16, 337, 351, 359, 507].

We begin with the geometry and rotational level diagram illustrated in Fig. 4.1. The 3D optical lattice is generated using four pairs of counter propagating beams, two forming the $xy$-lattice and two forming the $A$ and $B$ sublattices in the $z$-direction. For experimental convenience, we envision the two sublattices to be formed by beams with orthogonal linear polarizations of light. In this case a birefringent mirror can control the relative phase between the two reflected beams, which in turn determines the separation between sublattices.

To realize the rotational level diagrams of Fig. 4.1, we first propose to tune the rotational states $|1,0\rangle$ and $|1,1\rangle$ of all molecules to be approximately degenerate using applied DC electric and magnetic fields, oriented in the $z$-direction with amplitudes 1650 V/m and -490 G respectively[3]. The degeneracy between the $|1,0\rangle$ and $|1,1\rangle$ states, and, in turn, the sublattice symmetry between the $A$ and $B$ planes, can then be broken by using *different* intensities of light to form each sublattice. Owing to the AC polarizability of $^{40}$K$^{87}$Rb, the lattice beams not only trap the molecules in the designated geometry, but also induce $m_J$-dependent shifts in the molecules' rotational states proportional to the beams' intensities [349]. The individual intensities, $I_A$ and $I_B$, can therefore be tuned such that the transitions $|1,0\rangle_A \leftrightarrow |0,0\rangle_A$ and $|1,1\rangle_B \leftrightarrow |0,0\rangle_B$ are near-resonant with each other, yet off-resonant with all other transitions. Specifically, we calculate that $x$-polarized light with intensities $I_A =$0.43 kW/cm$^2$ and $I_B =$0.54 kW/cm$^2$ leads to the desired near-resonance, with an energy gap $\delta \sim 5$ kHz to the nearest rotational state outside the prescribed doublets. Energy levels are calculated as in Ref. [349], and we assume the $x$- and $y$-lattices are formed with $z$-polarized light of intensity .5 kW/cm$^2$. The molecule $^{40}$K$^{87}$Rb has a rotational splitting $\Delta = 2.2$ GHz and a measured dipolar interaction strength $t \sim 50$ Hz when trapped in a 3D optical lattice with 1064nm light [507]. This scheme therefore naturally leads to the desired separation of energy scales $t \ll \delta \ll \Delta$.

Energy levels in hand, let us turn to the implementation of the Floquet modulations (Fig. 4.7). To realize the $xy$-plane modulation, we can again rely upon the AC polarizability, using a two-dimensional intensity-modulated standing wave to directly tune the molecules' energy levels non-uniformly in both space and time. The energy shifts of the $|1,0\rangle_A$ and $|1,1\rangle_B$ states can be made equal [necessary to ensure the modulation is of the form of Eq. (4.13)] by tuning the polar angle of the light's polarization to $\theta = .96$ rad, owing to the anisotropic polarizability of $^{40}$K$^{87}$Rb [349]. An additional stationary standing wave on the even sites can cancel the site-dependent non-zero average of the modulation, preserv-

---

[3]The presence of the DC electric field mixes the $|0,0\rangle$ and $|1,0\rangle$ states, and gives rise to additional long-range density-density *interactions* in the hard-core boson model [513]. We may neglect these in our study of the single-particle physics of the Hopf insulator.
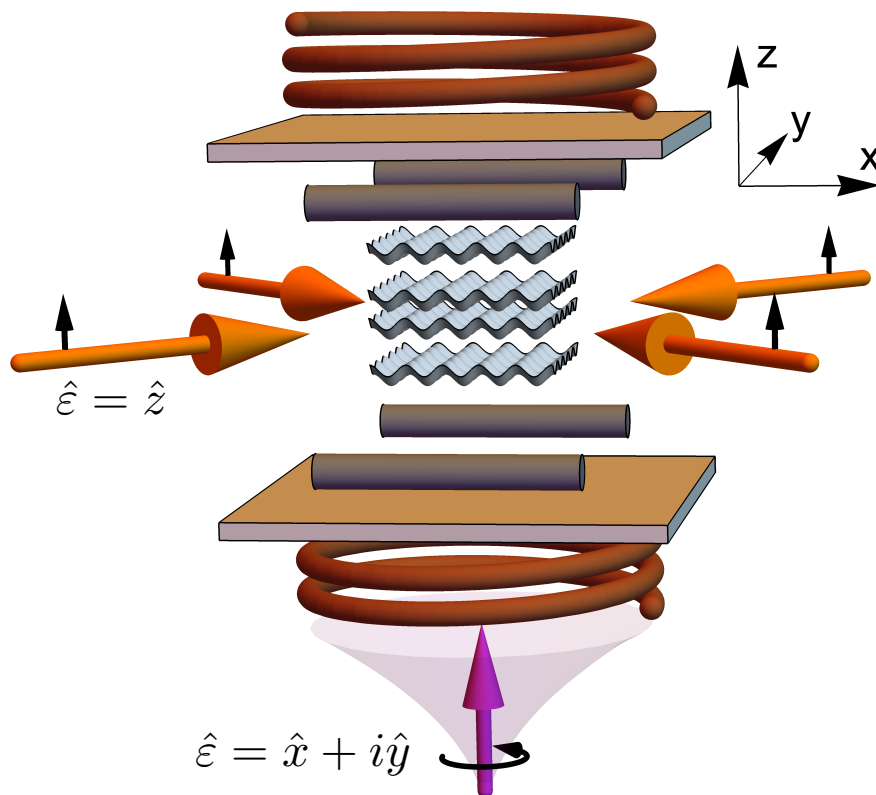
Figure 4.7: Schematic of the proposed experiment, highlighting the mechanisms for Floquet modulation. The lattice (light gray waves) is formed by three standing waves of laser light (beams not pictured). Stable electric field gradients are controlled an electrode system of tungsten rods (dark gray cylinders) and transparent plate electrodes (tan rectangles), while coils (brown spirals) generate a homogeneous magnetic field [116]. The $xy$-Floquet modulation is generated by $z$-polarized lasers forming a standing wave in the $(\hat{x} \pm \hat{y})$-directions (large orange arrows, left and right; polarization in small black arrows), using the AC polarizability of $^{40}$K$^{87}$Rb. The $z$-Floquet modulation is generated by a circularly-polarized laser in the $z$-direction (large purple arrow, bottom; polarization in small black arrow), which forms an intensity gradient along its direction of propagation due to the natural transverse spreading of a Gaussian laser beam.

ing translation invariance of the effective chemical potential. At a modulation frequency, $\Omega_{xy} \sim 500$ Hz, much greater than the dipolar interaction strength, $t_{nn} \sim 50$ Hz, the optimal modulation strength $g^{xy} = 1.2$ requires an intensity $\sim 10^{-2}$ kW/cm$^2$. An additional space-independent modulation of the two beams enables a difference between the two sublattices' modulations, achieving a nonzero $g_{SL}^{xy}$.

This method does not work for the $z$-gradient Floquet modulation, as a $z$-gradient in the light's intensity is necessarily accompanied by a polarization in the orthogonal $xy$-plane. In addition to shifting the molecules' energy levels, such a polarization would also induce mixing between rotational states, contaminating the desired hopping phase structure. Rather, we propose to achieve the $z$-gradient Floquet modulation by combining two independent sources of modulation [Fig. 4.8(c-e)]. First, we apply an oscillating electric field gradient of order $\delta E/\delta z \sim 1$ kV/cm$^2$. This gradient alone is not sufficient to realize the modulation of Eq. (4.13), because it shifts the energies of the the $|1,0\rangle_A$ and $|1,1\rangle_B$ states differently, owing to their different polarizability. We therefore combine this with a circularly-polarized beam tuned *near*, but off-resonant with, the $^3\Pi_{0+}$ electronic excited state of $^{40}$K$^{87}$Rb, which shifts the energy levels of the low-lying rotation states of interest via the AC Stark shift [Fig. 4.8(a-c)]. We imagine the beam to be traveling in the $z$-direction, with the natural transverse spreading of the beam along its propagation axis giving rise to a $z$-gradient in intensity $\delta I(z)/\delta z \sim I(z)/z$ [514]. To this end, we perform calculations of the AC polarizabilities of $^{40}$K$^{87}$Rb with circularly-polarized light as a function of detuning from the $b^3\Pi_{0+}$ state [Fig. 4.8(b)] using experimentally adjusted potential energy curves [22, 366] as well as parallel and perpendicular electronic polarizabilities [349], which we expand on in detail in the following section. For $\sigma^+$ light, the polarizabilities have poles at the resonant transition frequency to the excited $J = 2$ state, which allows the corresponding energy shifts to be precisely controlled by the detuning over a large range. Modulating the detuning about resonance (as a step function, to avoid any resonance-induced decay) precisely realizes the desired Floquet modulation. Quantitatively, we find that detunings $\Delta\nu \sim 1$ GHz lead to AC polarizibilities $\alpha/h \sim 1$ kHz/(W/cm$^2$), which in turn requires intensity gradients $\delta I/\delta z \sim 5$ W/($\mu$m cm$^2$) to achieve the optimal Floquet parameters at modulation frequency $\Omega_z \sim 5$ kHz $\gg \Omega_{xy}$. At a distance $z \sim 100\,\mu$m, the desired intensity gradient is thereby achieved with a modest intensity $I \sim .5$ kW/cm$^2$ and power $P \sim I(z) \times z^2 \sim 50$ mW [514].

We do not expect our proposed Floquet modulations to introduce substantial heating to the molecular system for a number of reasons. First, the modulations occur at a frequency significantly faster than the Hamiltonian energy scales, which exponentially suppresses many-body energy absorption [5]. Second, since the Hopf insulator's topology is characterized via its single-particle band structure, one only needs to excite a small number of molecules at any given time. At this single-particle level, the primary concern turns to heating from parametric processes associated with the laser intensity modulation. In this case, one can again utilize a separation of energy scales, by choosing the frequencies of the Floquet modulation to be far removed from any trap resonances (i.e. the trap frequency and its harmonics) such that no parametric heating will take place [348, 349]. Typical values of the trap frequency for $^{40}$K$^{87}$Rb experiments are $\sim 20$ kHz with a quality factor $\sim 20$ [348]; resonances are therefore

easily avoided both in our simple order-of-magnitude estimate, $\Omega_z \sim 5$ kHz and $\Omega_{xy} \sim 500$ Hz, as well as our more quantitative estimate in Fig. 4.6, using $\Omega_z = 600 t_{\text{nn}} \approx 30$ kHz and $\Omega_{xy} = 25 t_{\text{nn}} \approx 1.25$ kHz.

The edge modes of the dipolar Hopf insulator can be probed experimentally via molecular gas microscopy [117, 321]. Here, a tightly-focused beam applied near the edge induces local differences in the molecules' rotational splittings, enabling one to spectroscopically address and excite individual dipolar spins. The extent to which such an excitation remains localized on the edge during subsequent dynamics can be read out using spin-resolved molecular gas microscopy. For polar molecules separated by a distance of $1\,\mu$m, single-molecule addressing of the $|0,0\rangle \to |1,0\rangle$ transition has been estimated to require a beam of radius $1\,\mu$m and a reasonable power $10\,\mu$W [117]. The width of the edge region, typically large due to a wide harmonic confining potential, can be tuned via a number of recently developed techniques, including: box potentials [164], additional 'wall' potentials [541], or optical tweezers [301], allowing one to realize the three scenarios depicted in Fig. 4.6.

## 4.7 Details on AC polarizabilities for $z$-direction modulation

To effectively implement the Floquet modulation along $z$-direction, we use circularly polarized light tuned near a narrow transition, which allows light shifts to be precisely controlled by the detuning from the transition. Specifying to the molecule $^{40}\text{K}^{87}\text{Rb}$, we choose the dipole-forbidden transition $|X^1\Sigma^+, v = 0, J = 1, m_J\rangle \to |b^3\Pi_{0^+}, v = 0, J = 2, m_J + 1\rangle$ with 1028.7 nm [255] $\sigma^+$ light where $m_J = 0$ for the A sublattice and 1 for the B sublattice. With relatively weak laser intensity (on the order of $\text{W}/\text{cm}^2$), the light shift can be characterized by the AC polarizability of the molecular state of interest. The polarizability is calculated from two different contributions. The first and more important contribution comes from the resonant transition which has a strong dependence on the detuning, and the second contribution comes from all other transitions that has negligible dependence on the detuning in the range we are interested in. Here we assume the detuning is much larger than the spacings between $|X^1\Sigma^+, v = 0, J = 1, m_J\rangle$ states with $m_J = 0$ and $\pm 1$, and these spacings are much larger than the light shifts.

To characterize the contributions from the resonant transition, we follow the recipe in Refs. [64, 263, 443]. The generally complex dynamic polarizability for alkali-metal molecule in a rovibrational state of the ground $X^1\Sigma^+$ potential is given by

$$\alpha(h\nu, \hat{\varepsilon}) =$$
$$\frac{1}{\varepsilon_0 c} \sum_f \frac{E_f - E_i - ih\gamma_f/2}{(E_f - E_i - ih\gamma_f/2)^2 - (h\nu)^2} |\langle f|d\hat{R} \cdot \hat{\varepsilon}|i\rangle|^2 \tag{4.23}$$

$\hat{\varepsilon}$ and $\nu$ are the polarization vector and the frequency of the light, respectively, $c$ is the speed of light, $\varepsilon_0$ is the electric constant, $\hat{R}$ is the orientation of the interatomic axis, and $d$ is the
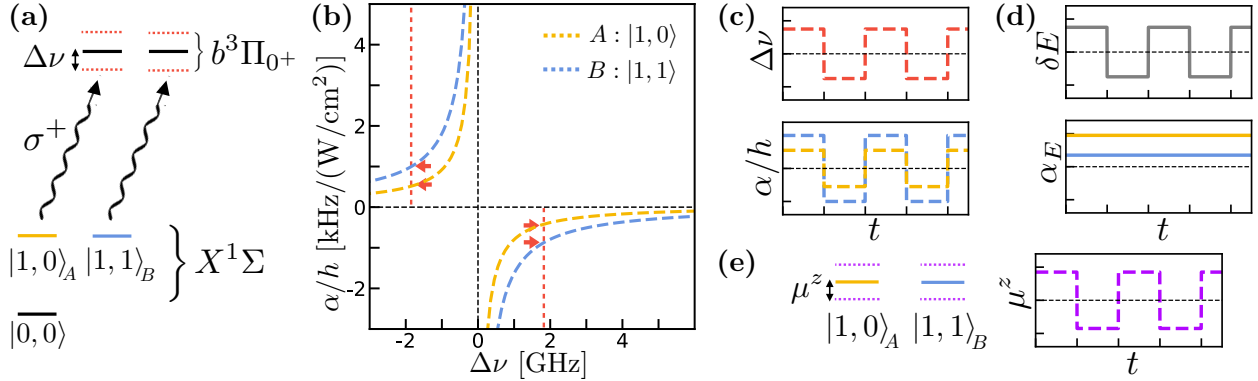
Figure 4.8:    Depiction of the two-component driving scheme used to implement the $z$-gradient Floquet modulation. **(a)** One component is realized using circularly-polarized ($\sigma^+$) light tuned near, but off-resonant with the electronic transitions $|X^1\Sigma, v = 0, J = 1, m_J\rangle \rightarrow |b^3\Pi_{0^+}, v = 0, J = 2, m_J + 1\rangle$, with detuning $\Delta\nu$. Here $X^1\Sigma, v = 0$ denotes the electronic ground state manifold of the molecule, and $b^3\Pi_{0^+}, v = 0$ the relevant electronic excited state manifold. This induces energy shifts in the electronic ground states of interest, $|1, 0\rangle_A$ and $|1, 1\rangle_B$, proportional to the AC polarizability $\alpha$ of $^{40}$K$^{87}$Rb at the particular detuning. **(b)** AC polarizibilities under circularly polarized $\sigma^+$ light as a function of the detuning $\Delta\nu$, calculated from first principles. Red dotted lines label two detunings that are oscillated between to achieve a step function Floquet modulation. Red arrows indicate the corresponding polarizibilities. **(c)** Simplified depiction of the detuning and resulting polarizibilities as a function of time $t$. In the dipolar simulations, we use a higher parameter step function, Eq. (4.19), which allows greater flexibility to optimize the band gap of the Hopf insulating phase. The AC light intensity is held constant in time (not depicted). **(d)** The other component of modulation consists of an electric field gradient $\delta E$ oscillated in time according to the same step function. The (DC) polarizibilities $\alpha_E$ of the $|1, 0\rangle_A$, $|1, 1\rangle_B$ states under this field are constant in time. **(e)** The polarizibilities and field amplitudes in **(c-d)** multiply to produce oscillating energy shifts $\mu^z$ (dotted purple) of the $|1, 0\rangle_A$, $|1, 1\rangle_B$ states. While each individual component of the $z$-gradient modulation produces a different magnitude shift for each state owing to the states' differing polarizibilities, the linear combination of both components can be chosen to produce equal shifts.

dipole operator. $i$ denotes the rovibrational state $|i\rangle$ of interest with energy $E_i$ in the ground
$X^1\Sigma^+$ potential, and the summation over $f$ denotes the summation over all rovibrational
states $|f\rangle$ other than $i$ with energies $E_f$ in all electronic potentials, and $\gamma_f$ describe the
natural linewidths of $|f\rangle$.

When the laser frequency is very close to the narrow dipole-forbidden transition, the
most significant contribution comes from that transition which has a pole at the resonant
frequency and weakens as the inverse function of the detuning. We treat all transitions from
$|X^1\Sigma^+, v = 0, J = 1, m_J\rangle$ to rovibrational states in the $b^3\Pi_{0^+}$ potential using Eq. (4.23).
The largest contribution by far comes from the transition to the excited $v = 0$ state due
to the similarity of its radial wavefunction to the ones in the ground potential. We use the
experimentally adjusted potential energy curves for both the excited $b^3\Pi_{0^+}$ state [22] and
the ground $X^1\Sigma^+$ state [366], and a spin-orbit modified transition dipole moment between
them [264]. Since the natural linewidths of the lowest rovibrational states in the $b^3\Pi_{0^+}$
potential are much smaller (on the order of kHz [255]) then the detunings we are interested
in (on the order of GHz), we take $\gamma_f = 0$.

The background contributions from all other transitions have negligible frequency depen-
dence close to the 1028.7 nm transition due to the large detunings from the corresponding
excited states. Thus we treat the background polarizabilities as constants throughout the
detuning range. We use the method in Ref. [265] with experimentally determined elec-
tronic parallel and perpendicular polarizabilities [349] to calculate the background polar-
izabilities at 1064 nm and assume them to be the same near the 1028.7 nm transition.
More specifically, we use $\alpha_\parallel/h = 10.0(3) \times 10^{-5}$ MHz/(W/cm$^2$) and $\alpha_\perp/h = 3.3(1) \times 10^{-5}$
MHz/(W/cm$^2$) determined for the wavelength of 1064 nm and obtain the background polar-
izabilities $\alpha_{bg,|1,0\rangle}/h = 4.64 \times 10^{-5}$ MHz/(W/cm$^2$) and $\alpha_{bg,|1,1\rangle}/h = 5.98 \times 10^{-5}$ MHz/(W/cm$^2$)
for $\sigma^+$ polarization.

Finally, we add the two parts together to arrive at the total AC polarizabilities shown in
Fig. 4.8(b).

## 4.8    Discussion and outlook

There has recently been a burst of theoretical interest in Hopf insulators and their possible
extensions, including non-hermitian generalizations [206], the survival of topology under
quantum quenches [215], crystal symmetries [19, 299], and generalizations to periodically-
driven Floquet systems [207, 428]. These ideas motivate the possibility of experimentally
realizing the Hopf insulator phase, which would allow one to test the above predictions, and,
more tantalizingly, could probe regimes of Hopf insulating physics that are much harder
for theory to handle. For instance, we have proposed using the presence of a gapless edge
mode at a smooth boundary, probed by spectroscopy, as a robust experimental diagnostic
of the Hopf insulating phase. Recent work suggests that at the (001)-edge this mode should
feature a nonzero Chern number associated with an unusual bulk-to-boundary flow of Berry

curvature [19]; numerous techniques to measure the Chern number have been developed [12, 377, 456, 497], which may allow one to detect this physics.

Our blueprint may also provide a basis from which to realize the various extensions of the Hopf insulator. In our proposal, we have already seen that polar molecules can realize certain crystalline symmetry-protected extensions of the Hopf insulator [19, 299], which can be detected independently from the ordinary (non-crystalline) Hopf insulator by looking at sharp edge terminations that respect the crystalline symmetry. Polar molecules might also be used to realize driven extensions of the Hopf insulator, for instance, the Floquet Hopf insulator [428]. Here, one subjects the system to periodic driving at a time-scale *comparable* to the hopping time, which can lead to a new Floquet Hopf insulating phase, characterized by a $\mathbb{Z} \times \mathbb{Z}_2$ pair of topological invariants that underlie an even richer spectrum of edge mode behavior than in the non-driven case. The Floquet Hopf insulator can be realized by strobing a flat band static Hopf insulator with periodic $\pi/2$-pulses of a staggered chemical potential [428]—the latter would be easily realized via a $\sim 100$ Hz oscillation of the lattice light intensity. Realizing a sufficiently flat band Hopf insulator is a less trivial task, but the bandwidth could be optimized via standard optimization techniques depending on the specific set of available experimental parameters. More speculatively, a flat band Hopf insulator might also be a natural launching ground into many-body generalizations of the Hopf phase (much as a flat band Chern insulator is a key ingredient for the fractional Chern insulator [50]), which are thus far unexplored territory.

Finally, the protocol outlined here and in Ch. 3 makes use of the high tunability and intricate real- and momentum-space structures afforded by recent advances in the manipulation of interacting dipolar molecules [323, 471]. Looking forward, the same approach suggests many promising avenues for realizing other exotic states presently residing at the forefront of theory [9, 180, 380, 390, 478, 532]. For instance, our use of a sublattice-dependent lattice light intensity to realize (pseudo)spin-orbit coupling via the $\Delta m = 1$ component of the dipolar interaction may prove fruitful in realizing other topological phases as well. As a simple example to demonstrate wider applicability, the exact same form of spin-orbit coupling $(t_{\mathbf{r}}^{AB} \sim e^{i\phi})$ in 2D gives rise to Chern insulating physics [380]. In polar molecule setups limited by the ability to fill only a (random) fraction of the full set of lattice sites, the Chern insulator might therefore provide a disorder-robust [9] stepping stone to realizing the Hopf insulator. We have also provided implementations of two independent Floquet engineering schemes: an even-odd patterning utilizing the molecules' AC polarizability under lattice light, and a truncation of the power-law dipolar interaction in the $z$-direction via a single circularly-polarized Gaussian laser beam. Floquet engineering has proven critical in other quantum simulation platforms, and these techniques may serve as building blocks for its use in polar molecules.

# Chapter 5

# Floquet Hopf insulators

In this Chapter, we introduce our first and only example of an intrinsically-dynamical phase of matter. We name this phase the "Floquet Hopf insulator". The Floquet Hopf insulator represents an extension of the Hopf insulator considered in Chs. 3 and 4 to periodically-driven, or *Floquet*, systems. Unlike the Hopf insulator realized in Chs. 3 and 4, the Floquet Hopf insulator cannot, even in principle, be realized in any static Hamiltonian.

The Floquet Hopf insulator follows in the footsteps of many other recent examples of intrinsically-dynamical phases [100, 245, 247, 259, 290, 310, 328, 461, 526]. In particular, we will show that it represents a particularly unique instance of the so-called Floquet topological insulators (FTIs): non-interacting, driven phases of matter, whose physical properties are characterized by a set of underlying quantized topological invariants [31, 90, 153, 183, 254, 297, 346, 388, 404, 407]. Unlike their non-driven counterparts, the topology of FTIs arises directly from the unitary time-evolution, leading to robustly protected gapless edge modes even when the stroboscopic time-evolution is topologically trivial.

To motivate the uniqueness of the Floquet Hopf insulator phase, we must first describe a common pattern in the classification of Floquet topological insulators thus far, which relates their topological invariants to those of static topological insulators with the same dimension and symmetries. Namely, a given FTI is found to possess all the invariants of its static counterpart, plus one additional invariant of *identical* classification. Intuitively, this is understood by extending the bulk-boundary correspondence to Floquet systems: under periodic modulation, the energy—now, quasienergy—becomes defined only modulo $2\pi$ (in units of the driving frequency) and thus an additional and identically classified edge mode emerges, associated with the bulk gap at quasienergy $\pi$. This result has been established rigorously in systems described by K-theory [404], and explored at great length in the context of specific symmetries and dimensionality [31, 90, 153, 254, 346, 407]. Nevertheless, one could wonder whether these arguments leave room for more unique topology in Floquet phases that escape this stringent bulk-boundary correspondence.

The Floquet Hopf insulator breaks this classification pattern: we will show in this Chapter that it is characterized by two *distinct* topological invariants: a 'static' $\mathbb{Z}$ invariant, and a uniquely Floquet $\mathbb{Z}_2$ invariant. The $\mathbb{Z}$ invariant of our system is precisely the Hopf invariant of
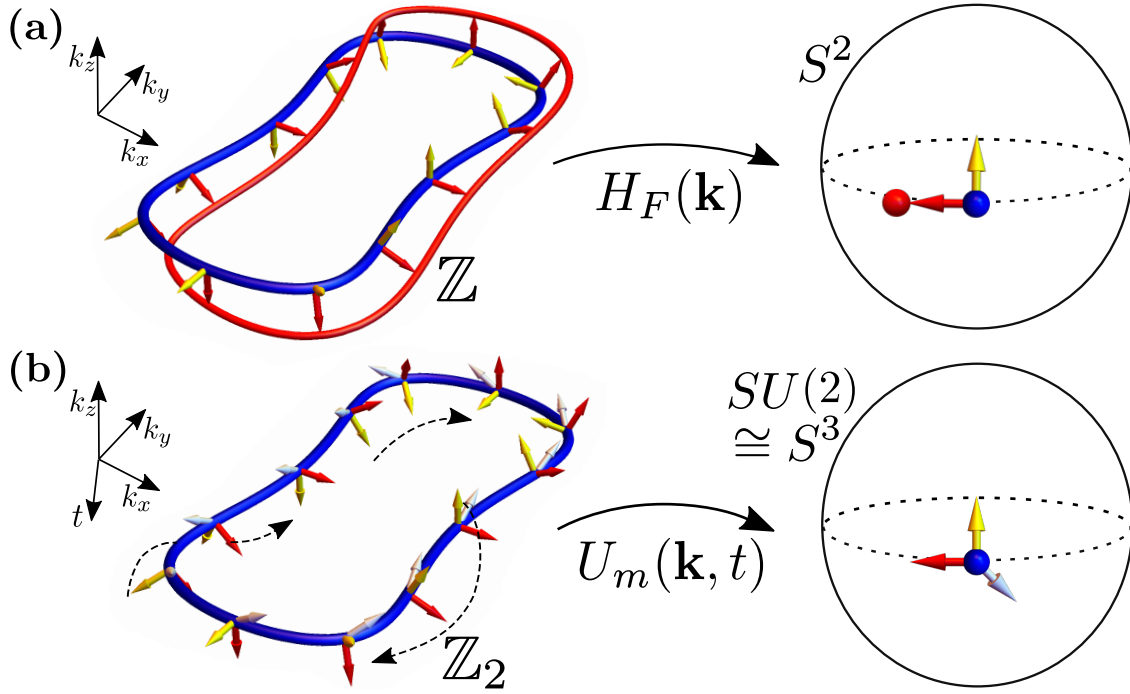
Figure 5.1:   Depiction of the Floquet Hopf insulator's two topological invariants. **(a)** The 'static' $\mathbb{Z}$ invariant is the Hopf invariant of the Floquet Hamiltonian $H_F(\mathbf{k})$, corresponding to the linking number of the pre-images of two points (blue, red) on the Bloch sphere. For nearby points, this equals the twisting of the Jacobian (colored arrows) along a single pre-image. **(b)** The 'Floquet' $\mathbb{Z}_2$ invariant classifies the micromotion operator $U_m(\mathbf{k}, t) \in SU(2)$, and is similarly interpreted as the Jacobian twisting (dashed black arrows) along a pre-image, with a reduced classification due to the larger dimensionality.

the static Hopf insulator. The $\mathbb{Z}_2$ invariant replaces the expected additional integer invariant, and characterizes the same topology that underlies the Witten anomaly in $(3+1)$D $SU(2)$ gauge theories [39, 143, 157, 498, 499]. In our context, it can be understood both as a twisting number extension of the Hopf invariant, as well as in terms of gapless topological defects of the Floquet evolution. These 'Hopf' defects may smoothly exchange even amounts of their topological charge, which leads to the reduced $\mathbb{Z}_2$ classification. Physically, the difference in invariants creates an atypical bulk-boundary correspondence, where gapless edge modes are topologically protected but may occur at either 0- or $\pi$-quasienergy, depending on non-universal properties of the boundary. In the remainder of this Chapter, we outline this physics in full detail. We begin with a brief summary of the homotopy classification of Floquet topological insulators generally (extending the homotopy classification of static topological insulators described in Ch. 2.2), and then proceed to the Floquet Hopf insulator.
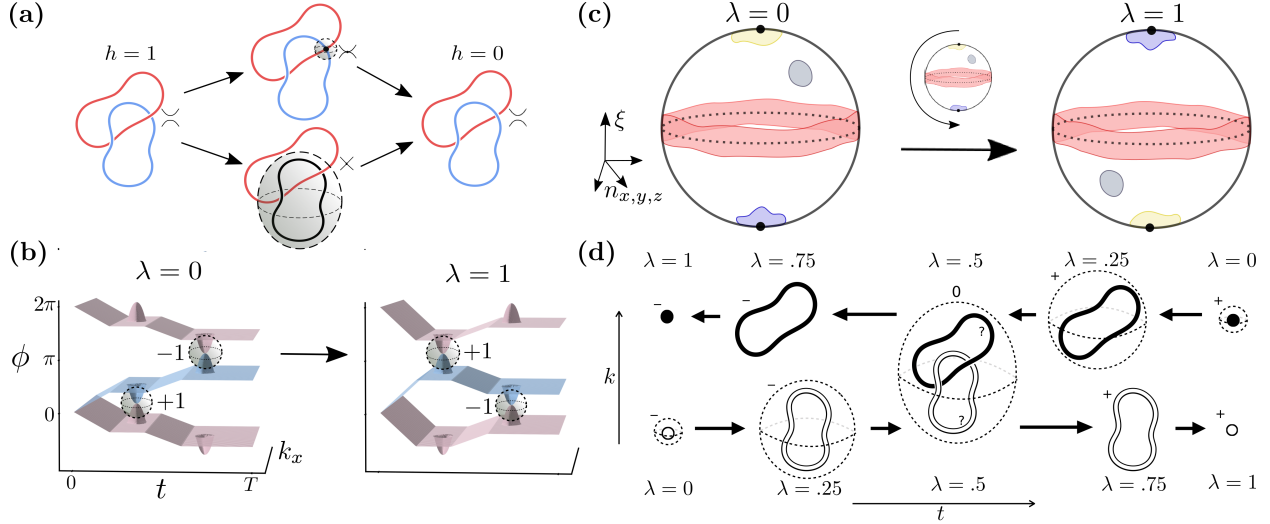
Figure 5.2: **(a)** A point Hopf defect (black point) has quadratic dispersion, functioning as a strand crossing that changes the linking number of any two eigenvectors' pre-images (red and blue). A loop Hopf defect (black loop) has linear dispersion, and can occur along a former pre-image. The defect charge is defined on a surface (gray, shaded) enclosing the defect. **(b)** Two Floquet evolutions with different defect charges but the same topological invariants, which are connected by a smooth deformation $\lambda \in [0,1]$ preserving the Floquet unitary's band gaps. **(c)** The deformation is a $\pi$ rotation of the 3-sphere parameterized by $(\mathbf{n}, \xi)$. Images of time-slices representing the initial 0-defect (yellow), $\pi$-defect (blue), trivial Hopf invariant (gray), Hopf invariant 1 (red), are displayed before and after the rotation. **(d)** During the deformation, the 0-defects (black outline) and $\pi$-defects (solid black) become loops that *link* in the Brillouin zone, at which point their individual charges are undefined. The total charge $h_0 + h_\pi$ is conserved, corresponding to the static $\mathbb{Z}$ invariant. Arrows indicate increasing $\lambda$.

## 5.1 Floquet topological insulators

We are concerned with non-interacting systems governed by a space- and time-periodic Hamiltonian, written in momentum-space as $H(\mathbf{k}, t) = H(\mathbf{k}, t + T)$, where $H(\mathbf{k}, t)$ is a matrix acting on the internal degrees of freedom that form the two bands of the system. Time-evolution is captured by the unitary operator $U(\mathbf{k}, t) = \mathcal{T}\left(e^{-i\int_0^t H(\mathbf{k}, t')dt'}\right)$, $0 \leq t < T$. Much like static insulators, one can view these unitaries in terms of the band-structures composed by their eigenvectors and eigenphases. For a two-band unitary we write

$$U(\mathbf{k}, t) = e^{i\phi} |z\rangle\langle z| + e^{i\phi'} |z'\rangle\langle z'|, \qquad (5.1)$$

where $\phi^{(\prime)}(\mathbf{k}, t)$, $|z^{(\prime)}(\mathbf{k}, t)\rangle$ depend on time as well as momentum, and the quasienergies $\phi^{(\prime)}(\mathbf{k}, t)$ are periodic.

Floquet topological insulators are Floquet-Bloch systems where the unitary is gapped *at time $T$*. The Floquet unitary, $U(\mathbf{k}, T)$, is equivalently described by the fictitious, time-independent Floquet Hamiltonian, $H_F(\mathbf{k}) = -i \log(U(\mathbf{k}, T))/T$. Similar to static TIs, two FTIs are in the same phase if one can smoothly interpolate between them without closing the gaps of the Floquet unitary. Focusing on systems with two gaps for simplicity, recent work has shown that—in all settings with non-driven analogues (i.e. in the absence of explicitly Floquet symmetries, e.g. time-glide symmetry [336])—such FTIs are characterized by *two* topological invariants, each with the same classification as a static TI of the same dimension and symmetries [31, 90, 153, 254, 336, 346, 404, 407].

Here, we note a finer distinction in the classification of FTIs with fixed band number. We decompose the unitary evolution into two components: the evolution over a full period, captured by the Floquet unitary $U(\mathbf{k}, T)$, and that within a period, captured by the micromotion unitary, $U_m(\mathbf{k}, t) \equiv U(\mathbf{k}, t) \left[ U(\mathbf{k}, T) \right]^{-t/T}$. From this decomposition, one sees that the classification factorizes into two, potentially distinct, invariants: a 'static' invariant classifying the Floquet Hamiltonian $H_F(\mathbf{k})$, and an intrisically Floquet invariant classifying the micromotion operator $U_m(\mathbf{k}, t)$. In $d$ space dimensions, the former classifies maps from the $d$D Brillouin zone to the set of gapped Hamiltonians, identical to the scheme for static TIs. The Floquet invariant classifies maps from the $(d{+}1)$D Floquet Brillouin zone, parameterized by $(\mathbf{k}, t)$, to $SU(n)$, for an $n$-band system without symmetries [407]. These invariants are identical in all cases previously considered. However, for systems with fixed band number they may differ.

## 5.2 Homotopy classification of Floquet Hopf insulator

We now introduce the Floquet Hopf insulator, a three-dimensional Floquet-Bloch system with two bands and no symmetries. The static invariant is the Hopf invariant of the Floquet Hamiltonian, which we briefly review. The gapped two-band Hamiltonian $H_F(\mathbf{k}) = \mathbf{n}(\mathbf{k}) \cdot \boldsymbol{\sigma}$ maps the 3D Brillouin zone to the Bloch sphere $S^2$. Neglecting "weak" lower-dimensional invariants [244, 335], such maps are classified by the homotopy group $\pi_3(S^2) = \mathbb{Z}$, thus possessing an integer topological invariant—the Hopf invariant. Expressed in terms of the eigenvectors $|z(\mathbf{k}, T)\rangle$ of the Floquet unitary, it takes the form $h_S = \frac{1}{2} \int d^3\mathbf{k} \, \epsilon^{ijk} \mathcal{A}_i \mathcal{F}_{jk}$, defining the Berry connection $\mathcal{A}_i = \frac{-i}{4\pi}(\langle z \rangle \partial_i z - \langle \partial_i z \rangle z)$ and curvature $\mathcal{F}_{jk} = \frac{-i}{4\pi}(\langle \partial_j z \rangle \partial_k z - \langle \partial_k z \rangle \partial_j z)$, and where $|z(\mathbf{k}, T)\rangle$ is related to the Bloch sphere $\mathbf{n}(\mathbf{k})$ by $\mathbf{n}(\mathbf{k}) = \langle z(\mathbf{k}, T)| \boldsymbol{\sigma} |z(\mathbf{k}, T)\rangle$.

The Hopf invariant has an intriguing visual interpretation as a linking number. To elaborate, consider the pre-image of any $\mathbf{n}'$ on the Bloch sphere, i.e. the set of all $\mathbf{k}$ that are mapped to $\mathbf{n}'$ by $\mathbf{n}(\mathbf{k})$. This is generically a 1D loop in the Brillouin zone. The topology of the HI enters when one considers two such pre-images. In the HI phase, any two pre-images are *linked*, with a linking number equal to the Hopf invariant. Intriguingly, this linking can be equivalently viewed as a *twisting* of the Jacobian of $\mathbf{n}(\mathbf{k})$ along a single pre-image [243, 372] (Fig. 5.1).
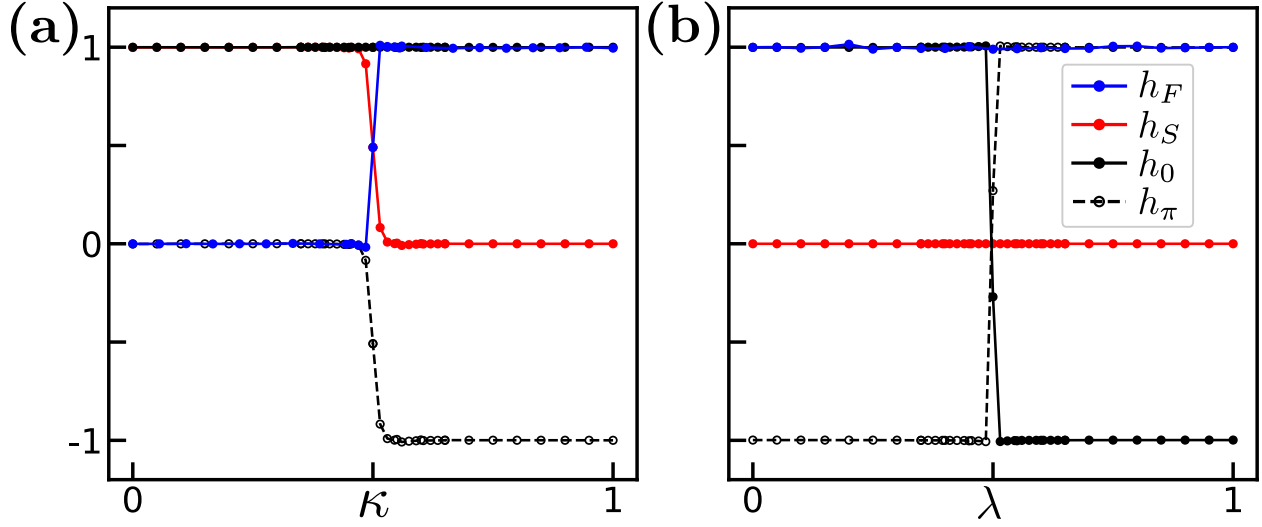
Figure 5.3:   Numerical calculation of the Floquet invariant, static invariant, and $0/\pi$-defect charges **(a)** across a phase transition $(h_S, h_F) = (1,0) \rightarrow (0,1)$ ($\kappa = 0 \rightarrow 1$) **(b)** along the smooth deformation exchanging defect charge.

We now turn to the Floquet invariant. The micromotion operator maps the 4D Floquet Brillouin zone to $SU(2)$, isomorphic to the 3-sphere $S^3$. Again neglecting weak invariants, this is classified by the group $\pi_4(S^3) = \mathbb{Z}_2$: a parity invariant, different from the integer Hopf invariant! This invariant was previously studied as the foundation of the Witten anomaly in $SU(2)$ gauge theories, where a formula for it was introduced [499]. In terms of the micromotion operator's eigenvectors $|z_m(\mathbf{k}, t)\rangle$ and their relative eigenphase $\Delta\phi_m(\mathbf{k}, t)$, we find (see Appendix A)

$$h_F = \frac{1}{4\pi} \int dt\, d^3\mathbf{k}\, \epsilon^{ijkl}\, \partial_i \Delta\phi_m(\mathbf{k}, t)\, \mathcal{A}_j\, \mathcal{F}_{kl} \bmod 2, \tag{5.2}$$

where the Berry connection and curvature are defined analogous to the non-driven case, now over space-time indices $\{k_x, k_y, k_z, t\}$. The Floquet invariant also relates to the Jacobian twisting along a 1D pre-image, now in (3+1)D (Fig. 5.1). The higher dimensionality leads to the reduced $\mathbb{Z}_2$ classification [372].

Combining the two invariants, we conclude that the Floquet Hopf insulator has a $\mathbb{Z} \times \mathbb{Z}_2$ classification. A system with arbitrary $(h_S, h_F)$ can be generated by strobing two flat band Hamiltonians according to

$$H_{(h_S, h_F)}(\mathbf{k}, t) = \begin{cases} \frac{2\pi}{T} H_{h_S - h_F}(\mathbf{k}) & 0 \leq t < T/2 \\ -\frac{\pi}{T} H_{h_S}(\mathbf{k}) & T/2 \leq t < T \end{cases}, \tag{5.3}$$

where $H_h(\mathbf{k})$ has Hopf invariant $h$ and energies $\pm 1$. To verify the static invariant, note that the Floquet unitary is given by $U(\mathbf{k}, T) = -e^{i\frac{\pi}{2} H_{h_S}}$, whose bands correctly have Hopf

invariant $h_S$. The Floquet invariant is also verified (see Appendix A): schematically, the contributions of the two halves of the evolution subtract, giving Floquet invariant $h_S - (h_S - h_F) = h_F$ mod 2.

## 5.3 Interpretation via topological defects and edge physics

It is illuminating to discuss how the Floquet Hopf insulator fits in the context of Ref. [346]. Here one again views the evolution in terms of bands, with particular attention to fixed time-slices. If the unitary $U(\mathbf{k}, t)$ is gapped at time $t$, one may define an instantaneous *static* topological invariant $\mathcal{C}(t)$ from its bands, exactly as one defines the static invariant of the Floquet unitary at time $t = T$. This invariant must be constant throughout each gapped region of the evolution, and can only change at times containing gapless points. Such points are *topological defects* of the evolution, and possess a defect charge equal to the total change in $\mathcal{C}(t)$ across the defect. They come in two varieties, 0- and $\pi$-defects, labelled by the quasienergy at which the gap closes. The total charges of the 0- and $\pi$-defects are locally conserved, and are thereby identified as the topological invariants of the evolution. For example, in the Floquet Chern insulator [407] the instantaneous Chern number changes at gapless Weyl points [478], and the integer charges of the 0- and $\pi$-Weyl points comprise a $\mathbb{Z} \times \mathbb{Z}$ classification [346].

Like other topological defects, Hopf topological defects possess an integer charge $h_{0/\pi}$ equal to the change in the instantaneous Hopf invariant across the defect. Two types of Hopf defect exist, each depicted in Fig. 5.2(a). The first occurs a single gapless point with a quadratic energy degeneracy. Interestingly, the Hopf invariant may also change across *loops* of gapless points, with linear degeneracy. The loops feature a Weyl cone [478] at each point, with the frame of the Weyl cone rotating by $2\pi\Delta h$ about the loop [299].

How does conservation of the two integer defect charges reconcile with the correct $\mathbb{Z} \times \mathbb{Z}_2$ classification? The answer lies in a smooth deformation that exchanges even charge between the 0- and $\pi$-defects, such that $(h_0, h_\pi) \to (h_0 - 2, h_\pi + 2)$. This process has no analogue in previously studied FTIs and keeps both band gaps of the Floquet unitary open, establishing the two configurations as the same phase. The total charge $h_0 + h_\pi$ is conserved in this process, while the individual charge $h_\pi$ is only conserved mod 2. This suggests the identifications

$$
\begin{aligned}
h_S &= h_0 + h_\pi \quad \in \mathbb{Z} \\
h_F &= h_\pi \bmod 2 \in \mathbb{Z}_2.
\end{aligned}
\tag{5.4}
$$

The former follows from the definition of defect charge: the invariant at $t = T$ equals the sum of all changes to it throughout the evolution. We explicitly describe the above deformation for the specific case of $(h_0, h_\pi) = (1, -1) \to (-1, 1)$, finding a continuous family of evolutions $U(\mathbf{k}, t; \lambda)$ with defect charges $(1, -1)$ at $\lambda = 0$ and $(-1, 1)$ at $\lambda = 1$ [Fig. 5.2(b)]. Recall that $SU(2)$ is topologically equivalent to the 3-sphere via the parameterization $U(\mathbf{k}, t) =$

$\xi(\mathbf{k}, t)\,\mathbb{1} + i\mathbf{n}(\mathbf{k}, t) \cdot \boldsymbol{\sigma}$, $\xi^2 + \mathbf{n}(\mathbf{k})^2 = 1$. The deformation acts as a time-dependent rotation of $U(\mathbf{k}, t)$ in the $\xi n_z$-plane: $U(\mathbf{k}, t; \lambda) = R_{\xi n_z}[\lambda\theta(t)]\{U(\mathbf{k}, t)\}$, where the rotation angle $\lambda\theta(t)$ interpolates from 0 at $t = 0$ to $\lambda\pi$ at times after the earliest defect.

To observe that this interpolates between the two configurations without closing the Floquet gap, we examine five regions of the $\lambda = 0, 1$ evolutions [Fig. 5.2(c)]. Throughout the deformation, the early and late gapped regions remain gapped with trivial topology. The middle region at $\lambda = 1$ is gapped with eigenvectors that can be smoothly deformed to $-\mathbf{n}$, which has the *same* Hopf invariant as the eigenvectors $\mathbf{n}$ at $\lambda = 0$[1]. Critically, this equivalence does *not* hold for TIs described by K-theory (e.g. the Chern insulator). Finally, the deformation interchanges the location of the 0- and $\pi$- defects. Since the intermediate invariant is unchanged, the defect charges are flipped by the deformation.

What allows the seemingly-conserved defect charges to change? Recall how defect charge is rigorously defined: one encloses the defect with a surface of gapped points, and computes the static topological invariant of the surface's eigenvectors [346]. As shown in Fig. 5.2(d), during the deformation the defects become loops of gapless points. At some value of $\lambda$, the 0- and $\pi$-defect loops *link* such that it is impossible to separately enclose each defect, causing the individual defect charges to be undefined. This defect linking arises directly from the linking of the HI (see Appendix A). After linking, the defects again have well-defined charges, which may differ from their initial values.

We compute the invariants and defect charges numerically in two scenarios (Fig. 5.3). Across a phase transition $(h_S, h_F) = (1, 0) \to (0, 1)$, both the invariants and defect charges change, following Eq. (5.4). In contrast, along the smooth deformation, the invariants remain robustly quantized while the defects exchange charge (see Appendix A).

Like its static counterpart [131, 335], the Floquet Hopf insulator features gapless edge modes at smooth boundaries between phases with different topological invariants (Fig. 5.4). An unusual situation occurs at boundaries where the static invariant changes, but the defect charge parities do not. Here, a gap closing is protected by the change in invariant, but may occur at either 0- *or* $\pi$-quasienergy, depending on details of the edge region. The anomalous $\mathbb{Z} \times \mathbb{Z}_2$ classification is precisely what allows this ambiguity: since the defect charges are only defined up to parity, neither quasienergy individually requires a gap closing, despite the change in topological invariant.

## 5.4 Discussion and outlook

At the highest level, the Floquet Hopf insulator demonstrates that the extension of static quantum phases of matter to driven systems is not always as simple as might be expected. In the context of topological physics, its most notable feature is the bizarre process by which edge modes can be exchanged between different band gaps, but never eliminated entirely.

---

[1]This follows from Eq. (3), since $\mathcal{F}_{jk}$ and $\mathcal{A}_i$ are each third-order in $\hat{\mathbf{n}}$, making the Hopf invariant sixth-order, and thus even, in $\hat{\mathbf{n}}$. The former follow from $\mathcal{F}_{jk} = \frac{1}{8\pi}\hat{\mathbf{n}} \cdot (\partial_j\hat{\mathbf{n}} \times \partial_k\hat{\mathbf{n}})$ and $\partial_j\mathcal{A}_k - \partial_k\mathcal{A}_j = \mathcal{F}_{jk}$.
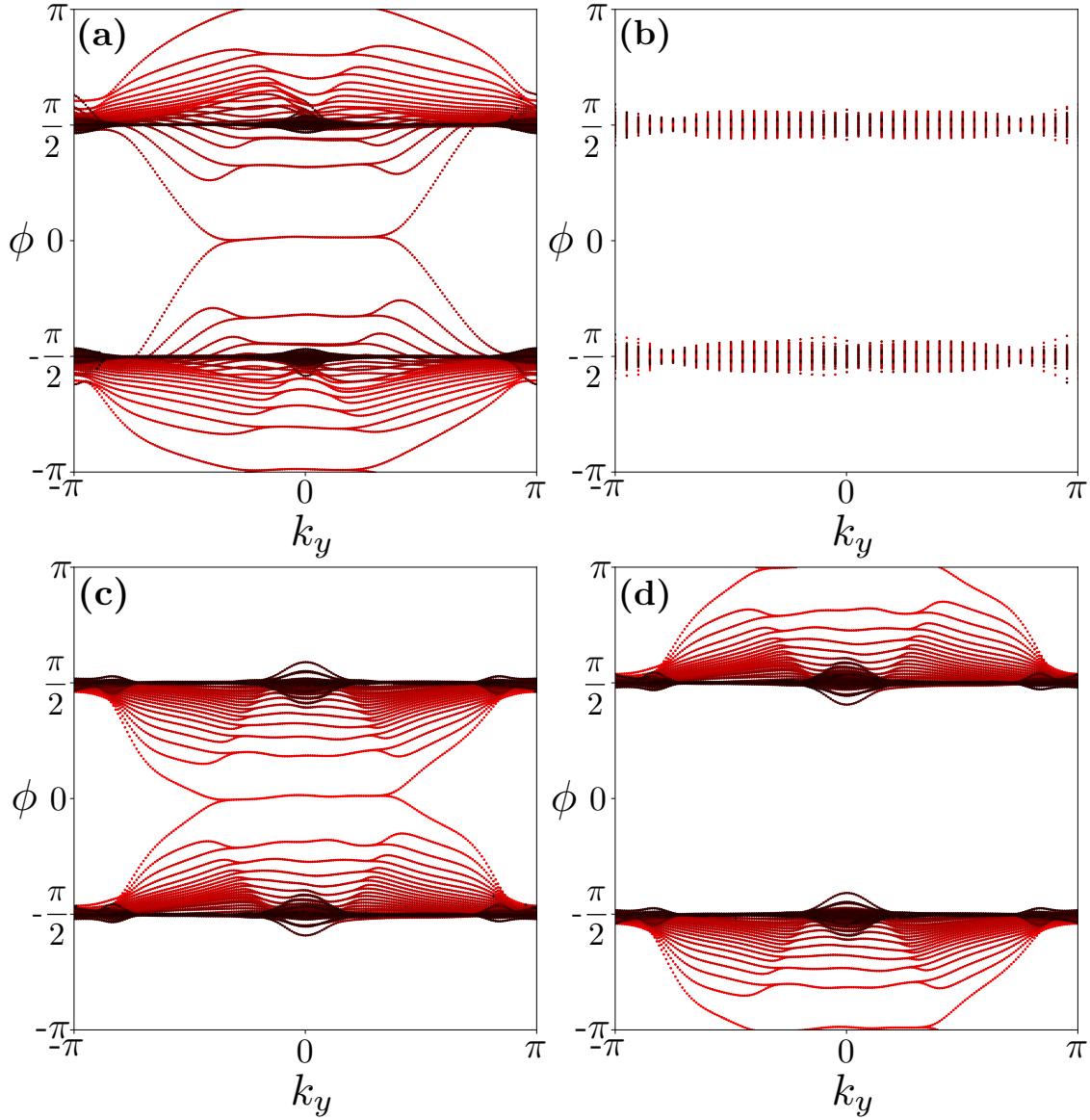
Figure 5.4:     Quasienergy spectra of the Floquet unitary at various smooth boundaries between Floquet Hopf insulator phases, solved via exact diagonalization (see Appendix A). Quasienergies are colored according to their eigenstates' average distance from the edge region, from localized at the edge (red) to far from the edge (black). **(a)** The boundary between $(h_S, h_F) = (0, 1)$ and the trivial phase $(h_S, h_F) = (0, 0)$, features gapless edge modes across both band gaps despite the Floquet Hamiltonian being trivial. **(b)** In contrast, we find no gapless edge modes between phases with different topological defect charges $(h_0, h_\pi) = (1, -1)$ and $(-1, 1)$, but the same topological invariants $(h_S, h_F) = (0, 1 \bmod 2)$, demonstrating the $\mathbb{Z}_2$ classification of the Floquet invariant. **(c,d)** Two different boundaries between the same two phases, $(h_S, h_F) = (2, 0)$ and the trivial phase $(h_S, h_F) = (0, 0)$, featuring gapless edge modes across either the 0- or $\pi$-gap.

Multiple open questions remain regarding the extent to which the Floquet Hopf insulator physics is observable in realistic physical systems. Many of these are common to all Floquet topological insulators [406]; for instance, concerning the robustness of the phase to interactions and, relatedly, the tendency of interacting systems to thermalize at late times. Experimental input may become essential to addressing these questions. To this end, the proposals in Chs. 3 and 4 to realize the static Hopf insulator may provide a first step towards a Floquet generalization. Specifically, we note that *any* realization of the static HI provides a direct path to realizing the Floquet Hopf insulator, using the stroboscopic construction of Eq. (5.3). For example, strobing a Hamiltonian with Hopf invariant 1 with a trivial Hamiltonian realizes the phase $h_S = 0, h_F = 1$. Time-evolving under a trivial Hamiltonian is straightforward: for instance, if the two bands arise from two sublattices, stroboscopic trivial time-evolution can be achieved by modulating the chemical potential on only a single sublattice. As a specific example, in the case of ultracold dipolar molecules [26, 117, 125, 338], a staggered chemical potential between two different spatial sublattices can be achieved by using different intensities of light. As in Ch. 4, recent advances in the context of KRb experiments [26, 117] suggest that the most interesting physical feature of Floquet Hopf insulator—its edge modes—can be probed using molecular gas microscopy.

# Chapter 6

# Braiding photonic topological zero modes

In this Chapter, we present the experimental realization of a non-interacting analogue of one of the longest sought features of topological phases: non-Abelian braiding. The experiment consists of 2-dimensional arrays of 1-dimensional photonic waveguides; classical light traveling through the waveguides will disperse across the waveguide array in a manner precisely analogous to quantum particles dispersing in (2+1)-dimensional spacetime. We will show that topologically-protected bound states in the waveguide array acquire a non-Abelian geometric (Pancharatnam-Berry) phase when braided around one another.

We begin this Chapter with a brief review of non-Abelian braiding and its importance in topological phases and topological quantum computing. Even before doing so, we preempt our results by noting that the non-Abelian braiding realized herein lacks certain hallmark associated with non-Abelian braiding in condensed matter settings. For one, the physics we consider does not require long-range entanglement between the bound modes being braided, and is (relatedly) not robust to the presence of interactions. Essentially, the entangled superpositions that lead to non-Abelian braiding in condensed matter systems are replaced here by unentangled superpositions that feature similar braiding properties. In addition, the braiding operations we find will realize a more restricted "gate set" than is found in non-Abelian braiding in condensed matter settings.

We also remark that in some ways, the presence of dynamics in this Chapter comes "by accident". Photonic systems are useful for our purposes due to their high tunability and inherent non-interacting nature. The same non-interacting nature then leads naturally to non-equilibrium physics, as there are no interactions to induce thermalization.

## 6.1 Overview of non-Abelian braiding

Quantum particles were long thought to fall into one of two classes: fermions or bosons. The two classes are distinguished by the effect on the many-particle wavefunction of ex-

changing any pair of such particles: for fermions, the wavefunction acquires a minus sign, whereas for bosons it does not. It was later understood [282] that, while these are indeed the only two possibilities in three or more spatial dimensions, the 2D case admits a remarkable generalization of the concept of bosons and fermions known as anyons [492, 493]. Mathematically, it is useful to think of anyons in $(2+1)$-dimensional spacetime, where their exchanges are viewed as "braids" [152, 347, 500] in which the particles' worldlines are wound around one another. In general, exchanging two anyons (i.e. performing a single braid) can yield any phase between 0 (bosons) and $\pi$ (fermions) [282]. Even more strikingly, in certain scenarios the phase acquired after performing a series of exchanges can depend on the order in which the exchanges occurred [152, 500]. Aside from its fundamental scientific importance, this "non-Abelian" braiding of anyons has attracted substantial interest in the condensed matter physics and quantum information communities as it can be used as a basis for robust quantum information processing [151, 253, 347].

Anyons have been theoretically predicted to arise in a variety of topological phases of matter, for example in fractional quantum Hall systems [29, 193], where they can exist as deconfined quasiparticle excitations, and in topological superconductors, where Majorana bound states nucleate at topological defects such as domain walls and vortices. New mechanisms for the realization of anyons have been proposed in non-interacting systems as well [306, 462], and have been discussed as a possible route for the realization of high magnetic permeability materials for photonic applications [462]. Despite substantial effort and progress [4, 25, 86, 239, 344, 495], as of yet there has been no conclusive experimental evidence of anyonic braiding, Abelian or non-Abelian.

Perhaps counterintuitively, the non-Abelian phases acquired upon braiding Majorana bound states in superconductors can be understood from the viewpoint of non-interacting particles, wherein the single-particle Schrödinger equation describes particles as waves. In this picture, the braiding phases are geometric phases [53, 54, 363] arising from the adiabatic variation of the phase texture of the bound-state wavefunctions as the vortices are wound around one another [225]. Therefore, non-Abelian braiding in its simplest incarnation can be viewed as a universal wave phenomenon accessible beyond electronic systems. This implies that it can be realized experimentally in the context of photonics, where a wide range of topological phenomena have been predicted and observed recently [189, 190, 360, 388, 482]. In fact, non-Abelian gauge fields have recently been observed in a photonic device [510], suggesting the possibility of photonic braiding. In general, photonic topological devices (as well as those in other bosonic systems) are expected to have entirely complementary applications to their condensed matter analogues.

In the remainder of this Chapter, we report on the measurement of the geometric phase arising from braiding topological defects in an array of photonic waveguides, fabricated using the femtosecond direct laser-writing technique [124, 453]. Our experiment utilizes classical light, which corresponds to a quantum coherent state of photons, consisting of a superposition of photon number states. Following the theoretical proposal of Ref. [222], we realize topological defects as vortices in a vector field that encodes the displacements of each waveguide in the array. The vortices realized in our experiment bind localized topological
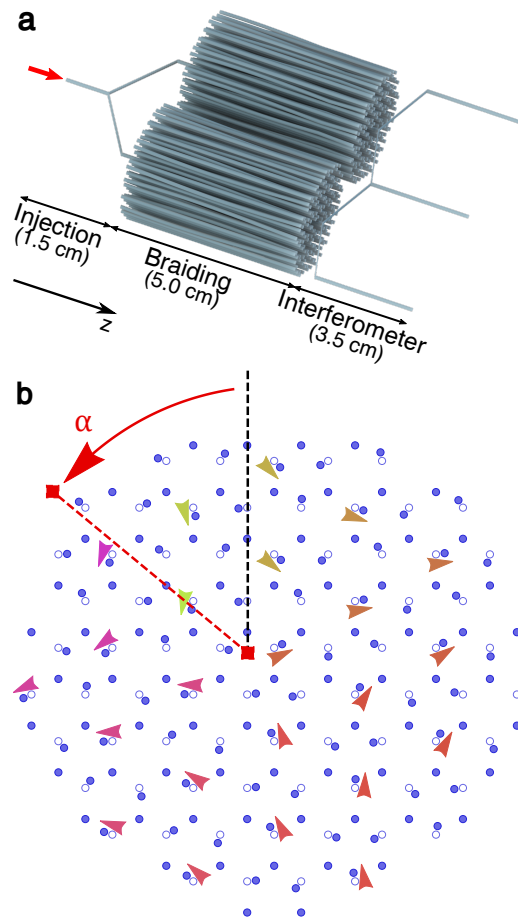
Figure 6.1: **Schematic of waveguide array.** **a,** Schematic of the waveguide array (dimensions not to scale). The experiment consists of three stages, occurring in succession in the propagation direction, $z$. First light is injected (indicated by the red arrow), split, and coupled into the vortex cores of the two lattices, each of which then undergoes a braiding operation. In the interferometer stage, light is extracted from the vortex cores and combined in a single waveguide at the center. The left and right arms are used to measure the intensity of light from each lattice after braiding. **b,** Depiction of one of the two lattices of the waveguide array at a fixed $z$-slice during the braiding stage. The waveguides (filled circles) are displaced from their honeycomb positions (empty circles) at an angle equal to the phase of the Kekulé order parameter $\Delta_{\mathbf{r}}(z)$ (arrows, drawn parallel to the displacements and colored according to their orientation). The order parameter contains a vortex of charge $-1$ (central red square) near the lattice center. The overall offset $\alpha$ of the order parameter's phase is varied as a function of $z$ and can be interpreted as the angle between the central anti-vortex and a fictitious vortex 'at infinity' (outer red square) that resides outside the waveguide array.

modes whose single-particle wavefunctions are identical to those of Majorana bound states in a 2D topological superconductor (Note however that the model we realize is in symmetry class BDI, while the topological superconductor is in class D [98]; thus the two systems have important fundamental differences). Consequently, at the non-interacting level the effect of braiding these vortices is the same as what is expected for Majorana bound states. We experimentally realize such vortices in the waveguide array and measure the effect of braiding one such vortex with a second that resides outside the array at an effectively infinite distance from the first. In order to eliminate the effect of dynamical phases, a 180° braiding operation is performed in two adjacent arrays, each with a vortex at its center. If the sense of rotation of the two arrays is the same (opposite), the relative phase at the core is found to be 0 ($\pi$). This observation matches the theoretically predicted geometric phase, providing a clear signature of braiding. The presence of the non-zero geometric phase here implies that braiding operations performed using such zero modes is non-Abelian, as shown explicitly in Ref. [222].

## 6.2 Photonic tight-binding model

We arrange the waveguides in a near-honeycomb lattice, with each waveguide displaced from its honeycomb position $\mathbf{r} = (x, y)$ by an $\mathbf{r}$- and $z$-dependent amount $\mathbf{u_r}(z)$. The diffraction of light through this waveguide array is governed by the paraxial wave equation

$$i\partial_z \psi(\mathbf{r}, z) = -\frac{1}{2k_0}\nabla_\mathbf{r}^2 \psi(\mathbf{r}, z) - \frac{k_0 \Delta n(\mathbf{r})}{n_0}\psi(\mathbf{r}, z), \tag{6.1}$$

where $\psi(\mathbf{r}, z)$ is the envelope function of the electric field $\mathbf{E}(\mathbf{r}, z) = \psi(\mathbf{r}, z)e^{i(k_0 z - \omega t)}\hat{x}$, $k_0 = 2\pi n_0/\lambda$ is the wavenumber within the medium, $\lambda$ is the wavelength of light, $\nabla_\mathbf{r}^2$ is the Laplacian in the transverse $(x, y)$ plane, and $\omega = 2\pi c/\lambda$. Here $n_0$ is the refractive index of the ambient glass and $\Delta n$ is the refractive index relative to $n_0$, which acts as the potential in equation 6.1. Since the displacements are small compared to the lattice spacing $a$ ($|\mathbf{u}| \leq .25a$) and vary slowly in the $z$-direction, the propagation of light through the waveguide array can also be described by a coupled-mode (i.e. tight-binding) equation:

$$i\partial_z c_\mathbf{r}(z) = \sum_{\langle \mathbf{r}' \rangle}[t + \delta t_{\mathbf{r},\mathbf{r}'}(z)]c_{\mathbf{r}'}(z) \tag{6.2}$$

where $c_\mathbf{r}(z)$ denotes the amplitude of light in waveguide $\mathbf{r}$, and the $z$-dependent hopping modification $\delta t_{\mathbf{r},\mathbf{r}'}$ arises from the change in waveguide separation due to the displacements $\mathbf{u_r}, \mathbf{u_{r'}}$. Identical to electrons in graphene, the two sublattices of the honeycomb lattice give rise to two energy bands in the spectrum, which touch at two gapless "Dirac" points at crystal momenta $\mathbf{K}_\pm = (\pm 4\pi/(3\sqrt{3}a), 0)$.

As shown in Fig. 6.1, we choose the displacements $\mathbf{u_r} = (u_\mathbf{r}^x, u_\mathbf{r}^y)$ corresponding to a Kekulé distortion [214, 222] controlled by the complex order parameter $\Delta_\mathbf{r}(z)$. The magnitude of

the order parameter determines the displacement amplitude, while its phase controls the displacement angle according to $\arg(u_\mathbf{r}^x + iu_\mathbf{r}^y) = \mathbf{K}_+ \cdot \mathbf{r} + \arg(\Delta_\mathbf{r})$. When $\Delta_\mathbf{r}(z)$ is constant in space, the Kekulé distortion couples the two Dirac cones, leading to a band gap proportional to $|\Delta|$ in the energy spectrum.

## 6.3 Vortices and braiding

Our work concerns order parameters that vary both throughout the lattice and in the $z$-direction. In particular, we focus on vortex configurations of the order parameter, $\Delta_\mathbf{r} = |\Delta| \exp(i[\alpha + q_v \arg(\mathbf{r} - \mathbf{r}_v)])$, where the phase of $\Delta_\mathbf{r}$ winds by $2\pi q_v$ about the vortex center $\mathbf{r}_v$ for a vortex of charge $q_v$. For $q_v = \pm 1$, such vortices are known to bind a single mid-gap photonic mode, localized near the vortex core [214, 222, 227]. The presence of this mode is protected by the nontrivial topology of the vortex configuration, as well as the time-reversal and chiral symmetries of the system [457]. (Here, chiral symmetry arises in the effective coupled mode equation due to the bipartite coupling between the two sublattices of the honeycomb lattice, similar to graphene [432].)

A system with multiple vortices carries a bound mode at each vortex. Away from the vortex centers, the order parameter varies slowly compared to the lattice scale, and the system remains gapped. This opens up the possibility of *braiding* the vortex modes as functions of $z$: if braiding is executed adiabatically in $z$, light bound in a vortex mode will not disperse into the gapped bulk.

Braiding has two effects on a vortex: moving its center $\mathbf{r}_v$, and altering the offset $\alpha$ in the local order-parameter phase. The latter arises due to the inherent nonlocality of vortices— as two vortices braid, each one changes its location in the space-dependent phase field of the other. The first vortex therefore experiences an effective offset $\alpha + q_{v,2} \arg(\mathbf{r}_{v,1} - \mathbf{r}_{v,2})$, which increases or decreases depending on the handedness of the braid and the charge of the second vortex. The effect of braiding on light trapped in the vortex mode is captured by the offset-dependence of the vortex-mode wavefunction $c_\mathbf{r}^v(\mathbf{r}_v, \alpha)$. Specifically, the wavefunction is *double-valued* in the offset,

$$c_\mathbf{r}^v(\mathbf{r}_v, \alpha + 2\pi) = -c_\mathbf{r}^v(\mathbf{r}_v, \alpha), \tag{6.3}$$

signifying that light trapped in the vortex mode gains a geometric phase of $\pi$ (i.e. a minus sign) after a full $2\pi$ braid.

The double-valueness of the wavefunction is reminiscent of Majorana wavefunctions at vortices in $p + ip$ superconductors [225, 387] or at superconductor-TI interfaces [155]. In fact, the parametric dependence on $\alpha$ of the photonic zero-mode wavefunction is identical to the dependence of the Majorana wavefunction on the phase of the superconducting order parameter [222]. However, since the underlying Hilbert space of the photonic system is *bosonic*, a single photonic zero mode has a well-defined notion of parity: an even or an odd number of photons can occupy it. In contrast, only a *pair* of Majorana zero modes has a well-defined (fermionic) parity. This crucial difference implies that braiding photonic
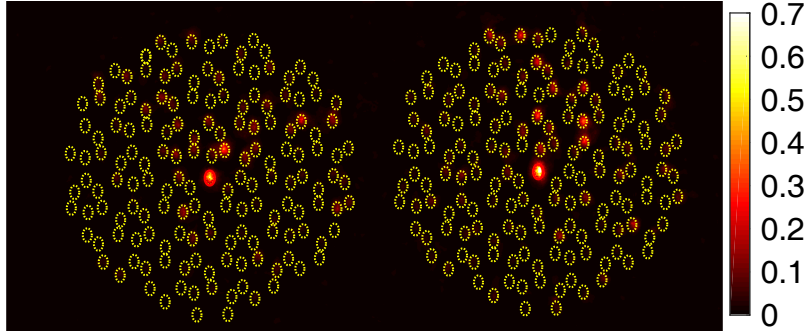
Figure 6.2: **Experimentally measured vortex mode.** Diffracted light measured at the output facet of a waveguide array device containing only the injection and braiding stages. Light stays largely confined to the vortex core, while excess light that does not overlap with the vortex mode diffracts throughout the array. A color bar is presented on the right, which indicates the light intensity at the output facet on a relative scale.

zero modes can implement only a subset of the possible operations realizable by braiding Majorana zero modes [222]. Nevertheless, the braiding are non-Abelian in both cases.

## 6.4 Observation of geometric phase induced by braiding

In this work, we provide a robust verification of the geometric phase $\pi$ gained by the photonic vortex modes under a $2\pi$ rotation of the order parameter $\alpha$. We detect this phase by performing two 'on-chip' interferometry experiments, in which we interfere light from two vortex modes that have undergone different rotations of $\alpha$. In each experiment, we fabricate a waveguide array containing two disconnected Kekulé-distorted honeycomb lattices, which constitute the 'left' and 'right' arms of an interferometer [see Fig. 6.1(a)]. The left and right lattices are initially identical and contain a single vortex in the Kekulé order parameter. As $z$ increases, the offset $\alpha$ of each lattice's order parameter is monotonically increased, or decreased, by $\pi$. The first experiment serves as a control, with $\alpha \to \alpha + \pi$ in an identical sense of rotation in both lattices. In the language of braiding, this increase of $\alpha$ can be interpreted as braiding a vortex 'at infinity' counterclockwise around each lattice by an angle $\pi$ [as indicated in Fig. 6.1(b)]. Since the left and right lattices are identical throughout, light trapped in the left vortex mode should gain the same phase as light trapped in the right (both dynamical and geometric phases), and they should interfere constructively after braiding.

The second experiment serves to detect the geometric phase from a $2\pi$ rotation of $\alpha$. Here the left lattice undergoes a rotation $\alpha \to \alpha + \pi$, while the right undergoes $\alpha \to \alpha - \pi$. This is analogous to braiding a vortex at infinity 180° counterclockwise about the left lattice, but clockwise about the right. Although the final left and right lattices are identical, the paths they have traversed differ by a $2\pi$ rotation of $\alpha$, and the relative phase gained by light

propagating through the two vortex modes will differ according to the geometric phase of this process. We note that, in theory, chiral symmetry pins the energy of the vortex mode precisely to the middle of the band gap [214, 222, 227], such that, despite their different $z$-evolutions, light in the left and right lattices is expected to gain the same dynamical phase during braiding. In practice, however, the waveguide lattice does not strictly realize chiral symmetry due to evanescent couplings between next-nearest-neighbor waveguides. Nevertheless, the effect of this chiral symmetry breaking is small as the next-nearest-neighbor couplings are suppressed exponentially relative to the nearest-neighbor couplings that preserve chiral symmetry. The accumulated dynamical phase is thus expected to be small. The symmetric braiding protocol proposed here further aids in the suppression of dynamical phases by ensuring that the path length of light through the waveguides in each arm of the interferometer is as similar as possible.

The entire experimental waveguide array consists of three stages, occurring in succession as a function of $z$ [see Fig. 6.1(a)]. The total length of the sample is 10 cm and the full width at half maximum of the major and minor axes of the waveguides are 8.91 and 5.83 $\mu$m, respectively. The first stage prepares the light to be coupled in-phase into the vortex modes of the left and right lattices. This is done using an 'on-chip' Y-junction beam-splitter [226]. We first precisely couple light from a tunable-wavelength laser through a lens-tipped fiber into a single waveguide at the input facet of the sample. We then split this waveguide into two in a symmetric manner, such that both waveguides trap light of nearly equal amplitude and phase. The two waveguides are subsequently guided to the center of the vortex cores of the left and right lattices.

The middle stage contains the braiding operation, and begins by abruptly initializing all other waveguides in each lattice. Light traveling in the waveguides of the first stage enters each lattice through a single waveguide near each vortex core, setting the initial condition for the braiding process. The portion of the input light that overlaps with the vortex mode remains localized in the mode, while the rest diffracts throughout the lattice. After initialization, the waveguide displacements are smoothly varied in $z$ to produce the desired change in order parameter for the experiment being performed. We emphasize that the individual waveguides move little during this process: each waveguide only undergoes rotation of its displacement angle from the undistorted honeycomb position.

In Fig. 6.2, we show the experimental output of a waveguide array containing only these first two stages. As expected, a high fraction of the light is localized near each vortex core, indicating that the initialization stage succeeded and the braiding process was performed sufficiently adiabatically. We also observe a nonzero intensity in nearly all waveguides of both lattices, indicating that light not bound to the vortex modes has diffracted throughout each lattice.

To detect the phase of the vortex mode light after braiding, we add a third, interferometric stage to the waveguide array [226], depicted in Fig. 6.1(a). The left and right lattices are abruptly terminated except for one waveguide each, chosen for its high overlap with the final vortex-mode wavefunction. Light localized in the left and right vortex modes continues to propagate through the remaining two waveguides. The intensity and phase of these

waveguides' light are proportional to those of the respective vortex modes after braiding. In order to read out the waveguides' relative phase, we first split each waveguide symmetrically into two, similar to the first stage. The two innermost waveguides of the left-right pair are then combined into a single waveguide, while the two outermost remain separate. The waveguide combination performs the interferometry: if the two input arms are in phase, they will excite the symmetric bound mode of the combined waveguide; if they are out of phase, they will not overlap the bound mode and thus the light will diffract away. The combined waveguide intensity $I_C$ thus indicates the intensity of in-phase light. The outermost left and right waveguide intensities $I_L$ and $I_R$ measure the waveguide intensities before combination.

Results of both experiments including the final stage are shown in Fig. 6.3. As anticipated, braiding the left and right vortices in the same direction results in a combined waveguide with higher intensity than the individual left and right waveguides, indicating constructive interference. On the other hand, braiding the left and right vortices in opposite directions results in nearly zero combined waveguide intensity, indicating near-complete destructive interference and therefore a relative phase of $\pi$ between the left and right vortex modes. To quantify these results, we define the *contrast* $\eta$ as the ratio between the combined waveguide intensity and the sum of the individual left and right waveguide intensities, $\eta = I_C/(I_L + I_R)$. In the ideal case, the contrast achieves the upper bound 1 for perfect constructive interference and the lower bound $(I_L + I_R - 2\sqrt{I_L I_R})/[2(I_L + I_R)]$ for perfect destructive interference. As plotted in Fig. 6.3(a), the experimentally observed contrast is indeed near 1 when braiding is performed in the same direction for both lattices, and is close to zero (the lower bound for symmetric intensities, $I_L = I_R$) when it is performed in opposite directions. Numerical simulations carried out using a split-step beam propagation method [195] conform to these predictions.

We verify that this phase difference arises as a geometric, rather than a dynamical, phase by demonstrating that the interference is insensitive to the wavelength of light used. The wavelength sets the 'time' scale in the paraxial equation, equation 6.2, so that different-wavelength light acquires different dynamical phases during propagation [184]. If the relative phase between the left and right waveguides arose as a dynamical phase, we would expect it to oscillate as the wavelength of light was varied rather than being quantized at $\pi$. Instead, as shown in Fig. 6.3, we observe quite consistent, non-oscillatory values of $\eta$ for both experiments over a large range of wavelengths, 1450–1650 nm. The principal sources of error/noise are interference from imperfectly in-coupled light in the braiding stage as well as diffracted light in the interferometric stage. We note that the random error is significantly lower in the sample in which we observe the $\pi$ phase associated with braiding [red points in Fig. 6.3(a)] compared to the 'control' sample in which we do not [blue points in Fig. 6.3(a)]. The measured $\eta$ values when the two vortices are braided in equal and opposite directions are $1.055 \pm 0.247$ and $0.053 \pm 0.037$, respectively. This indicates that the relative phase is fixed to $\pi$ and supports its identification as a geometric phase.
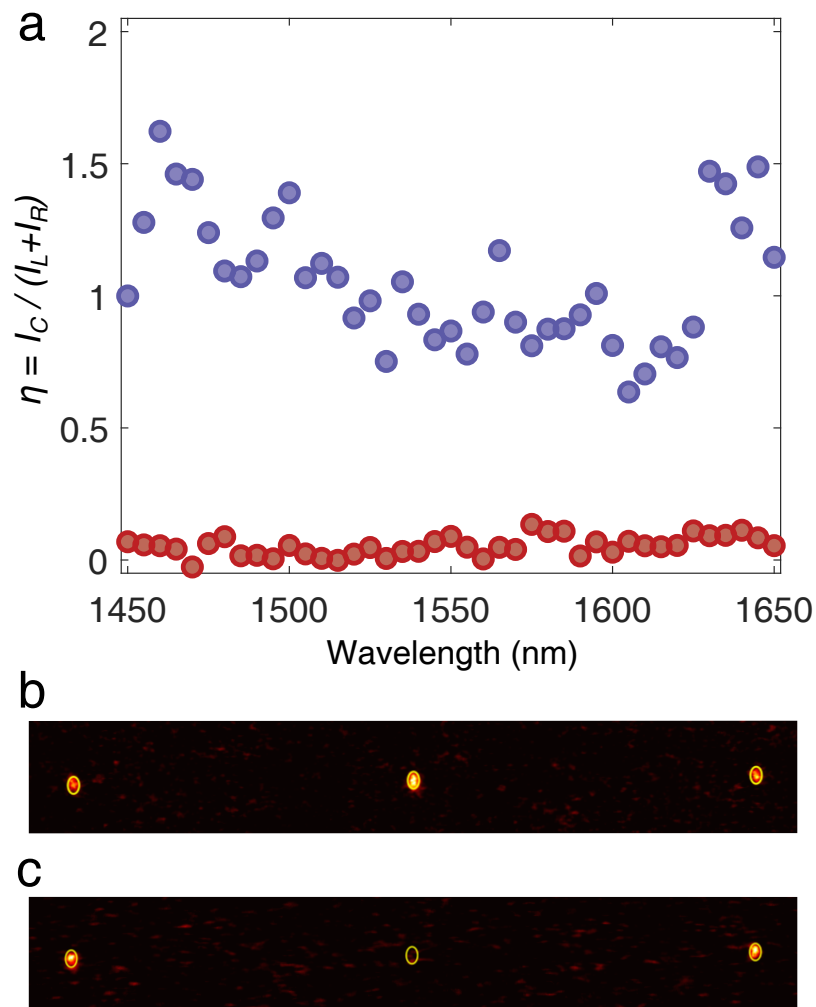
Figure 6.3: **Interferometric measurement of Berry phase associated with braiding. a,** The experimentally observed contrast when the two lattices are braided in the same direction (blue) or in opposite directions (red). The geometric phase $\pi$ picked up by light near the vortex core in the latter case is evidenced by the near-zero contrast due to destructive interference. **b-c,** Diffracted light measured at the output facet when the two lattices are braided in the same and opposite directions, respectively.

## 6.5   Methods

We fabricated the waveguides using femtosecond direct laser writing technique. Using a 800 nm Titanium:sapphire laser and amplifier system (Coherent:RegA 9000 with pulse duration 270 fs, repetition rate 250 kHz, and pulse energy 820 nJ), we wrote the waveguides in a borosilicate glass (Corning Eagle XG borosilicate glass) with refractive index $n_0 = 1.473$ at $\lambda = 1550$ nm. The shape and size of the focal volume were controlled by first sending the laser writing beam through a beam-shaping cylindrical telescope and then focusing it inside the glass chip using a $\times 50$, aberration-corrected microscope objective (NA = 0.55). The waveguides were fabricated by translating the glass chip through the focal volume of the laser beam using a high-precision three-axis Aerotech motion stage (model ABL20020).

Experiments are performed by butt-coupling a single-mode optical fibre to waveguides at the input facet of the chip, which subsequently couples light into the waveguide array. A tunable mid-infrared diode laser (Agilent 8164B), which can be tuned through the 1,450âĂŞ1,650 nm wavelength range, was used as an input light source. After a total propagation distance of 10 cm within the array, the diffracted light is observed at the output facet from the waveguide array using a 0.2 NA microscope objective lens and a near-infrared InGaAs camera (ICI systems).

## 6.6   Discussion and outlook

In summary, we have used a photonic lattice of evanescently-coupled waveguides to directly measure the braiding of vortices. The $\pi$ phase that we observe here directly implies that non-Abelian braiding operations can be carried out. In a similar spirit to Chs. 3 and 4, the results of this Chapter highlight the potential of dynamical quantum systems to realize long-sought static topological phenomena. However, as we noted in Ch. 5, the stability of such phenomena to interactions is often much weaker in dynamical compared to statics settings.

Using similar ideas to those here, we note that non-Abelian braiding can in principle be realized in any platforms governed by classical wave equations, such as systems of acoustic phonons, or polariton condensates. We also hope that our work motivates the exploration of non-Abelian braiding physics in interacting bosonic systems (via optical nonlinearity or mediation by Rydberg atoms, for example), as well as the search for applications beyond robust quantum information processing.

# Part III

# Many-Body Quantum Information Dynamics

# Chapter 7

# Verified quantum information scrambling on a trapped ion quantum processor

In Part II, we studied how periodically-driven dynamics can provide new avenues to stabilize quantum states with interesting quantum information properties (i.e. topological phases of matter). In this Part, we turn to the dynamics of quantum information itself. As discussed in Ch. 2.3, we will be particularly interested in how initially local information is stored within a quantum system after time-evolution.

The works contained in this Part seek to address a few broad questions regarding quantum information dynamics. How do quantum information dynamics differ across different "types" of many-body quantum systems? (Chs. 10 and 11) How can one measure these dynamics in laboratory experiments? (Chs. 7, 8, 9 and 10) And finally, what can we use our knowledge of quantum information dynamics for? (Chs. 9 and 10, as well as Chs. 12 and 13 in Part IV) A flood of recent works over the past decade (including, in small measure, those contained in this Part) have to a substantial degree answered the first two questions. The third question remains an exciting open direction in the years to come.

In this and the following Chapter, we present two experimental measurements of quantum information scrambling. These experiments seek to address a crucial question: How can experimental measurements distinguish the effects of scrambling from the effects of experimental noise and decoherence? For example, the most common probe of scrambling are out-of-time-ordered correlators (OTOC's) [277, 315, 436], which take the form $\langle \hat{V}^\dagger \hat{W}^\dagger(t) \hat{V} \hat{W}(t) \rangle$, where $\hat{V}$ and $\hat{W}$ are unitary operators acting on separate subsystems. As scrambling proceeds, $\hat{W}(t)$ becomes increasingly nonlocal, causing the OTOC to decay [398], which is taken as an experimental indication of scrambling [163, 289, 325, 486].

However, it is difficult to distinguish between information scrambling and extrinsic decoherence in the OTOC's temporal decay. For example, non-unitary time-evolution arising from depolarization or classical noise processes naturally lead the OTOC to decay, even in the absence of quantum scrambling. A similar decay can also originate from even slight mismatches between the purported forward and backwards time-evolution of $\hat{W}(t)$ [163, 450, 518]. While full quantum tomography can in principle distinguish scrambling from

decoherence and experimental noise, this requires a number of measurements that scales exponentially with system size and is thus impractical.

In this and the following Chapter, we implement a quantum circuit that provides a positive test for the scrambling features of a given unitary process [517, 518]. This approach utilizes many-body dynamics to conditionally teleport a quantum state through the circuit, providing an unambiguous litmus test for scrambling while projecting potential circuit errors into an ancillary observable. The quantum circuit implementing this teleportation was originated in the context of the black hole information paradox [201, 361]. Its discovery suggested a resolution to the paradox where quantum information that has crossed the event horizon can be retrieved assuming that the dynamics of the black hole are both unitary and scrambling [160, 317, 449]. These experiments thus also serve as proof-of-principle demonstrations of this "many-body" teleportation.

In this Chapter, we implement the many-body teleportation circuit on a 7-qubit ion trap quantum computer. The teleportation circuit involves a tunable 3-qubit unitary operation; the teleportation fidelity probes the scrambling dynamics of the unitary. Measured teleportation fidelities are typically $\sim 80\%$, and enable us to experimentally bound the scrambling-induced decay of the corresponding OTOC measurement.

## 7.1 Hayden-Preskill recovery protocol

The intuition behind our approach lies in a re-interpretation of the black hole information paradox, depicted schematically in Fig. 7.1. Let us suppose that an observer (Alice) throws a secret quantum state into a black hole. Is it then possible for an outside observer (Bob) to reconstruct this state by collecting the Hawking radiation emitted at a later time? Assuming the dynamics of the black hole can be modeled as a random unitary operation $\hat{U}$, it has been shown that Bob must wait for at least half of the lifetime of the black hole to reliably recover Alice's quantum state [361]. But what if Bob possesses a quantum memory that was previously entangled with the black hole? In this case, it was shown that the decoding of Alice's quantum state could be performed by collecting only a few Hawking quanta [202].

To this end, an explicit decoding protocol has been recently proposed [517, 518], which enables Bob to decode Alice's state using only his quantum memory, an ancillary EPR (Einstein-Podolsky-Rosen) pair, and knowledge of the effective black hole unitary $\hat{U}$ [318]. The protocol requires Bob to apply $\hat{U}^*$, the complex conjugate of the black hole's evolution operator, to his own quantum memory and one half of the ancillary EPR pair. Following this, Bob performs a projective EPR measurement on the emitted Hawking radiation and the corresponding subsystem of his own quantum memory. This projective measurement plays the role of teleporting Alice's secret quantum state to the reference qubit in Bob's ancillary EPR pair. The successful decoding of Alice's quantum information is possible due to the maximally scrambling dynamics of the black hole, which ensure that the information about Alice's secret state is distributed, almost immediately, throughout the entire black
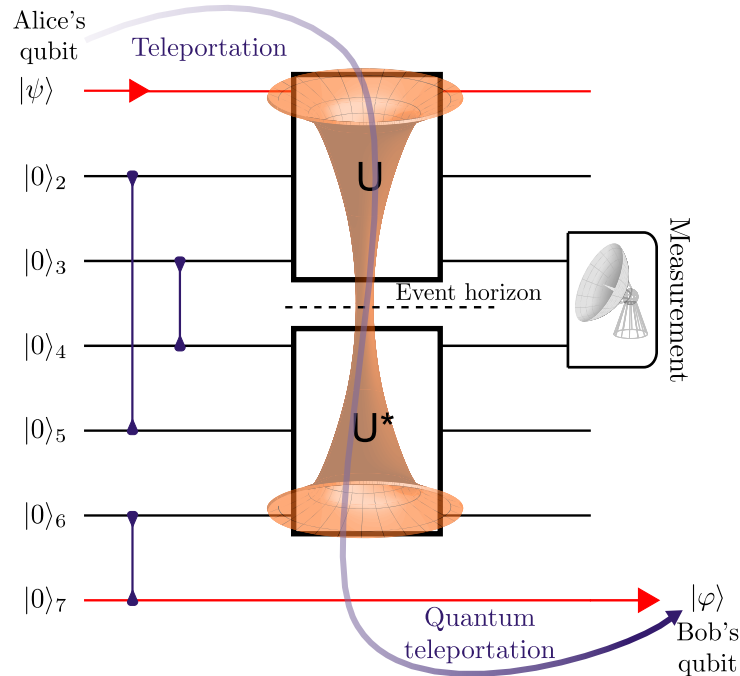
Figure 7.1: Schematic of our 7-qubit scrambling circuit with Alice's qubit, $|\psi\rangle$, as an input. Alice's information is scrambled throughout the entire system by maximally scrambling Clifford circuits $\hat{U}$ and $\hat{U}^*$. If the scrambling operations are performed without errors, Bob will be able to teleport Alice's qubit to himself ($\langle\varphi|\psi\rangle = 1$) by a projective measurement. The underlay depicts an interpretation of the unitaries $\hat{U}$ and $\hat{U}^*$ as representing a two-sided black hole connected by a wormhole [160, 317, 449]; in the experiment, the two-sided black hole is modeled via multiple EPR pairs as in [318]. The maximally scrambling nature of the black hole dynamics is captured by $\hat{U}$ and $\hat{U}^*$. The vertical lines indicate that qubit pairs $\{3, 4\}$, $\{2, 5\}$, and $\{6, 7\}$ are initially prepared as EPR pairs, $|\text{EPR}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

hole [202, 213]. Thus, the fidelity of quantum teleportation provides a fail-safe diagnostic for true quantum information scrambling.

Unlike a direct measurement of OTOCs, this protocol can explicitly distinguish scrambling from either decoherence or a mismatch between forward and backward time-evolution (i.e. the encoding and decoding unitaries: $\hat{U}$ and $\hat{U}^*$). Moreover, the success probability of Bob's projective measurement represents an independent measure of the average experimental value of the OTOC, which includes the effects of both noise and decoherence [518]. By comparing the teleportation fidelity and the success probability, one can quantitatively and unambiguously bound the amount of quantum scrambling in the unitary operation $\hat{U}$.

## 7.2 Distinguishing quantum information scrambling from noise

We experimentally implement the above teleportation protocol on a $7-$qubit fully-connected quantum computer [126] using a family of 3-qubit scrambling unitaries $\hat{U}_s$. Our quantum computer is realized with a crystal of trapped atomic $^{171}$Yb$^+$ion qubits, defined by the hyperfine "clock" states, as described in Appendix B. We confine nine ions in the linear ion trap and use the nearly equally-spaced middle seven ions for the circuit. We can drive single qubit gates on any of the seven qubits with a typical fidelity of 99.0(5)% and entangling two-qubit gates on any pair of qubits with a typical fidelity of 98.5(5)% (see Appendix B). Projective measurements of the qubits in any basis are performed with standard fluorescence techniques [356], with a qubit readout fidelity of 99.4(1)%. In combination, EPR state preparation and measurement can be performed with a fidelity of 98(1)% and are generated by a compiler that pieces together native, one- and two-qubit gates to produce the desired gates in a modular fashion [126].

A schematic of the experiment is depicted in Fig. 7.1. The first qubit is prepared in a designated single-qubit state $|\psi\rangle$ and acts as Alice's secret quantum state. We initialize 6 additional qubits in 3 non-adjacent EPR pairs, $|\text{EPR}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Two of these pairs correspond to the black hole entangled with Bob's quantum memory, and the third corresponds to Bob's ancillary EPR pair. We perform a scrambling unitary $\hat{U}_s$ on Alice's qubit and the two black hole qubits, and the decoding unitary $\hat{U}_d = \hat{U}_s^*$ on the two quantum memory qubits and one ancilla qubit. The explicit form of these unitaries and their decompositions into two-qubit gates is detailed in Appendix B. We complete the decoding protocol by projectively measuring any designated pair of qubits – a chosen Hawking-radiated qubit and Bob's complement of it – onto an EPR pair. In the absence of decoherence and errors, the probability $P_\psi$ of a successful projective measurement is directly related to the OTOC by:

$$P_\psi = \int d\phi \int d\hat{O}_H \left\langle \hat{O}_A^\dagger \hat{O}_H^\dagger(t) \hat{O}_A \hat{O}_H(t) \right\rangle, \tag{7.1}$$

where $\hat{O}_A \equiv |\psi\rangle\langle\psi|\phi$ acts on Alice's qubit, $\int d\phi \int d\hat{O}_H$ denotes an average over single-qubit

quantum states $\phi$ and local unitary operators $\hat{O}_H$ acting on the Hawking-radiated qubit, and $\hat{O}_H(t) = \hat{U}_s^\dagger \hat{O}_H \hat{U}_s$. If the EPR projection is successful, the decoding of Alice's quantum state can be quantified via the teleportation fidelity: $F_\psi = |\langle \varphi | \psi \rangle|^2$, where $|\varphi\rangle$ is the final state of the ancillary qubit. To characterize the nature of different scrambling unitaries, we repeat this protocol for initial states $|\psi\rangle \in \{|0_x\rangle, |1_x\rangle, |0_y\rangle, |1_y\rangle, |0_z\rangle, |1_z\rangle\}$, where $|0(1)_\alpha\rangle$ denotes the positive (negative) eigenstate of the Pauli operator $\sigma_\alpha$.

We begin by illustrating the challenge associated with interpreting conventional OTOC experimental measurements [163, 289, 325, 486]. In particular, we perform a control experiment with a non-scrambling unitary in the presence of deliberate experimental errors (Fig. 7.2a): specifically, we take $\hat{U}_s$ to be the identity operation, and introduce single-qubit rotational errors (parameterized by strength $\theta$) following the operation of $\hat{U}_s$, but not the decoding operation $\hat{U}_d$, creating a mismatch between forward and backward time-evolution. To allow for a fair comparison with the case of maximally scrambling unitaries, we implement the identity operator as a combination of one- and two- qubit gates of comparable complexity (and total number). As we increase the size of the mismatch error, we see that the average OTOC (as measured by $P_\psi$) decays, consistent with the expected sensitivity of the OTOC to experimental noise (Fig. 7.2). Crucially however, the decoding fidelity remains constant near 50%, the expected fidelity for an unknown qubit state, confirming that no scrambling has taken place.

Taken together, the measured teleportation fidelity ($F_\psi$) and the average OTOC ($P_\psi$) enables us to quantify the error-induced decay via an effective noise factor [518],

$$\mathcal{N} \equiv d_A \left[ (d_A + 1) \langle F_\psi P_\psi \rangle - \langle P_\psi \rangle \right], \tag{7.2}$$

where $d_A$ is the dimension of Alice's quantum state (in our case, $d_A = 2$) and the average is performed over all initial states $\psi$. Note that $\mathcal{N} = 1$ in the ideal case and $\mathcal{N} = 0.25$ (e.g. $1/d_A^2$) in the fully decohered case. To this end, the decay of this noise factor from unity indicates the presence of error-induced OTOC decay in our quantum circuit. As expected, the observed $\mathcal{N}$ decreases with increasing mismatch (Fig. 7.2e), reflecting the deliberate error-induced decay of the OTOC, despite the lack of any quantum scrambling dynamics. Moreover, the measured value of $\mathcal{N} \sim 0.60 - 0.75$ at zero-mismatch ($\theta = 0$) reflects the inherent errors in the experiment, which are expected from the many gates comprising the EPR preparation, unitary operation, and EPR measurement.

With the control experiment in hand, we now characterize information scrambling for a family of unitary operators $\hat{U}_s(\alpha)$ that continuously interpolate (Fig. 7.2b) between the identity operator ($\alpha = 0$) and a maximally scrambling unitary ($\alpha = 1$), as described in Appendix B. The gate decomposition of the $\hat{U}_s(\alpha)$ operation varies only in single-qubit rotations about the $z$-axis, which are performed classically with negligible error. Similar to the previous case, we observe the average OTOC to decay as the scrambling parameter, $\alpha$, is tuned from 0 to 1, as shown in Fig. 7.2c. However, unlike the case of the deliberate mismatch-error in Fig. 7.2a, the OTOC decay is accompanied by an increase in the decoding teleportation fidelity, indicating the presence of true quantum information scrambling.
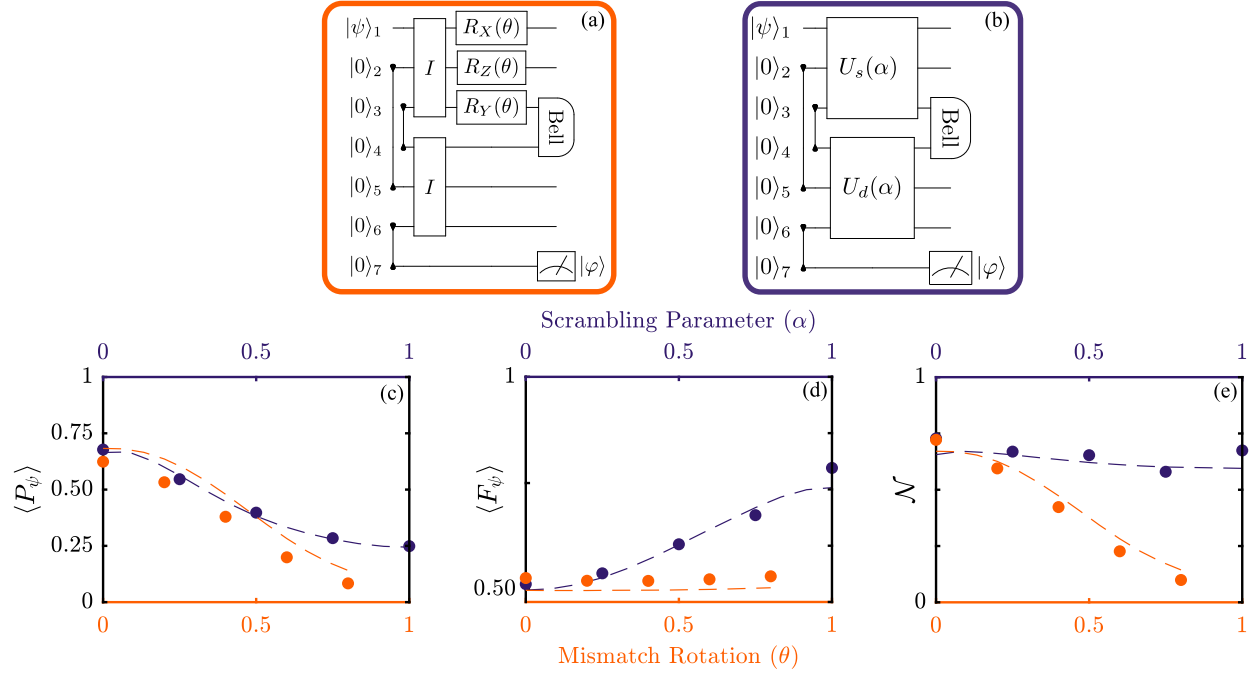
Figure 7.2: (a) Circuit designed to demonstrate that a mismatch between $\hat{U}_s$ and $\hat{U}_d$ naturally leads to the decay of the OTOC without enabling teleportation. Following the $\hat{U}_s = I$ operation, we perform three additional independent rotations $R_X$, $R_Y$, and $R_Z$ on the qubits by angle $\theta$. Accompanying data (orange) for the averaged successful projective measurement ($\langle P_\psi \rangle$), averaged teleporation fidelity ($\langle F_\psi \rangle$) and noise factor ($\mathcal{N}$) as a function of $\theta$ are depicted in panels (c-e). (b) Circuit designed to probe the OTOC and teleportation fidelity as a function of the scrambling parameter $\alpha$ with $\alpha = 0$ representing no scrambling and $\alpha = 1$ representing full scrambling. Accompanying data (purple) for $\langle P_\psi \rangle$, $\langle F_\psi \rangle$ and $\mathcal{N}$ as a function of $\theta$ are depicted in panels (c-e). (c-e) For the mismatch circuit shown in (a), we find that the teleportation fidelity remains near its minimal value, $\langle F_\psi \rangle \sim 0.5$, for all $\theta$, consistent with our expectation that scrambling is not occurring. However, one observes that the OTOC (as measured via $\langle P_\psi \rangle$) decays to nearly zero, which would nominally suggest scrambling. This is precisely the challenge with interpreting OTOC measurements as an indicator of scrambling in noisy experiments. Finally, as expected, we observe that the noise parameter $\mathcal{N}$ decays as the mismatch grows. For the tunable scrambling circuit shown in (b), we find that the teleportation fidelity increases as we increase our scrambling parameter $\alpha$. This increase in teleportation fidelity is accompanied by a decrease in the OTOC (as measured via $\langle P_\psi \rangle$), indicating that the OTOC's decay is caused, at least in part, by true scrambling dynamics. Finally, the noise parameter $\mathcal{N}$ remains relatively constant because the complexity and therefore experimental errors associated with implementing $\hat{U}(\alpha)$ are mostly $\alpha$-independent. Dashed lines represent theory curves that are obtained via numerical simulations of the circuit assuming a one-parameter coherent error model (see Appendix for details). Error bars indicating statistical uncertainties are smaller than the data points and are omitted.
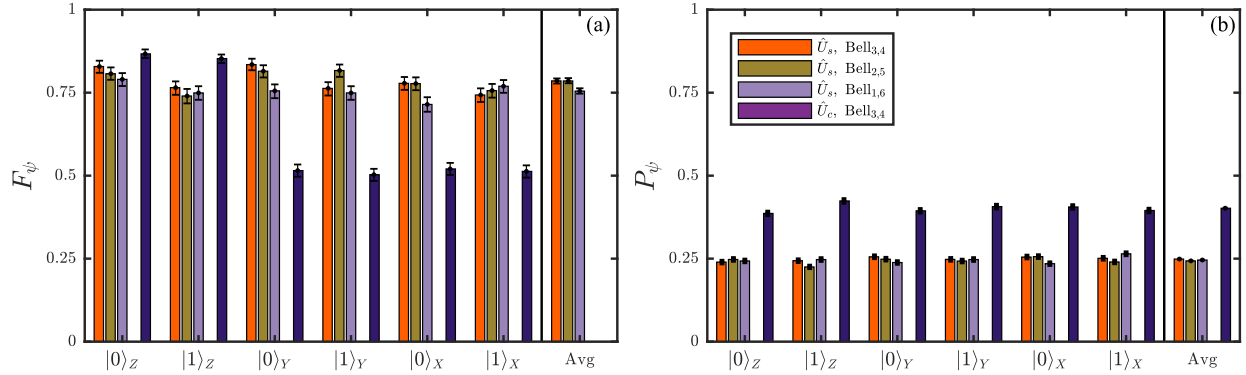
Figure 7.3: (a) Teleportation fidelities ($F_\psi$) for maximally scrambling ($\hat{U}_s(\alpha = 1)$) and classically scrambling ($\hat{U}_c$) unitaries are presented for all teleported states as well as all subsystems $\{3,4\}$, $\{2,5\}$, and $\{1,6\}$ that was used for the projective measurement (indicated as different bar colors). In the case of the maximally scrambling unitary, all basis states and all measurement Bell pairs lead to successful teleportation, demonstrating the full delocalization of Alice's quantum state. In the case of the classical scrambling unitary, we projectively measure on subsystem $\{3,4\}$. Only the z-basis states are successfully teleported. Data (for $\hat{U}_s$) averaged over all six teleported states is shown in the final column. Data is depicted with the same color scheme as in (b). b) Measurements of $P_\psi$ from the experiments described in (a). The probabilities averaged over all basis states constitutes the average experimental OTOC. Error bars indicate statistical uncertainties.

Measurement of a relatively constant noise factor confirms that the experimental error does not depend strongly on the parameter $\alpha$ and thus cannot fully account for the decay of the OTOC. In our system, the error scales with the number and type of gates, which are constant across the interpolation.

Using our experimentally measured noise factor $\mathcal{N}$, we can bound the true, scrambling-induced decay of the OTOC for error-free time-evolution via the unitaries $\hat{U}_s(\alpha)$. Assuming that extrinsic decoherence is negligible (i.e. that coherent errors dominate the experiment), we find that the ideal average OTOC can upper-bounded by [518]: $4\langle P_\psi \rangle^2/\mathcal{N}^2$. Therefore, we can experimentally upper-bound the value of the OTOC for the maximally scrambling unitary, $\hat{U}_s(\alpha = 1)$, by approximately 0.47(2).

## 7.3 Dependence on Hawking-radiated qubit and initial state

In order to demonstrate that our scrambling unitaries are indeed delocalizing information across the entire system, we show that teleportation succeeds independent of the chosen

subsystem partition. To do this, we select three different pairs of qubits to be projectively measured in the final decoding step, corresponding to the three qubits acted upon by the scrambling unitary $\hat{U}_s$. Decoding succeeds with a fidelity of $70 - 80\%$ for all projectively measured pairs and all initial states (Fig. 7.3). By contrast, the same protocol applied to the non-scrambling identity operator results in a nontrivial decoding fidelity for only a single pair, in which case the entire protocol reduces to the standard setup for quantum teleportation (Fig. 7.3). Taken together, these measurements demonstrate the full delocalization of Alice's initially local information (the secret qubit state) via the maximally scrambling unitary $\hat{U}_s$.

Intriguingly, time-evolution that is not maximally scrambling may nevertheless scramble some subset of information. For example, in strongly disordered systems, localization can lead to the scrambling of phase information but not the scrambling of population [345]; this situation would correspond to a sort of "classical scrambler" wherein teleportation only occurs for $z$-basis states. Such classical scramblers may provide insight into connections between quantum and classical chaos; moreover, measurements of state- and unitary-dependent scrambling may help to more precisely diagnose errors in digital many-body quantum simulations. To this end, we implement a classical scrambling unitary, $\hat{U}_c$, that completely delocalizes phase, but not population, information: $[\hat{U}_c, \sigma_z^i] = 0$ for all qubits $i$. This commutation implies that no measurement on a black hole qubit can distinguish the initial state $|\psi\rangle$ of Alice's qubit from the phase-reversed state $\sigma_z |\psi\rangle$, so that perfect teleportation is possible only for the classical eigenstates $|0_z\rangle$, $|1_z\rangle$ (see Appendix B for details), as observed in Fig. 7.3.

Thus far, our decoding protocols have all been probabilistic: they rely upon an EPR projective measurement to teleport the unknown quantum state. While the success probability of this EPR projection enables us to quantify the error-induced decay of the OTOC, one can also implement a deterministic version of the decoding protocol [517, 518] (Fig. 7.4a). The intuition behind this deterministic decoder is to perform a search for the EPR pair through an implementation of Grover's search algorithm [179] instead of post-selecting on a projective measurement. Scrambling remains the focus of this version; while the deterministic decoding fidelity can be lower bounded by 4-point OTOCs, their precise measurement actually corresponds to certain averages of higher-point OTOCs (in our case of a qubit input, 8-point OTOCs). Such higher-point OTOCs [436] can serve as more fine-grained measures of chaos beyond four-point OTOCs [187], have applications as probes of unitary 4-designs and may be able to diagnose higher-moments of quantum randomness [399]. Within the Grover-search variant of our decoding protocol, three states are deterministically decoded with an average fidelity of $77(2)\%$ (Fig. 7.4b). Finally, we perform a further variation of this protocol that reintroduces a projective measurement as a means to purify errors from the experiment. Here, the same three initial states were decoded with an average fidelity of $78(2)\%$ (Fig. 7.4b). That this purification leads to the same teleportation fidelity despite the additional gate depth (Fig. 4a, purple box) suggests that the fidelity of the EPR Grover search is roughly $85\%$. To demonstrate the generality of our approach, these Grover-based protocols were performed with a different class of maximally scrambling unitaries than the previous probabilistic protocol, as described in Appendix B.
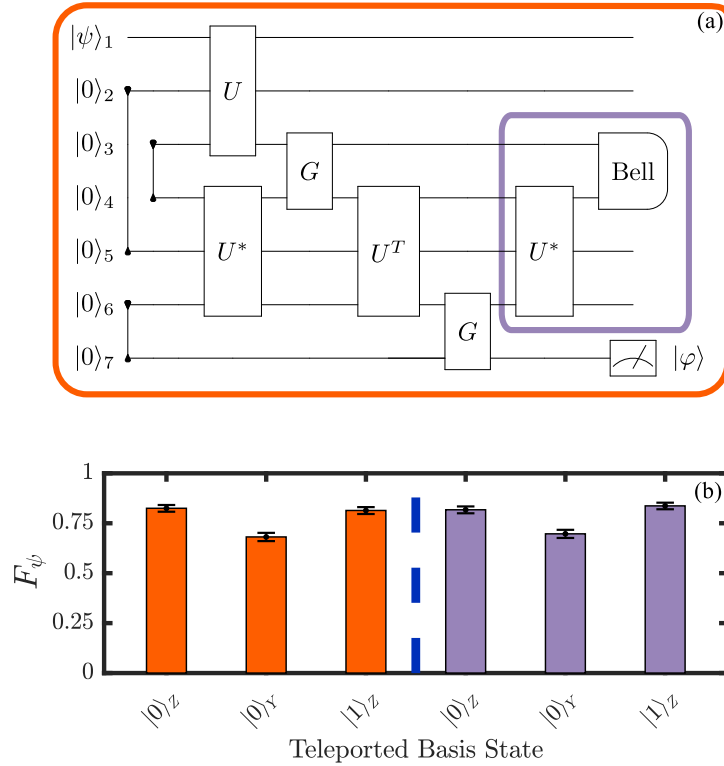
Figure 7.4: (a) Circuit diagram for our deterministic teleportation scheme that utilizes a built in Grover's search protocol. Ideally, this search finds the desired EPR state with perfect fidelity; by adding in post-selection (purple box), we can quantify the performance of the search. (b) Depicts the measured teleportation fidelity for different initial states utilizing the Grover search protocol, both with and without post-selection. The average fidelity without post-selection (orange) is $77(2)\%$ and with post-selection (purple) is $78(2)\%$.

## 7.4 Discussion and outlook

Our work opens the door to a number of intriguing future directions, many of which are explored later in this dissertation. First, by experimentally scaling to larger circuits, one should be able to study the scrambling dynamics of many-body quantum systems of interest (e.g. Hamiltonian systems), complementary to the Clifford scramblers studied here. Indeed, see Ch. 9 for an in-depth exploration of teleportation across wide classes of many-body dynamics in a closely related quantum circuit. Second, while we have focused on the challenges of distinguishing between scrambling and decoherence in this work, our protocol suggests that scrambling circuits may provide a near-term method to benchmark and verify large-scale quantum simulations [105]. In particular, by probing the teleportation fidelity as a

function of both input state complexity and circuit, one can imagine identifying certain many-body noise mechanisms that would be invisible to typical single- and two-qubit randomized benchmarking methods. To do so requires a precise knowledge of how experimental noise affects teleportation and scrambling dynamics more generally. We provide the latter in Ch. 10, and the former in Ch. 9. Finally, it has recently been speculated that quantum simulations can become a natural way to probe the physics of quantum gravity [446]. These arguments hinge on the holographic principle, which suggests that any bulk quantum gravity phenomena can in fact be encoded on the boundary of a quantum mechanical system. To this end, it would be intriguing to modify our teleportation experiment to send a probe qubit into the interior of a black hole before retrieving it via a dynamical back-reaction, thereby providing a way to probe the interior geometry of a black hole. Such an experiment is at the heart of the teleportation protocol studied in Ch. 9.

# Chapter 8

# Quantum information scrambling on a qutrit superconducting quantum processor

In this Chapter, we present an experimental measurement of quantum information scrambling in a quite different physical system: a superconducting qutrit processor. As in the previous Chapter, we utilize a particular quantum teleportation circuit [517, 518] to robustly characterize the scrambling dynamics in the presence of experimental noise. In this Chapter we also verify that the experimental dynamics through an alternative method: quantum process tomography. This method takes advantage of the fast cycle times of superconducting platforms to provide a complete characterization of the dynamics. However, unlike the teleportation circuit, it is not scalable to larger system sizes.

The work in this Chapter also serves a dual purpose: as the first demonstration of qutrit-based superconducting quantum processor. Indeed, while the majority of current generation quantum processors are based on qubits, qutrit-based (and more generally qudit-based [307]) systems have long been known to exhibit significant advantages in the context of quantum technology: they have been touted for their small code-sizes in the context of quantum error correction [87], high-fidelity magic state distillation [88], and more robust quantum cryptography protocols [47, 75]. To date, advantages of individual qutrits have been explored experimentally in fundamental tests of quantum mechanics [276] and in certain quantum information protocols [119, 147, 272, 400], while entanglement between qutrits has been demonstrated in probabilistic photonic systems [223, 308]. A qutrit platform capable of implementing deterministic high-fidelity gates would be a powerful tool for both quantum simulation and information processing.

In what follows, we present a prototypical multi-qutrit processor based on superconducting transmon circuits (Fig. 8.1), and—as a proof-of-principle demonstration—use it to (i) perform a maximally-scrambling two-qutrit unitary and (ii) verify its scrambling behavior using a five-qutrit quantum teleportation circuit. The superconducting qutrit processor we develop here features long coherence times; multiplexed readout of individual qutrits; fast,
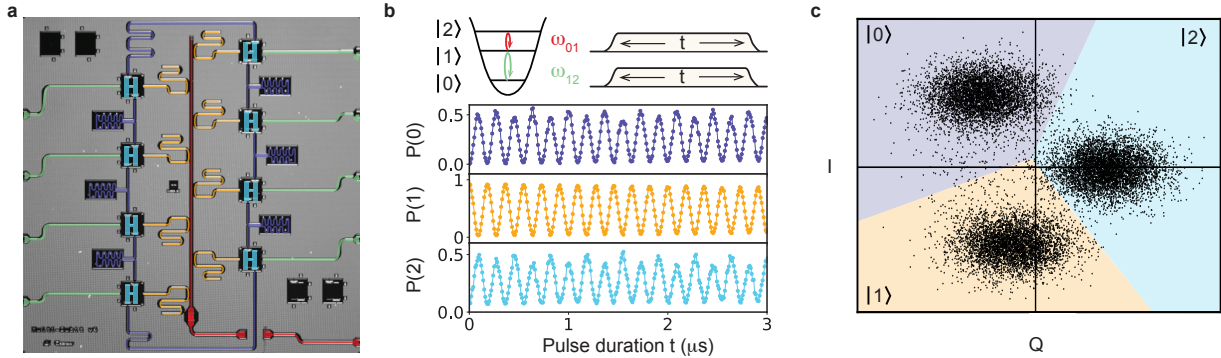
Figure 8.1: Superconducting qutrit processor. **(a)** Optical micrograph of the five-transmon processor used in this experiment. Transmon circuits (light blue) couple to an integrated Purcell-filter and readout bus (red) via individual linear resonators (gold), enabling multiplexed state measurement. Exchange coupling between nearest-neighbor transmons is mediated by resonators (purple), while microwave drive lines (green) enable coherent driving of individual qubits. **(b)** Coherent Rabi dynamics of a single qutrit induced by simultaneous microwave driving at frequencies $\omega_{01}$ and $\omega_{12}$. Achievable Rabi frequencies are in the range of tens of MHz, three orders of magnitude faster than decoherence timescales. **(c)** Example single-shot readout records of an individual qutrit, generally achievable with fidelities above 0.95. This is largely limited by decay during readout.

high-fidelity single-qutrit operations; and two types of two-qutrit gates for generating entanglement. Using this gateset on our processor, we construct a maximally scrambling qutrit unitary and characterize it using quantum process tomography. Finally, to demonstrate the ability of our platform to perform genuine multi-qutrit algorithms, we perform a five-qutrit teleportation protocol which serves as an additional verification of genuine two-qutrit scrambling behavior. This protocol is inspired by the Hayden-Preskill variant of the black-hole information paradox [202, 517]. While we have chosen quantum scrambling as a demonstration of our processor, our work—the first realization of a multi-qutrit algorithm—opens the door more broadly to the experimental study of quantum information processing utilizing qutrit-based logic.

## 8.1   Qutrit superconducting quantum processor

Nearly all current quantum processors are based on collections of the simplest possible quantum system: qubits, consisting of two states per site. In contrast, qudits – featuring $d > 2$ states – can store an exponentially greater amount of information compared to qubits and can therefore, in certain cases, implement quantum algorithms using smaller systems and fewer multi-site entangling gates [80, 169]. In cavity QED experiments, continuous variable

quantum information has been encoded [342, 422], protected [355] and entangled  [162] in linear cavity modes, by coupling them to an artificial non-linear atom.  Encoding discrete quantum information in an intrinsically non-linear system has the advantage that energy transitions have distinct frequencies that can be driven individually and with faster gates.

For qutrit systems specifically, many quantum information protocols have been proposed that would yield a significant advantage over qubits [33, 47, 75, 87, 88].  Qutrits have been successfully realized in various physical degrees of freedom including the polarization state of multiple photons [275], spin $S = 1$ states of solid-state defect centers [137], hyperfine states in trapped atoms and ions [411, 433] and the lowest energy states of an anharmonic oscillator in superconducting circuits [56].  However, no platform to date has executed ternary quantum circuits, and a universal two-qutrit gate has not yet been demonstrated.

Our superconducting qutrit processor features eight transmons [258, 426] connected in a nearest-neighbor ring geometry (Fig. 8.1a) of which we use five (denoted $Q_1, \ldots, Q_5$) to realize quantum circuits.  Each transmon encodes a single qutrit and is coupled both to a dedicated microwave control line (for performing gates, Fig. 8.1b) and its own readout resonator (for state measurement, Fig. 8.1c).  Transmons are quantum nonlinear oscillators that can be operated as qubits, using only their two lowest-lying energy states $|0\rangle$ and $|1\rangle$. While their higher energy states ($|2\rangle$, $|3\rangle$, etc.) in principle enable transmons to be operated as higher-dimensional qutrits or qudits [56] , the experimental implementation of multi-qutrit algorithms is challenging due to a lack of two-qutrit entangling gates and increased noise associated with the higher transmon states.

## High-fidelity single-qutrit operations

In order to implement high-fidelity single-qutrit operations in transmons, one must overcome multiple sources of noise and coherent errors that are naturally introduced upon including the $|2\rangle$ state of transmon in the computational subspace.  The primary such sources include:

($i$) *Relaxation due to spontaneous emission*—Energy relaxation or $T_1$ processes in transmon qubits cause unwanted, incoherent transitions from the $|1\rangle$ state to the $|0\rangle$ state, and arise from lossy interfaces or electromagnetic channels.  For transmon qutrits, the addition of the $|2\rangle$ state introduces another decay channel, from $|2\rangle \to |1\rangle$[1].  Due to bosonic enhancement, in which the spontaneous emission rate scales linearly with photon number, the time constant associated with $|2\rangle \to |1\rangle$ decay is roughly half that of the $|1\rangle \to |0\rangle$ transition.

To address this, we use a high-quality fabrication process and careful microwave engineering to minimize the effects of energy relaxation.  The fabrication recipe, detailed in Appendix C, is optimized to remove lossy oxide layers at the substrate-metal and metal-metal interfaces through three separate buffered-oxide etches and an ion-mill step.  In addition, finite-element simulations are used to identify and mitigate each loss channel associated

---

[1]In principle, one would also have to consider the transition from $|2\rangle \to |0\rangle$, but this channel is largely suppressed, due to parity, in transmons.

with microwave radiation. Specific mitigation strategies include: an integrated, broadband
Purcell filter suppressing leakage into the readout bus (Fig. 8.1a); readout and bus resonator
geometries designed to shift the frequencies of lossy higher modes away from those of qutrits
and resonators of band; and extensive wire-bonding connecting the ground planes on either
sides of the readout bus, coupling resonators, readout resonators, and control lines. As a
results of these techniques, the average $T_1$ times on the chip are $56.0 \pm 10$ $\mu$s for the $|1\rangle \to |0\rangle$
transition, and $34.8 \pm 4$ $\mu$s for the $|2\rangle \to |1\rangle$ transition.

($ii$) *Dephasing due to the charge-sensitivity of higher transmon levels*—The transmon
was originally developed to reduce the dependence of energy levels on offset charge noise.
While typical values of transmon parameters (i.e. the Josephson energy $E_J$ and the charging
energy $E_C$) result in low charge-noise sensitivity in the qubit subspace, the charge dispersion
increases rapidly with increasing energy levels [258, 369]. For qutrits specifically, the charge
dispersion of the $|2\rangle$ state is at least an order of magnitude greater than that of the $|1\rangle$ state,
resulting in a charge-limited dephasing time ten times lower than that of the qubit subspace.

To mitigate charge noise in the $|2\rangle$ state, we tune the transmon parameters even further
than typical into the "transmon regime": specifically, we choose a ratio of Josephson and
charging energies $E_J/E_C \approx 73$ (typically, this ratio is near 50 for transmon qubits [40,
235, 414]). This reduces the charge dispersion of the $|2\rangle$ and $|1\rangle$ states to 12 kHz and 250
Hz, respectively. As a result, we realize dephasing times, averaged over the five qutrits, of
$T_2^* = (39 \pm 21)$ $\mu$s [$(14 \pm 5)$ $\mu$s] for the $|0\rangle \to |1\rangle$ [$|1\rangle \to |2\rangle$] transitions, which can be further
extended with a Hahn-echo to $T_{2,\mathrm{echo}} = (61.2 \pm 11)$ $\mu$s [$(28 \pm 5)$ $\mu$s].

($iii$) *Increased cross-talk due to frequency crowding*— If pulses, intended to apply unitary
operations on a single qutrit, are not well localized in space to the desired qutrit, they can
induce unintended unitary operations on nearby qutrits. This 'cross-talk' is increasingly
troublesome when including the $|2\rangle$ state, as the frequency spectrum of state transitions
becomes more crowded due to the inclusion of the $|1\rangle \to |2\rangle$ transition frequencies.

On our processor, we find significant cross-talk between the microwave drive lines for
most transmons. When driving Rabi oscillations on a given qutrit, this cross-talk has two
unwanted effects:

1. All other qutrits are off-resonantly driven. Depending on the relative frequencies be-
   tween the qutrits, this can manifest either as an unwanted shift in state populations
   or as an AC Stark shift.

2. Microwave fields leaking onto neighboring qutrit(s) will result in an unwanted cross-
   resonance interaction, making the desired Rabi frequency dependent on the state of
   the neighboring qutrit(s) [394].

For a given drive frequency, our cross-talk can be characterized in terms of a five-by-five
complex-valued matrix $C(\omega)$, which relates the field amplitudes $\mathbf{o}(\omega)$ seen by each of the
five qutrits to the input field amplitudes $\mathbf{i}(\omega)$ on each drive line: $\mathbf{o}(\omega) = C(\omega)\mathbf{i}(\omega)$. While
this cross-talk matrix strongly depends on frequency, it does not exhibit any nonlinearity at
the powers used in the experiment. Therefore, we compensate for the cross-talk by inverting

the matrix $C(\omega)$ at each frequency to yield linear combinations of drive lines which result in non-zero fields only at a single qutrit.

To measure the cross-talk matrix $C(\omega)$ of our system, we drive two control lines simultaneously, and, for each pair of driven control lines, determine the relative amplitudes and phases of driving that cancel the resulting field at each qutrit on the chip. Depending on the relative frequencies between the drive field and the qutrit transition, we use either an AC Stark shift or a Rabi oscillation as the diagnostic of an unwanted microwave field. This measurement is repeated for each of ten drive frequencies of interest (i.e. the $|0\rangle \leftrightarrow |1\rangle$ and $|1\rangle \leftrightarrow |2\rangle$ transition frequencies of all five qutrits), each pair of lines, and each qutrit. While this method proved effective for our system, it can become prohibitively measurement-intensive for future processors with hundreds of qudits. Looking forward, in order to scale to such systems it will be crucial to pinpoint and mitigate the source of this cross-talk at the hardware level.

With the above three issues—relaxation, dephasing, and cross-talk—taken care of, we now turn to our realization of high-fidelity single-qutrit gates. In qutrits, these require the combination of a microwave pulse (to perform the rotation in the relevant subspace) and a virtual phase-gate. This phase gate is necessary because a microwave field that addresses, e.g. the $|0\rangle \rightarrow |1\rangle$ transition, also generates a relative phase $e^{i(\phi_G + \phi_D)} |2\rangle$ on the idling state, where $\phi_G$ is the geometric (Berry) phase associated with the rotation in the $\{|0\rangle, |1\rangle\}$ subspace and $\phi_D$ is the dynamical phase due to the AC Stark shift of the idling $|2\rangle$ state. These phases represent logical Z errors in the computational subspace, which we measure and correct.

Single-qutrit gates in our system are performed within 30 ns (see Appendix C for the universal set of gates used in our experiment). We characterize them by performing qubit-based randomized benchmarking in two different qubit subspaces [378]. These yield fidelities on par with state-of-the-art qubit processors: $f_{01} = 0.9997 \pm 0.0001$ and $f_{12} = 0.9994 \pm 0.0001$ for gates within the $\{|0\rangle, |1\rangle\}$ and $\{|1\rangle, |2\rangle\}$ subspace, respectively. While this benchmarking method demonstrates that single-qutrit coherence is no longer a major bottleneck for high-fidelity operations, it should be noted that this method is not sensitive to certain sources of errors, including phase errors in the idle state and multi-qutrit errors. In the future, a full characterization of qutrit operations will require the development of genuine qutrit randomized benchmarking protocols.

Multiplexed, individual qutrit readout is performed dispersively via a linear readout resonator coupled to each transmon. As shown in Fig. 8.1c, with a suitable choice of the readout frequency and amplitude all three states of the transmon can be resolved in a single shot. Averaged over all qutrits, our readout fidelity is $F_{\text{avg}} = 0.96 \pm .02$. In ensemble measurements (e.g. when performing tomography), we can correct for this readout imperfection by storing a 'confusion matrix' $M_{ij}$ of the conditional probabilities to measure a state $|i\rangle$ given an actual state $|j\rangle$. Applying its inverse to the measurement results allows one to infer the actual state populations.

## Two-qutrit entanglement

Having outlined our implementation of high-fidelity single-qutrit operations, we next demonstrate two methods for generating controllable two-qutrit entangling gates in our system: the first based on the cross-resonance interaction [102, 394], and the second based on the longitudinal, or cross-Kerr, interaction. Both of these operate on neighboring qutrits, giving our full eight-qutrit processor an effective ring geometry; restricting to five qutrits, this reduces to a linear geometry. The two types of interaction give rise to two distinct forms of entangling gates: the cross-resonance interaction leads (up to a phase) to a conditional subspace rotation, while the cross-Kerr interaction leads to a controlled-SUM gate—the qudit analogue of the controlled-NOT gate. We now describe both of these gates in detail.

*Cross-resonance gate*—The cross-resonance interaction has been used extensively in superconducting qubit systems to generate high-fidelity controlled-NOT gates [234, 236, 435]. The origin of the interaction is a weak transverse coupling between two qubits that are highly off-resonant such that the coupling is almost negligible. When one of the two systems (the control) is driven at the frequency of the other (the target), the combination of the coupling and the drive induces Rabi oscillations on the target qubit, with a Rabi frequency that depends on the state of the control qubit. This results in two Rabi frequencies, $\omega_0$ and $\omega_1$, for the two control qubit states, $|0\rangle$ and $|1\rangle$, both of which also depend on the drive power. This interaction is naturally described by the Hamiltonian

$$H_{cr}^{\text{qubit}} = \omega_0 \, |0\rangle\langle 0| \otimes \sigma_x + \omega_1 \, |1\rangle\langle 1| \otimes \sigma_x. \tag{8.1}$$

Higher-order deviations from this Hamiltonian can largely be mitigated with dynamical decoupling, local rotations, or pulse shaping. This Hamiltonian is locally equivalent to a $\sigma_z \otimes \sigma_x$ interaction, and periodically entangles and disentangles qubits at a rate $\Delta\omega = |\omega_0 - \omega_1|$. When the two oscillations are completely out of phase ($t\Delta\omega = \pi$), the resulting unitary

$$U_{\text{CNOT}} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma_x \tag{8.2}$$

is locally equivalent to a controlled-NOT gate.

On our chip, the cross-resonance interaction arises from the static capacitive coupling between neighboring transmons. To extend the cross-resonance gate to *qutrit* systems, we must consider how this interaction behaves in the full two-qutrit Hilbert space. We will show that we can realize a similar gate to that of qubit systems, where we swap the $|0\rangle$ and $|1\rangle$ populations of the target qutrit conditional on the control qutrit being in the $|1\rangle$ state.

The dynamics of the two-qutrit cross-resonance interaction are shown in Fig. 8.2(a-b). Compared to qubits, we have two additional considerations: the effect of the cross-resonance drive on the $|2\rangle$ state of the target qutrit, and the Rabi oscillations induced on the target qutrit by the $|2\rangle$ state of the control qutrit. The first effect results in an overall phase being applied to the $|2\rangle$ state of the target via the AC Stark shift (Fig. 8.2b), since the effective Rabi drive is off-resonant with the $|0\rangle, |1\rangle \rightarrow |2\rangle$ transitions of the target qutrit (in contrast to the $|0\rangle \rightarrow |1\rangle$ transition, with which it is resonant). The second effect extends the qubit
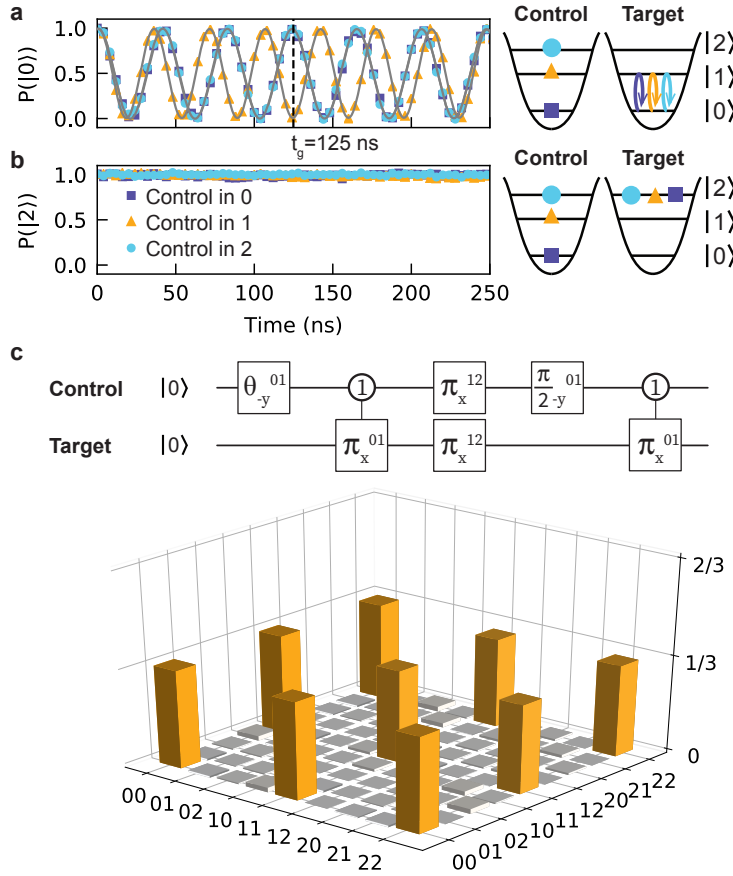
Figure 8.2:   Two-qutrit EPR pair generation via the cross-resonance interaction **(a)** Nearest-neighbor qutrits coupled by an exchange interaction can be entangled via the cross-resonance effect, where one qutrit (the control) is microwave-driven at the $|0\rangle \leftrightarrow |1\rangle$ transition frequency of the other (the target). Resulting Rabi oscillations of the target qutrit exhibit a Rabi frequency dependent on the state of the control qutrit. Here, we drive the control with a field whose amplitude is chosen to make the Rabi frequencies corresponding to control states $|0\rangle$ and $|2\rangle$ identical, resulting in a unitary operation which, after 125 ns, interchanges states $|0\rangle$ and $|1\rangle$ of the target qutrit when the control qutrit is the $|1\rangle$ state. **(b)** When the target qutrit is in the $|2\rangle$ state, the cross-resonance interaction is off-resonant and does not affect the population. **(c)** Top: Sequence used to prepare an EPR pair with two applications of the cross-resonance gate. $\pi_x^{01}$ represents be a rotation of $\pi$ about the x axis in the $\{|0\rangle, |1\rangle\}$ subspace of the qutrit. The number in the circle for the two-qutrit gate represents the condition (state of the control qutrit) for which a gate is applied on the target. Bottom: The density matrix (reconstructed via state tomography) of the resulting EPR pair, with a state fidelity of $F_{\mathrm{EPR}} = 0.98 \pm 0.002$.

cross-resonance Hamiltonian to qutrits according to

$$
\begin{aligned}
H_{cr}^{\text{qutrit}} &= \omega_0 \, |0\rangle\langle 0| \otimes \pi_x^{01} + \omega_1 \, |1\rangle\langle 1| \otimes \pi_x^{01} \\
&\quad + \omega_2 \, |2\rangle\langle 2| \otimes \pi_x^{01},
\end{aligned} \tag{8.3}
$$

where we have added the (drive-dependent) frequency $\omega_2$, which induces Rabi oscillations
in the target qutrit when the control qutrit is in the $|2\rangle$ state. In the above Hamiltonian,
$\pi_x^{01}$ is the $3 \times 3$ Gell-Mann matrix that swaps the $|0\rangle$ and $|1\rangle$ states, and does nothing to
the $|2\rangle$ state. We emphasize that this model for the interactions, while physically motivated
and consistent with our experimental observations, is only approximate. Much like for the
two-qubit cross-resonance gate [312, 466], further study is needed to develop a precise model
of the dynamics.

The qutrit cross-resonance interaction Eq. (8.3) involves three distinct Rabi frequencies,
corresponding to the three states of the control qutrit (Fig. 8.2a). At low drive power, these
frequencies scale linearly with the drive amplitude, while at higher powers the dependence
becomes nonlinear. We utilize this nonlinearity to choose a drive power such that two of the
Rabi frequencies are equal, $\omega_0 = \omega_2$. Next, similar to the qubit implementation, we choose
the gate time $t_g$ according to $t_g |\omega_0 - \omega_1| = \pi$. Up to state-dependent phases, this gives rise
to a conditional-$\pi$ gate (also known as a generalized controlled-X gate [133]):

$$
\begin{aligned}
U_{\text{C}\pi} &= (|0\rangle\langle 0| + |2\rangle\langle 2|) \otimes I \\
&\quad + |1\rangle\langle 1| \otimes (|0\rangle\langle 1| + |1\rangle\langle 0| + |2\rangle\langle 2|).
\end{aligned} \tag{8.4}
$$

Our experiment features a frequency difference $\omega_0 - \omega_1 = 4$ MHz and a corresponding gate
time $t_g = 125$ ns. During this gate, simultaneously with the cross-resonance drive applied
to the control qutrit, we apply a concurrent drive to the target qutrit which ensures that
it ends up in a computational basis state at $t_g$. This target drive, with Rabi frequencies in
the tens of MHz, has the additional benefit of suppressing dephasing in the target qutrit's
$\{|0\rangle, |1\rangle\}$ subspace during the gate [146].

In order to benchmark the performance of the conditional-$\pi$ gate, we use two applications
of the gate to create a maximally entangled two-qutrit EPR pair:

$$
|\text{EPR}\rangle = \frac{|00\rangle + |11\rangle + |22\rangle}{\sqrt{3}}. \tag{8.5}
$$

We measure the success of EPR preparation using qutrit state tomography [56], enabled by
our high-fidelity local qutrit pulses and qutrit readout. From this we reconstruct the density
matrix of our final state (Fig. 8.2c) and compute an EPR fidelity of $F_{\text{EPR}} = 0.98 \pm 0.002$,
which we observe to be mostly limited by decoherence.

Finally, while the conditional-$\pi$ gate is fast and capable of generating high-fidelity en-
tangled states, it acts only in a two-qubit subspace of the full two-qutrit Hilbert space.
Consequently, despite the conditional-$\pi$ gate being universal for qutrit computation when
combined with arbitrary single qutrit gates [134], implementing general two-qutrit unitaries

|  | $Q_1/Q_2$ | $Q_2/Q_3$ | $Q_3/Q_4$ | $Q_4/Q_5$ |
|---|---|---|---|---|
| $\alpha_{11}$ | -279 | -138 | -276 | -262 |
| $\alpha_{12}$ | 160 | 158 | -631 | -495 |
| $\alpha_{21}$ | -528 | -335 | 243 | -528 |
| $\alpha_{22}$ | -743 | -342 | -748 | -708 |

Table 8.1: Measured cross-Kerr interaction strengths between nearest-neighbor pairs of transmons, in units of kHz.

will require a high number of such gates. This can be undesirable given the finite error associated with each gate. To address this, in what follows we introduce a *second* qutrit entangling gate, the controlled-SUM gate, which we will implement via the cross-Kerr interaction instead of the cross-resonance interaction. Ultimately, we will use both gates to perform the full five-qutrit teleportation protocol, applying the conditional-$\pi$ gate to EPR preparation/measurement and the more powerful controlled-SUM gate to a complex two-qutrit scrambling operation.

*Cross-Kerr gate*—As neighboring qutrits in our processor are capacitively coupled to one another, they are subject to an "always-on" cross-Kerr entangling interaction, which we harness to perform controlled-phase and controlled-SUM gates (the two are locally equivalent). The always-on cross-Kerr interaction, analogous to the familiar dispersive interaction between a qubit and cavity, is modeled by the dispersive cross-Kerr Hamiltonian

$$H_{\text{cK}} = \alpha_{11} |11\rangle\langle 11| + \alpha_{12} |12\rangle\langle 12| + \tag{8.6}$$
$$\alpha_{21} |21\rangle\langle 21| + \alpha_{22} |22\rangle\langle 22| ,$$

where we work in a rotating frame that cancels the Hamiltonian's action on the $|0j\rangle , |j0\rangle$ states. The cross-Kerr coefficients $\alpha_{ij}$ represent the rate at which phases accumulate on the $|ij\rangle$ state during idling. Their values are sensitive to both the magnitude of the capacitive coupling between transmons as well as the transmon spectra. While in principle one could design a processor with pre-specified values of $\alpha_{ij}$, this is impractical given current technology because the transmon spectra depend on the critical current of their Josephson junction, which cannot be precisely set in fabrication [269]. To maintain robustness against such variability, we therefore design the cross-Kerr gate to work regardless of the particular values of the coefficients. For reference, these values for our particular processor are shown in Table C.2.

We will now demonstrate how to combine a generic cross-Kerr Hamiltonian with single-qutrit rotations to form a controlled-SUM gate. The foundation of our construction is the local equivalence between the controlled-SUM gate:

$$U_{\text{CSUM}} = \sum_{n=1}^{d} |n\rangle\langle n| \otimes X^n \tag{8.7}$$

and the controlled-phase gate $U_{C\phi}$:

$$U_{C\phi} = \sum_{n=1}^{d} |n\rangle\langle n| \otimes Z^n, \tag{8.8}$$

which we will soon show is achievable via the cross-Kerr interaction. Here

$$X = \sum_{j=1}^{d} |j+1\rangle\langle j|, \quad Z = \sum_{j=1}^{d} e^{i2\pi j/d} |j\rangle\langle j| \tag{8.9}$$

are the generators of the set of qudit Pauli operators. The two gates are related via conjugation by a single-qutrit Hadamard $H$ gate:

$$(I \otimes H^\dagger)U_{C\phi}(I \otimes H) = U_{\text{CSUM}}, \tag{8.10}$$

which, as for qubits, interchanges the Pauli $X$ and $Z$ operators. This realization also easily generalizes to slight modifications of the controlled-SUM gate: one can realize instead a controlled-MINUS gate by inverting the Hadamard, and one can interchange the control and the target qudits by switching which qudit is Hadamard conjugated.

Specifying to qutrits, the controlled-phase gate $U_{C\phi}$ imparts a phase $2\pi/3$ to the $|11\rangle$, $|22\rangle$ states and the opposite phase $-2\pi/3$ to the $|12\rangle$, $|21\rangle$ states. We have developed two methods for implementing this gate via the cross-Kerr interaction. The first, described in what follows, is conceptually simpler and uses only four single-qutrit pulses, but is prone to local dephasing and to errors arising from a nonzero cross-Kerr interaction with qutrits outside the target-control pair. The second method addresses these errors via dynamical decoupling in the specific context of the five-qutrit teleportation protocol, and is described in conjunction with the protocol in Section 8.2.

The essential concept behind these constructions is to intersperse time-evolution under the native cross-Kerr Hamiltonian with single-qutrit pulses exchanging the $|1\rangle$ and $|2\rangle$ states. The simplest construction uses four segments of cross-Kerr evolution. Denoting a swap pulse on qutrit $q$ as $\pi_{x,q}^{12}$, the total pulse sequence is given by the decomposition:

$$\begin{aligned}
U_{C\phi} = e^{-iT_A H_{cK}} \cdot \Pi_0^{12} \cdot e^{-iT_B H_{cK}} \\
\cdot \Pi_1^{12} \cdot e^{-iT_C H_{cK}} \cdot \Pi_0^{12} \cdot e^{-iT_D H_{cK}} \cdot \Pi_1^{12},
\end{aligned} \tag{8.11}$$

where $T_A$, $T_B$, $T_C$, $T_D$ denote the time durations of each segment of cross-Kerr evolution. To determine these times from the interaction coefficients, note that each evolution segment serves to add a phase $\phi_{ij} = \alpha_{ij}T$ to the $|ij\rangle$ state (here, $i, j \in \{1, 2\}$), while the swap pulses interchange the indices $1 \leftrightarrow 2$ on the acted upon qutrit. Our choice of swap gates guarantees that each state spends exactly one segment under each interaction coefficient, and returns to itself at the end of the gate. The unitary thus amounts to a phase applied to each state equal to a linear combination of the four evolution times, with the transformation matrix

determined by the interaction coefficients. For generic interaction coefficients this linear transformation from interaction times to applied phases is full rank, which guarantees that the method can in principle generate *any* combination of two-qutrit phases.

With our particular interaction coefficients, this controlled-phase gate implementation takes a total time $\sim 1.5$ $\mu$s for each of the qutrit pairs $(Q_1, Q_2)$ and $(Q_3, Q_4)$. We characterize the performance of the full controlled-SUM gate via quantum process tomography on the full two-qutrit subspace (Fig. 8.3b) and achieve a fidelity of 0.889, primarily limited by decoherence occurring throughout the cross-Kerr time evolution.

Combined with the single-qutrit control demonstrated in the previous section, each of the conditional-$\pi$ gate and the conditional-SUM gate enable universal quantum computation on our qutrit processor [134]. Having both gates on hand provides additional flexibility in gate compilation as the different entangling gates may be more adept in different scenarios.

## 8.2 Qutrit quantum information scrambling

We now turn to proof-of-principle experiments demonstrating the capabilities of our qutrit processor, specifically by demonstrating and verifying quantum scrambling. We will demonstrate the implementation of a two-qutrit scrambling unitary, and verify that it scrambles using a five-qutrit quantum teleportation algorithm. Quantum scrambling, loosely defined as the delocalization of initially local operators induced by many-body time dynamics, is an area of active current research; it is likely that quantum simulations performed on future processors will deepen our understanding of this phenomenon and many-body quantum dynamics more generally. To date, experimental realizations of scrambling have been performed with qubits; however, many interesting scrambling-related phenomena occur in higher-dimensional systems; see, e.g., [182, 341, 540]. Using quantum scrambling to demonstrate the capabilities of our processor thus also represents a step towards such simulations of these high-dimensional composite systems.

Specifically, we implement a simple Clifford scrambling unitary $U_s$, which effects the following permutation of the nine two-qutrit computational basis states [518]:

$$U_s \left| m, n \right\rangle = \left| 2m + n, m + n \right\rangle . \tag{8.12}$$

As described in Ref. [518], the scrambling behavior of this unitary can be seen in the effect it has on the single-qutrit Pauli operators $X$ and $Z$:

$$U \left( Z \otimes I \right) U^\dagger \;=\; Z \otimes Z^2 \tag{8.13}$$
$$U \left( I \otimes Z \right) U^\dagger \;=\; Z^2 \otimes Z^2 \tag{8.14}$$
$$U \left( X \otimes I \right) U^\dagger \;=\; X^2 \otimes X \tag{8.15}$$
$$U \left( X \otimes I \right) U^\dagger \;=\; X \otimes X. \tag{8.16}$$

Each single-qutrit Pauli operator is transformed into a two-qutrit operator, as required by scrambling. Additionally, the transformations above show that $U_s$ is also a Clifford scrambler,
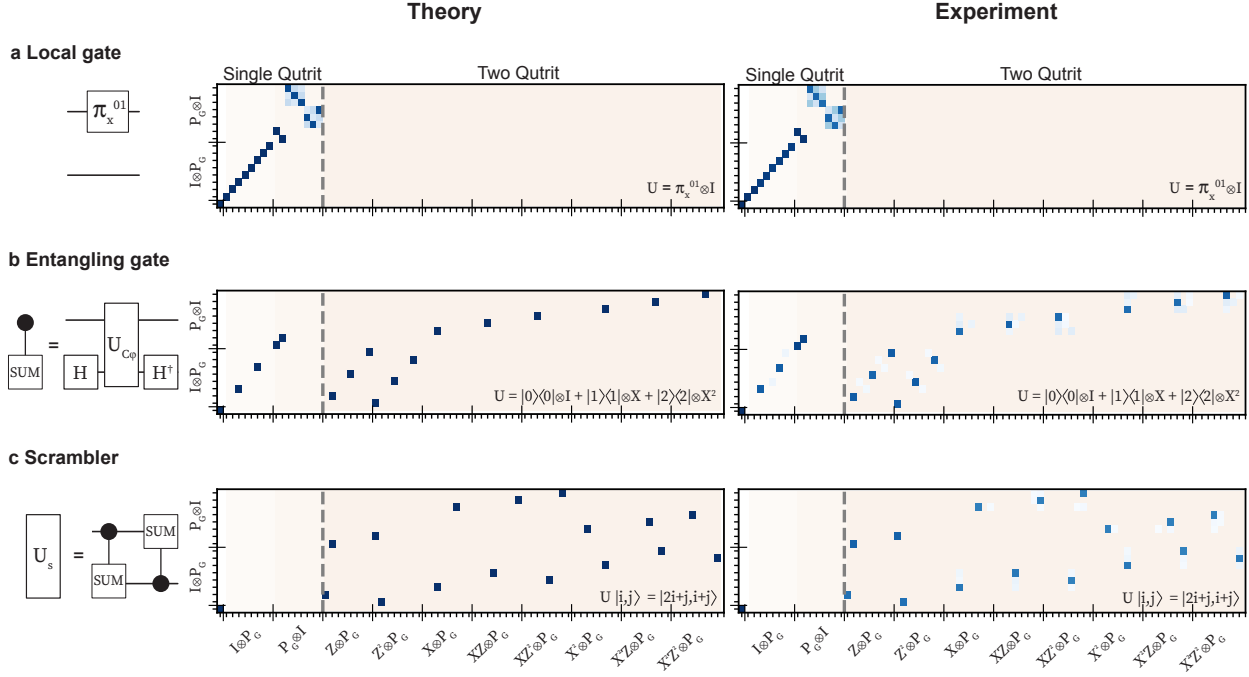
Figure 8.3:    Quantum process tomography of the two-qutrit scrambling unitary. Results
of process tomography both experimental and ideal. Plotted is part of the Pauli transfer
matrix (where the horizontal axis only includes the single qutrit operators) which shows
the unitary's action on single-qutrit Pauli operators when the unitary is a local gate (**a**),
non-scrambling entangling gate (**b**) and the scrambling-unitary (**c**), . This directly verifies
the key characteristic of scrambling by $U_s$, that it maps all non-identity single-qutrit Pauli
operators to two-qutrit operators.

i.e. a scrambling unitary which is a member of the Clifford group: this is evidenced by the
fact that the action of $U_s$ on Pauli operators is closed.

For a bit of context, we note that scrambling is not possible in a two-qubit system: there
is no unitary operation that completely delocalizes all single-qubit operators. A system of
two qutrits is the smallest bipartite system in which maximal scrambling is possible.

In the next two sections, we verify the maximally-scrambling nature of our implemen-
tation in two ways: (i) explicitly through quantum process tomography, and (ii) through a
five-qutrit teleportation protocol inspired by quantum gravitational physics.

## Verifying scrambling through quantum process tomography

On our processor, the scrambling unitary $U_s$ is constructed by applying two controlled-
SUM gates in sequence, switching the control and target qubit between the two (Fig. 8.3c).

In order to fully characterize our implementation of the scrambling unitary $U_\mathrm{s}$ we perform full quantum process tomography. For a two-qutrit unitary, this is a highly measurement-intensive procedure, which entails reconstructing, via two-qutrit state tomography, a $9 \times 9$ output density matrix for a complete set of 81 input states. This requires millions of measurements – each involving state preparation, application of $U_\mathrm{s}$, and state measurement (in a different basis) – with the precise number of repetitions determined by the desired statistical uncertainty. For experimental platforms with a duty cycle in the Hz range (e.g. trapped ions) this procedure would be prohibitively long; in contrast, superconducting circuits can reach repetition rates up to $\sim 100$ kHz, such that the full process matrix can be measured in approximately 1 hour. This gives superconducting platform the unique ability to not only quantify the process fidelity of the implemented scrambling unitary but also to measure precisely *how* it scrambles.

We implement the scrambling unitary using two sequential controlled-SUM gates based on the cross-Kerr gate described in the previous section. In Fig. 8.3c, we depict the results of quantum process tomography on our implementation. Through this tomography, we find that the fidelity of the scrambling operation on our hardware is 0.875, with two dominant error mechanisms: (i) dephasing and (ii) amplitude-damping during the cross-Kerr evolution.

Quantum process tomography also allows us to directly visualize the entire action of the unitary via its effect on local operators; i.e., we can directly visualize (Fig. 8.3) the relations (13)-(16). In particular, in Fig. 8.3c we can verify explicitly that $U_s$ transforms all single-qutrit operators into fully two-qutrit operators – the definition of a maximally scrambling unitary. We can also verify that it does so as a Clifford unitary, i.e. all Pauli operators are transformed into different Pauli operators. For comparison, we also illustrate the quantum process maps, both theoretical and experimental, of a single-qutrit unitary that does not delocalize any information (Fig. 8.3a), as well as a controlled-SUM gate, which is entangling but not fully scrambling (Fig. 8.3b). In the former, we verify that no single-qutrit operators are delocalized, while in the latter we observe that all single-qutrit operators are delocalized except for $Z^{(\dagger)} \otimes I$ and $I \otimes X^{(\dagger)}$, which commute with the controlled-SUM gate.

## Verifying scrambling through quantum teleportation

While process tomography provides an elegant and exhaustive way to "image" scrambling performed by our two-qutrit unitary, such an approach is infeasible for verifying scrambling in larger systems, as the number of required measurements scales exponentially with system size. As an example of a protocol which can circumvent this difficulty, and thus could be used on larger quantum processors, we turn to a teleportation protocol that quantifies the scrambling behavior of a unitary by using scrambling to enable the teleportation of quantum states [274, 517, 518]. Compared to tomography, a downside of the teleportation protocol is that it is limited as an essentially 'one-parameter' diagnostic of scrambling: it quantifies the average amount that operators on a given qutrit are transformed to have support on an additional qutrit. It does so via an average of OTOCs [163, 213, 277, 284, 289, 315, 398, 436, 486] associated with the unitary over local operators; in this context, maximal scrambling
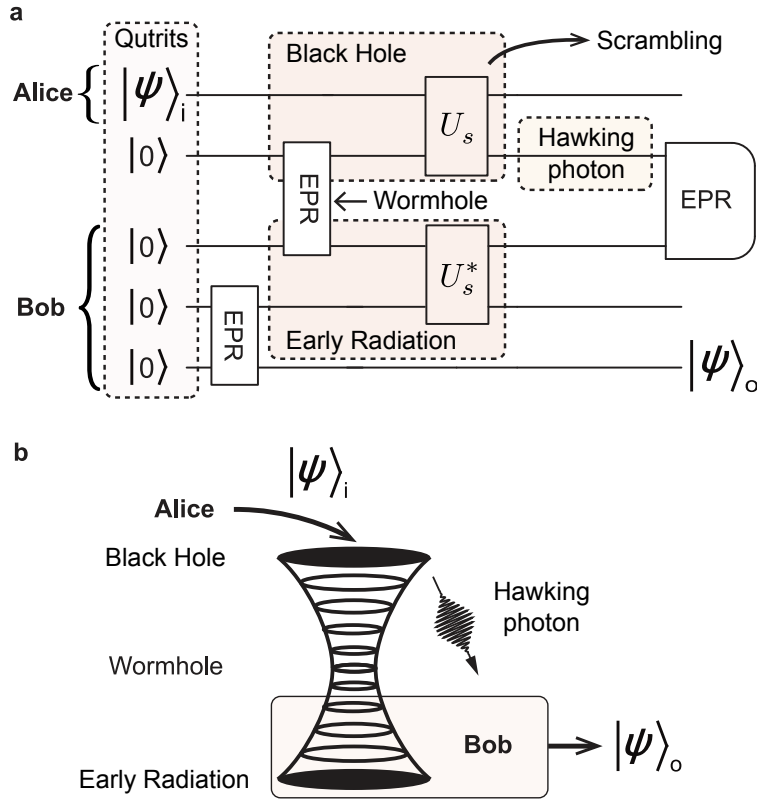
Figure 8.4:   Quantum teleportation circuit. Five-qutrit teleportation protocol used to test for scrambling by the two-qutrit unitary $U_s$, and its interpretation in the context of black hole physics. **(a)** We begin with the first qutrit in a quantum state $|\psi\rangle$, and prepare the remaining qutrits into two EPR pairs. The first qutrit $Q_1$ is 'scrambled' with half of the first EPR pair by the unitary $U$, while the conjugate $U^*$ is applied to $Q_3$ and $Q_4$. An EPR measurement on a pair of qutrits ($Q_2$ and $Q_3$) exiting each of $U$, $U^*$ serves to teleport the initial quantum state to $Q_5$ if and only if $U$ is scrambling. **(b)** In the context of the Hayden-Preskill thought experiment, $Q_1$ corresponds to Alice's diary, which subsequently falls into a black hole ($Q_2$) that is maximally entangled with its past Hawking radiation ($Q_3$), held by an outside observer Bob. In the gravitational picture, this shared entanglement functions similarly to a geometric wormhole between the black hole and its early radiation. The black hole dynamics are modelled as a scrambling unitary $U$, followed by emission of a Hawking photon ($Q_2$). Finally, Bob applies $U^*$ and measures in the EPR basis to recover Alice's diary.

by $U_s$ is captured by the fact that the average OTOC decays to its minimum allowed value (1/9 for a two-qutrit system) [518]. Crucially, the protocol is constructed in such a way that faithful teleportation of a quantum state $|\psi\rangle$ *requires* quantum information scrambling to occur, as well as the absence of experimental errors. The teleportation fidelity can then in turn be used to upper-bound the OTOC, even in the face of such errors.

As shown in Fig. 8.2, the verification protocol requires both the scrambling unitary, $U_s$, and its time-reversal, $U_s^*$, to be performed in parallel on separate pairs of qutrits. The qutrit pairs undergoing these two time-evolutions are initially highly correlated: two qutrits out of the four (one from each pair) begin as a maximally entangled EPR pair. After applying $U_s, U_s^*$, the probability of measuring this same pair in the EPR state decreases to 1/9 – the same value as in a maximally mixed state – due to scrambling with the additional qutrits. However, simply observing this decrease is *not* enough to robustly demonstrate scrambling, since the same decrease could also have resulted from deterioration of the EPR pair under experimental decoherence.

Here, teleportation comes to our aid: we place one of the remaining two qutrits ($Q_1$, the 'input') in an arbitrary pure single-qutrit state $|\psi\rangle$, and the other ($Q_4$) in an EPR pair with a fifth qutrit ($Q_5$, the 'output'). In the absence of experimental error, maximally scrambling dynamics guarantee that whenever $Q_2$ and $Q_3$ happen to be measured in their initial EPR state, the state $|\psi\rangle$ of $Q_1$ is teleported to $Q_5$. Unlike measuring a low EPR probability, high-fidelity teleportation can only arise from quantum information scrambling, not as a result of decoherence, making the teleportation fidelity a robust diagnostic of information scrambling [517, 518]. Furthermore, while experimental error will lead to a decay of the teleportation fidelity from unity, *any* measured fidelity above the classical limit (0.5 for qutrits) places a non-trivial upper-bound on the averaged OTOCs, and thus the scrambling behavior, of $U_s$ [518].

This association between scrambling and teleportation originated in black hole physics, and the teleportation protocol explored here is in fact inherited directly from this context. Its most straightforward interpretation is based on the Hayden-Preskill variant of the black hole information paradox [202, 517], as outlined in Fig. 8.2. Here, one pictures an observer Alice who has dropped a 'diary' consisting of the quantum state $|\psi\rangle$ into a black hole. Meanwhile, an outside observer Bob wonders whether this information is recoverable, as opposed to being irreversibly destroyed by the black hole. In seminal works [202, 517], it was shown that *if* (*i*) the black hole's dynamics are approximated as a fully scrambling quantum unitary $U_s$ and (*ii*) Bob possesses a large resource of entanglement with the black hole (e.g. a collection of its early Hawking radiation), then the state $|\psi\rangle$ can in fact be recovered via quantum operations on any few degrees of freedom emitted from the black hole (e.g. one additional Hawking photon).

In the gravitational picture, the shared entanglement between the black hole and its early radiation functions similarly to a geometric wormhole between the two systems. Indeed, in a close variant of this protocol, teleportation has a precise interpretation as Alice's state traversing a two-sided wormhole, in the specific case where the unitary corresponds to time-evolution under many-body dynamics with a holographic gravity dual [71, 160, 317]. Quite
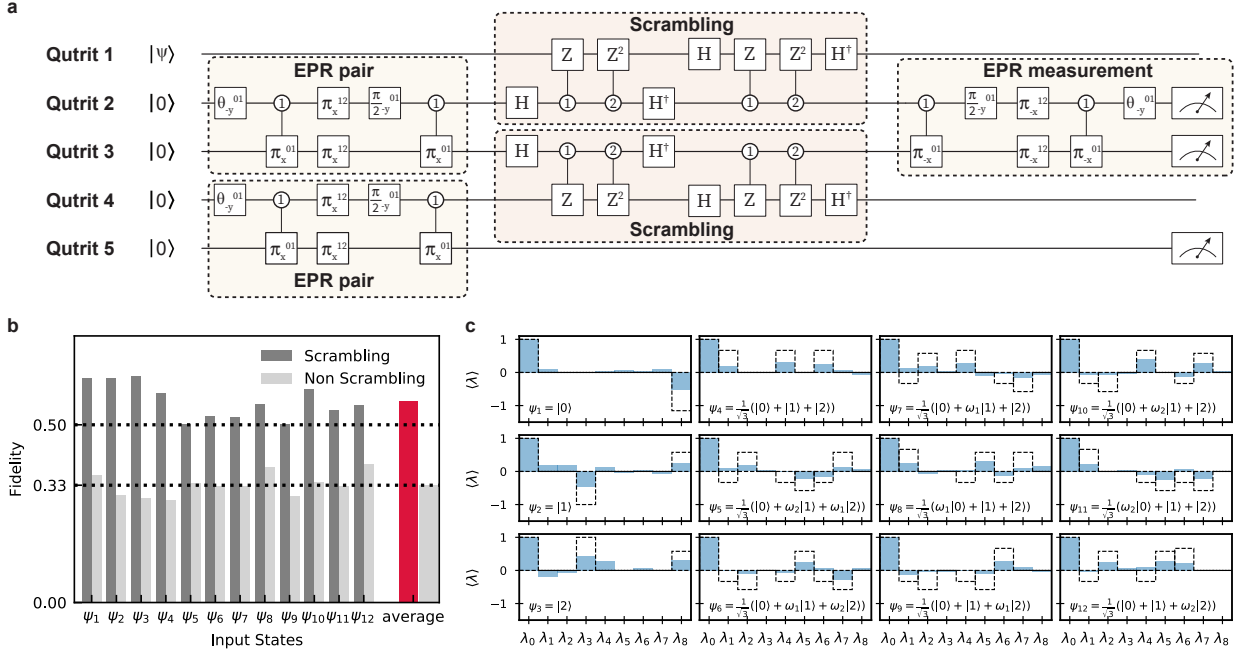
Figure 8.5:   Results of the five-qutrit teleportation protocol. **(a)** An expanded view of the five-qutrit teleportation protocol in Fig. 8.2, showing the native operations used to realize each portion of the protocol.  **(b)** Measured teleportation fidelities for twelve teleported states, which combine to give an unbiased estimate of the teleportation fidelity averaged over *all* single-qutrit pure states.  An average fidelity above $1/2$ – the classical limit for teleportation of qutrits – verifies non-zero quantum information scrambling by the maximally scrambling unitary, despite the presence of experimental error.  When the scrambling unitary is replaced with an identity operation with same complexity as the scrambler, the average teleportation state fidelity reduces to $1/3$, the same as a random guess when one does not have access to the input state.  **(c)** Representation of each of the twelve reconstructed density matrices after teleportation, expressed in the basis of Gell-Mann matrices $(\lambda_0 - \lambda_8)$ with dotted lines showing the ideal result.

surprisingly, recent work has revealed that the *same* protocol also features a more generic, non-gravitational mechanism for teleportation, which displays remarkably similar features to gravitational teleportation but is based only on the spreading of operators [429, 517]. This second form of teleportation encapsulates the physics observed here.

Experimentally, the teleportation protocol requires three essential ingredients:  initial preparation of EPR pairs, implementation of $U_s$ and $U_s^*$, and measurement in the EPR basis. We have discussed our implementation of the first two of these ingredients in previous sections. As our architecture only natively allows for computational-basis measurements, we

realize the EPR measurement on $Q_2$ and $Q_3$ by performing the EPR creation sequence in reverse, which transforms $|\text{EPR}\rangle$ to the computational basis state $|00\rangle$. We then measure all five qutrits in the computational basis and use the measurements outcomes showing $Q_2$ and $Q_3$ in $|00\rangle$ as a herald for successful teleportation. We verify success by tomographically reconstructing the quantum state of $Q_5$ after EPR measurement (by inserting state tomography pulses on $Q_5$ before the final measurement), represented by the density matrix $\rho_{\text{out}}$.

All three of the steps described above work particularly well on pairs of qutrits in "isolation", i.e. when the pair's neighboring qutrits are left in their ground states. However, this is necessarily *not* the setting of any five-qutrit quantum circuit, including the teleportation protocol. Here, we have to contend with the "always-on" nature of the cross-Kerr interaction, which leads to unwanted entangling dynamics between all neighboring qutrit pairs, even as we attempt to perform our quantum gates. To combat these unwanted interactions, we develop a novel set of dynamical decoupling sequences optimized for qutrits (see Appendix C).

We perform teleportation of twelve different single-qutrit quantum states $|\psi\rangle$, corresponding to all possible eigenstates of single-qutrit Pauli operators, i.e. three eigenstates each for $X, Z, XZ, XZ^2$ (the remaining Pauli operators are either inverses of these or the identity). We calculate the teleportation fidelity from the reconstructed density matrix of the output qubit as $F_\psi = \langle\psi|\,\rho_{\text{out}}\,|\psi\rangle$. Our chosen set of states is known as a 'state 2-design', which guarantees that the average teleportation fidelity $F_{\text{avg}} = (1/12)\sum_\psi F_\psi$ measured for these twelve sampled states is in fact equal to the average teleportation fidelity over *all* possible states [399, 518]. Furthermore, this average teleportation fidelity in fact allows us to upper bound the average OTOCs associated with the implemented unitary by $(4F-1)^{-2}$, without making *any* assumptions about the nature of the noise affecting our quantum processor [518].

The results of teleportation for two different choices of unitary dynamics are shown in Fig. 8.5. First, as a control, we perform the protocol with the identity operation in place of $U_s$. To ensure that the two have similar magnitudes and modes of experimental error, we implement the identity with precisely the same number and type of gates as the maximally scrambling unitary; the two differ only in the magnitude of certain software-defined phase gates. Since the identity operator does not scramble quantum information, we observe trivial teleportation fidelities $F_\psi \approx 1/3$ for all input states (Fig. 8.5b). Indeed, using quantum state tomography we verify that the final state of qubit $Q_5$ is near maximally mixed regardless of input state.

Finally, we perform the teleportation protocol with the maximally scrambling unitary $U_s$, which in theory should allow for perfect teleportation of any input quantum state. In contrast to the identity operation, we observe that all but one of the input states are teleported with fidelity $F_\psi > 1/2$, with an average fidelity of $F_{\text{avg}} = 0.568 \pm 0.001$. This allows us to experimentally upper bound the averaged OTOC by $0.618 \pm 0.004$. This marks a significant decay from its non-scrambling value of unity and verifies that our platform is capable of coherently simulating multi-qutrit scrambling dynamics.

## 8.3   Discussion and outlook

As in the previous Chapter, our work serves as a robust proof-of-principle measurement
of quantum information scrambling dynamics, and opens the door to many larger-scale ex-
plorations of many-body scrambling in future works. In addition, our works suggested a
number of intriguing future directions specific to qutrit quantum processors. First, our plat-
form opens the door to exploring the potential advantages of ternary quantum logic, including
a more efficient decomposition of the Toffoli gate [169] as well as magic state distillation pro-
tocols that outperform conventional qubit-based strategies in terms of both error-threshold
and yield [88]. Second, the dynamical decoupling techniques introduced here naturally apply
to other spin-1 systems including solid-state defect centers and multi-level atomic, molecular
and trapped ion systems [99, 433]. Third, one can imagine further enlarging the qudit di-
mension by further leveraging the intrinsic anharmonicity of transmons [162, 342], enabling
the study of many-body phases and entanglement dynamics in higher-spin quantum sys-
tems [362]. Fourth, building upon recent excitement surrounding quantum supremacy pro-
tocols using pseudorandom quantum circuit sampling [30], it would be natural to investigate
analogous qudit-based protocols, where supremacy might be achieved using a substantially
smaller number of elements. Finally, we note that the quantum teleportation experiment
performed can be interpreted as an implementation of quantum error-correction. In this
language, the teleportation protocol is equivalent to a three-qutrit quantum error-correcting
code, which protects information from the erasure of any one of the three qutrits [107]. Ex-
ploring other forms of error-correction in qudit-based processors is a natural extension of our
work.

# Chapter 9

# Many-body quantum teleportation via operator spreading in the traversable wormhole protocol

In this Chapter, we present a new mechanism for quantum teleportation using many-body quantum information dynamics. This teleportation represents a moderate extension of the teleportation considered in Chs. 7 and 8. Nonetheless, we will demonstrate that it features much greater power both as a quantum communication tool—e.g. it can teleport more qubits, for more types of scrambling dynamics—and as an experimental protocol for probing quantum information dynamics.

As introduced in Ch. 2, quantum teleportation leverages entanglement to transmit quantum information between distant locations [45, 48, 357, 391, 393]. Typically, one thinks about teleportation in the context of a few, well-controlled degrees of freedom. For example, two distant observers might share a pair of maximally entangled qubits (i.e. an Einstein-Podolski-Rosen (EPR) pair [353]), enabling a measurement by one observer to teleport an unknown quantum state to the other.

Recently, a confluence of seminal results has unveiled several novel instances of teleportation in strongly-interacting, *many-body* systems [38, 61, 71, 159, 160, 274, 314, 317, 350, 517, 518]. The circuit in Chs. 7 and 8 represents one example. Similar to conventional quantum teleportation, these protocols utilize shared entanglement as well as measurement and classical communication. However, they differ from conventional quantum teleportation in a few key aspects. Most notably, prior to teleportation, the initial quantum state is *scrambled* by the application of a many-body unitary. At first glance, this coexistence of scrambling— broadly speaking, the increasing complexity of initially simple quantum information under many-body time dynamics [213, 315, 398, 431, 436]—and teleportation might seem counter-intuitive. Indeed, one often thinks of teleportation as a directed quantum channel moving information between two specific locations; in contrast, scrambling disperses quantum information across all of the degrees of freedom in a system. As discussed in Ch. 7, the most natural way to reconcile these two perspectives is through the language of quantum error
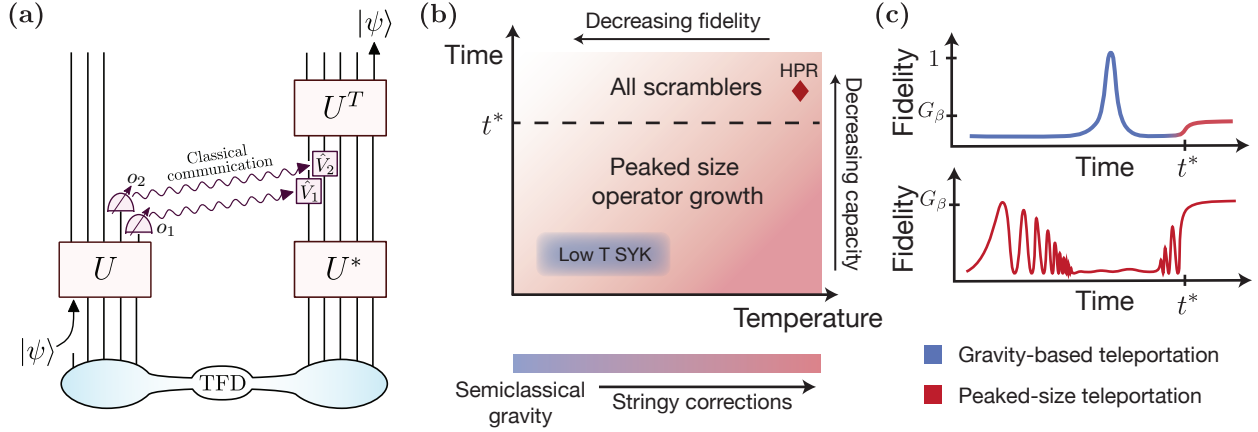
Figure 9.1: **(a)** Teleportation protocol, proceeding from bottom to top. To teleport, a subset of the left qubits are measured in the $\hat{O}_i$ basis, and operations $\hat{V}_i = e^{ig o_i \hat{O}_i / K}$ conditioned on the measurement results $o_i$ are performed on the right (purple). **(b)** The protocol hosts two mechanisms of teleportation: peaked-size (red) and gravitational (blue). The channel capacity of peaked-size teleportation decreases with increasing time (dark to light red), while its fidelity decreases with decreasing temperature (dark to light red, again). At high temperature and late times, it is equivalent to teleportation in the HPR protocol (red diamond). Gravitational teleportation occurs at low temperatures in systems dual to semi-classical gravity (e.g. the SYK model), and exhibits the same channel capacity but *higher* fidelity compared to peaked-size teleportation. Increasing the strength of stringy corrections to the gravity theory interpolates between gravitational and peaked-size teleportation. **(c)** The two mechanisms display distinct time profiles for the teleportation fidelity at fixed coupling strength, $g$. In systems dual to gravity (top), the fidelity features a single $\mathcal{O}(1)$ peak near the scrambling time (gravitational, blue), and a late time revival (peaked-size, red) to a fidelity suppressed by the two-point function $G_\beta$ [317]. In generic thermalizing systems (bottom), the fidelity oscillates between 0 and $G_\beta$ with phase proportional to the operator size, may subsequently decay if sizes become not peaked, and revives at late times.

correction [202]: by encoding, via scrambling, one observer's local information into non-local correlations across a many-body system, one can in fact teleport this information with access only to any few of the system's qubits.

The most notable example of many-body teleportation is the so-called traversable wormhole (TW) protocol, discovered in the context of quantum gravity [38, 71, 159, 160, 317, 350]. From the bulk gravitational perspective, this protocol consists of a particle traveling from one side of a wormhole geometry to the other; the wormhole is rendered traversable by the application of a coupling between the two sides. In the boundary theory, the wormhole geometry corresponds to a highly entangled thermofield double (TFD) state shared between two copies of a many-body system, and the coupling is implemented via measurement and

feed-forward operations [Fig. 12.1(a)]. Crucially, for this bulk-boundary correspondence to hold, the Hamiltonian describing the boundary system must exhibit "coherent", gravitational scrambling dynamics—this is realized, most notably, in the Sachdev-Ye-Kitaev (SYK) model at low temperatures [249, 316].

Interestingly, recent work has uncovered a number of instances of many-body teleportation *without* gravitational dynamics. For example, teleportation in the TW protocol was recently demonstrated analytically in the SYK model at *high* temperatures [159], and numerically in chaotic spin chains at late times [71, 350]; in both cases, the microscopic mechanism for teleportation remains an outstanding puzzle. In addition to the TW protocol, the many-body teleportation protocol in Chs. 7 and 8 was introduced in the context of the Hayden-Preskill variant of the black hole information paradox [202, 517]. This so-called Hayden-Preskill recovery (HPR) protocol allows for many-body teleportation via *generic* scrambling dynamics, including as we saw, Clifford scrambling unitaries. Although the two protocols bear some structural similarity, the HPR protocol is exponentially less efficient for teleporting multiple qubits. To this end, understanding the precise relationship between these protocols remains an essential open question.

In this Chapter, we present a unified framework for many-body teleportation from the perspective of the growth of operators under scrambling time-evolution. Most significantly, this framework leads to the identification of a new teleportation mechanism—dubbed *peaked-size teleportation*—which succeeds for a wide variety of physical systems and encapsulates all known examples of many-body teleportation outside of the gravitational regime. We emphasize that peaked-size teleportation represents a *distinct* teleportation mechanism compared to "gravitational" teleportation. Although the same TW protocol can host either mechanism, the features of peaked-size teleportation differ markedly from those of gravitational teleportation [Fig. 12.1(c), Table I]. Crucially, this distinction implies that the TW protocol can act as a litmus test for identifying intrinsically gravitational dynamics. More broadly, our results pave the way towards utilizing the TW protocol as a powerful experimental tool for characterizing the growth of operators in strongly interacting systems.

## 9.1 Summary of results

Due to the excessive length of this Chapter, we begin by providing a technical overview of our main results and the organization of the remaining Chapter. This summary is intended to introduce the overarching concepts of our work, such that the remaining sections are self-contained and can be read according to individual preference. A more detailed, section-by-section guide to the reader is included at the end of this summary.

In **Section 9.2**, we begin with a general description of the TW protocol [Fig. 1(a)]. In this protocol, locally encoded quantum information is inserted into one side of an entangled thermofield double (TFD) state and teleported to the other side through a combination of (*i*) unitary evolution of each side individually, and (*ii*) a simple two-sided coupling that acts on a large subsystem of each side. The coupling is quite flexible in form, and corresponds to

unitary evolution, $e^{igV}$, under a two-sided interaction

$$V = \frac{1}{K} \sum_{i=1}^{K} O_{i,l} O_{i,r}^*,$$
(9.1)

where $\{O_i\}$ are *any* set of $K$ local, non-identity operators applied to the left ($l$) and right ($r$) side of the system. This coupling can be performed as either a quantum gate, or through local measurements of $O_i$ on the left side, followed by classical communication and feed-forward operations on the right side [Fig. 1(a)].

In **Section 9.3**, we discuss the general requirements for successful teleportation in the TW circuit. In particular, we relate the teleportation fidelity to the following correlation functions of the two-sided system [317]:

$$C_Q(t) \equiv \langle \text{TFD}| \, Q_r(-t) e^{igV} Q_l(t) \, |\text{TFD}\rangle$$
(9.2)

where $Q(\pm t)$ is a time-evolved operator initially acting on the qubit(s) to be teleported. Our analysis leads to two conditions on these correlators that, when combined, are necessary and sufficient for teleportation to succeed with unit fidelity:

1. The magnitudes of the correlators must be maximal for every $Q$.

2. The phases of the correlators must be the same for every $Q$.

Here, $Q$ runs over a complete basis of operators on the qubits to be teleported. We show that Condition 1 is naturally satisfied, even without the coupling $V$, if the TFD state is at infinite temperature, in which case it reduces to an extensive set of maximally entangled EPR pairs. On the other hand, Condition 2 requires that the coupling acts non-trivially on the operators $Q$.

In **Section 9.4**, we describe the relation between the coupling, $V$, and the growth of time-evolved operators, $Q(t)$. For the purposes of teleportation, this growth is characterized by the *size distribution* of the operators [379, 381, 397], which provides a finer-grained measure of quantum information scrambling compared to more conventional quantities such as out-of-time-ordered correlators (OTOCs) [277, 315, 436]. Specifically, writing $Q(t)$ as a sum over Pauli strings, $Q(t) = \sum_R c_R(t) R$, we define the size distribution as:

$$P(\mathcal{S}) = \sum_{\mathcal{S}[R]=\mathcal{S}} |c_R(t)|^2,$$
(9.3)

where the sum is over Pauli strings, $R$, of size, $\mathcal{S}$ (equal to the string's number of non-identity components). By probing correlations between the two sides of the doubled Hilbert space, the coupling $V$ directly measures the operator size [379].

In **Section 9.5**, we introduce the *peaked-size* mechanism for many-body teleportation. This mechanism is possible whenever the size distributions of time-evolved operators, $Q(t)$, are tightly peaked about their average size. In this scenario, the exponentiated coupling,

$e^{igV}$, applies approximately the same phase, proportional to the size, to each coefficient, $c_R$, and therefore to the entire operator, $Q(t)$. We show that these applied phases are sufficient to align the correlators' phases for all operators $Q$, thereby achieving Condition 2. We also demonstrate that the magnitudes of the correlators are unchanged by the coupling when size distributions are tightly peaked. This implies that peaked-size teleportation achieves perfect fidelity at infinite temperature, where Condition 1 is automatically satisfied; at finite temperature, peaked-size teleportation can still occur, but with a reduced fidelity (Table I).

In **Sections 9.6-9.7**, we analyze examples of peaked-size teleportation across a wide variety of interacting, many-body dynamics. We demonstrate that the capabilities of peaked-size teleportation—most notably, the fidelity and the number of qubits that can be sent (i.e. the channel capacity)—depend on the temperature, coupling strength, evolution time, and the specific scrambling dynamics of the model under study (Table I).

More specifically, in **Section 9.6**, we provide general arguments that all scrambling systems exhibit peaked-size teleportation at *late times*, after the system's scrambling time $(t \gtrsim t_s)$. In this regime, operators have become fully delocalized across the system, so that their size distributions are peaked about a typical, extensive value. We also show that late time peaked-size teleportation is limited to transmitting only a single qubit.

In **Section 9.7**, we show that many scrambling quantum systems also feature peaked-size teleportation at *intermediate times*, i.e. after the local thermalization time but before the scrambling time $(t_{th} \lesssim t \lesssim t_s)$. We begin with ergodic short-range interacting systems in $\geq 1$D, which we show naturally possess peaked-size distributions due to thermalization within the bulk of a time-evolved operator's light cone. In contrast, the size distributions of operators in all-to-all coupled (0D) systems are not intrinsically peaked; nevertheless, peaked sizes can be engineered by non-locally encoding the quantum information before insertion into the teleportation circuit. Interestingly, in both of these classes of dynamics, we find that multiple $(\sim \mathcal{O}(K))$ qubits can be teleported simultaneously via the peaked-size mechanism, in contrast with the unit channel capacity of late time teleportation. We substantiate these claims through extensive numerical and analytic studies on a variety of physical models: random unitary circuits (RUCs) in dimensions $d = 0, 1$, and 2 [340], the SYK model, and (in Section 9.9) experimentally relevant spin chain Hamiltonians [52].

In **Section 9.8**, we discuss the interplay between peaked-size and gravitational teleportation. Notably, we expect gravitational teleportation to occur only at low temperatures, where certain quantum mechanical models (e.g. the SYK model) are known to possess a dual description in terms of conformal matter coupled to gravitational dynamics. From the perspective of operator growth, the unique feature of gravitational teleportation is the presence of non-trivial phase *winding* in a variant of the size distribution [71]. Crucially, this effect enables gravitational teleportation to satisfy Condition 1, and thereby achieve high teleportation fidelity at low temperatures, in sharp contrast with peaked-size teleportation (Table I).

Intriguingly, while it may seem that there is a sharp distinction between peaked-size and gravitational teleportation, we find that this is not always this case. In particular, we show that varying the temperature of the SYK model provides a continuous interpolation

| Model | Teleportation mechanism | Protocol parameters | Maximum per qubit fidelity |
|---|---|---|---|
| All scrambling systems at late times (Refs. [317, 517], Section 9.6) | peaked-size | $g = \pi \bmod 2\pi$ | $\sim G_\beta$ |
| $\geq$ 1D RUCs & chaotic spin systems (Sections 9.5,9.7,9.9) | peaked-size | $\eta_d g \mathcal{S}(t)/N = \pi \bmod 2\pi$ | $\sim G_\beta$ |
| 0D RUCs, with encoding (Section 9.7) | peaked-size | $\eta_d g \mathcal{S}(t)/N = \pi \bmod 2\pi$ | $\sim 1$ |
| High-$T$ SYK, with encoding (Ref. [159], Sections 9.7) | peaked-size | $\eta_d g \mathcal{S}(t)/N = \pi \bmod 2\pi$ | $\sim 1$ |
| Low-$T$ SYK / AdS$_2$ gravity (Refs. [71, 159, 160, 317], Fig. 12.1) | gravitational | $ge^t/N \sim 1$ (SYK) $\\ gG_N e^t \sim 1$ (AdS$_2$) | $\sim 1$ |
| AdS$_2$ gravity with strong stringy corrections, with encoding (Section 9.8) | peaked-size | $g\mathcal{S}(t)/N \sim \pi \bmod 2\pi$ | $\sim G_\beta$ |

Table 9.1: Summary of our expectations for teleportation in a variety of physical models. For each model, we specify the associated teleportation mechanism, the optimal value of the coupling strength $g$, the optimal teleportation fidelity, and the channel capacity. Here $G_\beta$ is the imaginary time two-point function (Section 9.5), $\mathcal{S}(t)$ is the size of a time-evolved operator, $K$ is the number of measured qubits [Fig. 12.1(a)], $\eta_d = 1/(1-1/d^2)$ is an order one constant determined by the local qudit dimension $d$ [Sec. 9.4], and $G_N$ is Newton's constant. We refer to the Summary of Results and the cited sections for further details.

between gravitational teleportation at low temperature and peaked-size teleportation at high temperature. In the dual picture, perturbing away from the low temperature limit corresponds to adding *stringy* corrections to the gravity theory [181, 316, 437]. Following this intuition, we show that teleportation in a gravity theory with strong stringy corrections [317] bears a remarkable qualitative similarity to peaked-size teleportation, thus providing a first step towards a bulk understanding of this phenomenon.

Finally, in **Section 9.9**, we discuss experimental applications of the TW protocol for probing many-body dynamics. In particular, we demonstrate that the protocol can function as a diagnostic tool for scrambling dynamics in near-term quantum simulators, enabling one to starkly distinguish between generic thermalizing systems and gravitational dynamics. To this end, we provide detailed blueprints for realizing the protocol in two complementary experimental platforms—Rydberg atom arrays [52, 174, 273, 311, 319, 496] and trapped

ions [35, 59, 91, 158, 333]. Specifically, the observation of a high teleportation fidelity at low temperatures would be a tantalizing experimental indicator of gravitational scrambling dynamics. In addition, gravitational dynamics exhibit unique qualitative features as a function of both evolution time and protocol parameters [Fig. 12.1(c), Table 9.1]. More broadly, our analysis suggests that the TW protocol can provide insights into many-body dynamics outside the gravitational regime. In particular, we demonstrate that the fidelity of peaked-size teleportation probes higher moments of operator size distributions [381].

*Guide to the reader*—Considering the wide scope of results presented in this work, we encourage readers to skip to sections that align with their specific interests and refer to the above summary for context. To this end, we highlight below the nature of each section and provide recommendations for readers of different backgrounds. Sections 9.2-9.4 introduce the formal tools and derivations necessary for rigorously understanding our results. These sections will be of interest to readers with a background in quantum information who wish to understand the precise connection between teleportation and operator sizes. Sections 9.5-9.7 introduce peaked-size teleportation and analyze its realization in several example systems. Since many these systems are experimentally accessible, these sections will be most relevant to members of the quantum simulation and many-body physics communities. Section 9.8 focuses on the interplay of peaked-size teleportation and gravitational physics, both in the SYK model and from a bulk gravitational perspective. For brevity, background material on gravitational physics is relegated to references, making this section best suited for experts at the intersection of quantum information and quantum gravity. Finally, Section 9.9 contains a summary of the experimental signatures of the TW protocol, detailed blueprints for Rydberg atom and trapped ion implementations, and a discussion of the protocol's behavior under experimental error. This section will be of interest to AMO experimentalists and all readers interested in near-term realizations of many-body quantum teleportation [375].

## Relation to previous works

To further elaborate on the broad context of our results, a brief summary of the relevant prior studies and their relation to our work is provided as follows.

*Gravitational teleportation in the TW protocol*—Traversable wormhole teleportation was originally introduced in Refs. [160, 317] in the context of gravitational physics, where it was realized that a coupling of the form $V$ enables a traversable channel between the boundaries of a two-sided black hole. The explicit quantum mechanical circuit implementing this teleportation [Fig. 12.1(a)] was later introduced in Refs. [71, 159], alongside exact calculations for the teleportation fidelity in the large-$q$ SYK model [159]. While the emphasis of our work is not on the bulk interpretation of gravitational teleportation—indeed, the peaked-size teleportation mechanism is intended to *contrast* with the gravitational mechanism—it will be helpful to recall the main results from the gravitational perspective.

We focus on the specific case of two-dimensional anti-de Sitter space, which is the bulk dual of the SYK model at low temperatures [252, 317]. In the simplest case (ignoring gravitational backreaction), the two-sided correlator, Eq. 9.2, can be explicitly calculated

and is given by [317]:

$$C_Q(t) = \left( \frac{1}{2 - g^{\frac{\Delta_O}{2^{2\Delta_O+1}}} G_N e^{2\pi t/\beta}} \right)^{2\Delta_Q}. \tag{9.4}$$

Here, $G_N$ is Newton's constant, $\beta = 1/T$ is the inverse temperature of the black hole, $\Delta_O$ is the conformal dimension of the coupling operators $O_i$ [Eq. (9.1)], and $\Delta_Q$ is the conformal dimension of the operator $Q$. In the context of the SYK model, $G_N$ is inversely proportional to the number of Majorana fermions, $N$, and the black hole temperature is equal to the temperature of the TFD state [159, 317].

For our purposes, the most notable feature of the correlator is that it exhibits a sharp peak at time $t \approx G_N \log(g)$ [Fig. 12.1(c)], corresponding to the moment a particle inserted on one side of the black hole emerges on the other side. While in the above formula [Eq. (9.4)], the correlator diverges at this time, in the large-$q$ SYK model, this divergence is regularized and the correlator peaks at its maximal value of unity [159]. Thus, at time $t \approx G_N \log(g)$, the correlator satisfies Condition 1 for successful teleportation; in Ref. [159], it was shown that Condition 2 is also satisfied for certain conformal dimensions of the operators $Q$. In combination, this leads to unit teleportation fidelity.

Another notable feature of gravitational teleportation is the ability to teleport multiple qubits simultaneously, as discussed in Ref. [317]. In the gravitational picture, multi-qubit teleportation has an intuitive explanation: particles corresponding to different qubits pass through the black hole in parallel, without interacting with one another. However, for sufficiently many qubits, the effects of gravitational backreaction become important, leading to a predicted channel capacity of $\mathcal{O}(K)$.

*HPR teleportation*—An independent, but closely related, set of protocols for many-body teleportation was introduced in Ref. [517] for the recovery of information in the Hayden-Preskill thought experiment [202]. Unlike previous works on traversable wormholes, in Ref. [517] teleportation succeeds for any fully scrambling unitary dynamics (i.e. at late times, $t \gtrsim t_s$), with no reliance on gravitational physics. However, the channel capacity of HPR teleportation is fundamentally limited: multi-qubit teleportation requires a protocol whose circuit depth grows exponentially in the number of qubits to be teleported [517].

In Appendix D.2, we show that a deterministic variant of the HPR protocol (for single-qubit teleportation) is in fact *equal* to the TW protocol in Fig. 12.1(a), restricted to infinite temperature and with a particular choice of the coupling operators, $O_i$. Furthermore, in Section 9.6 we show that teleportation at *late times* via the peaked-size mechanism is equivalent to this variant of HPR teleportation. However, peaked-size teleportation is more powerful than HPR teleportation in the sense that: (*i*) it succeeds for a much larger class of couplings, $V$, (*ii*) it can succeed at intermediate times, and (*iii*) at such times, it is capable of sending multiple qubits with no change in the protocol's complexity, an exponential improvement over the HPR protocol.

*Previous many-body teleportation experiments*—Many-body quantum teleportation has recently been demonstrated in both trapped ion [274] and superconducting qutrit [61] experiments. Both Refs. [61, 274] implement a probabilistic variant of the HPR protocol, which

differs slightly from the TW protocol, while Ref. [274] also implements the deterministic
variant discussed above. In all cases, the scrambling dynamics, $U$, are generated by digital
quantum gates acting on a small number of qubits. Teleportation is performed for a single
qubit and a fully scrambling unitary, placing the experiments in the same physical regime
as late-time, peaked-size teleportation.

Our work demonstrates that experiments in the *TW protocol* at *intermediate times* can
access new regimes of many-body quantum teleportation, with the potential to provide more
information about the scrambling dynamics under study. Most notably, such experiments
can distinguish between teleportation in generic many-body systems (via the peaked-size
mechanism) versus systems with a gravity dual (via the gravitational mechanism), which is
not possible in the HPR protocol.

*SYK teleportation in the TW protocol*—In Ref. [159], the two-sided correlator of the
TW protocol [Eq. (9.2)] was calculated exactly for the large-$q$ SYK model (defined in Sec-
tion 9.7). As anticipated in Ref. [317], the correlator at low temperatures—where the model
is dual to gravity—agrees with the gravitational result [Eq. (9.4)] up to the previously men-
tioned regularization. More surprisingly, it was shown that teleportation with unit fidelity
is also possible at high temperatures—where the model is not dual to gravity. As we will
see in Section 9.7, all features of high temperature teleportation in the SYK model are in
precise agreement with the peaked-size mechanism; our work thus provides a microscopic
understanding for this previously unexplained result.

*Gravity in the lab*—Ref. [71] discusses various instances of teleportation in the TW pro-
tocol. The authors distinguish two teleportation mechanisms: (*i*) an "operator transfer"
mechanism, which occurs at intermediate times in gravitational systems and is capable of
teleporting multiple qubits, and (*ii*) a "state transfer" mechanism, which occurs at late times
in all scrambling systems, and is capable of sending only a single qubit. Moreover, they intro-
duce a microscopic interpretation for the teleportation mechanism in gravitational systems,
termed "size winding", which we connect to in Section 9.8.

In our terminology, the first teleportation mechanism corresponds to gravitational tele-
portation, while the second mechanism corresponds to peaked-size teleportation at late
times[1]. In our work, we provide a microscopic interpretation for late time teleportation
(i.e. the peaked-size mechanism) and demonstrate that it is equivalent to teleportation in
the HPR protocol. In addition, we demonstrate that peaked-size teleportation is a more
general phenomenon that also occurs at intermediate times in many systems, where we show
that it is capable of teleporting multiple qubits.

In a follow-up work, Ref. [350], whose pre-print was posted concurrently with that of
this work, the same authors elaborate on their previous results and provide more detailed
examples and calculations. These agree with our own results in areas of overlap.

---

[1]The terminology of Ref. [71] can be understood using our two Conditions for teleportation. Specifically,
operator transfer corresponds to situations that satisfy Condition 1, but not necessarily Condition 2, as
occurs in gravitational teleportation [see Eq. (9.4)]. State transfer corresponds to situations that satisfy
both Conditions.

## 9.2   Introduction to diagrammatic notation

We begin by introducing a diagrammatic "tensor network" notation for depicting the teleportation circuit. Adapted from Ref. [517], this notation provides a precise visual framework for analyzing teleportation in Section 9.3 and will be convenient for deriving rigorous results on the teleportation fidelity in Section 9.5.

To begin, we represent a quantum ket $|\psi\rangle$ and bra $\langle\psi|$ as:

$$|\psi\rangle = \Big|_{|\psi\rangle} \quad , \quad \langle\psi| = \Big|^{\langle\psi|} \tag{9.5}$$

Note that time proceeds upwards—an initial state $|\psi\rangle$ terminates the bottom of a leg, while a final projection $\langle\psi|$ terminates the top. Similarly, much as in Fig. 12.1(a), we represent an operator, for instance the many-body unitary $U$, as a box with input (bottom) and output (top) legs:

$$U = \boxed{\begin{smallmatrix} \text{C} & U & \text{C} \\ \text{A} & & \bar{\text{A}} \end{smallmatrix}} \tag{9.6}$$

Here we have decomposed the input and output into two subsystems, A and its complement for the input, C and its complement for the output, in reference to the teleportation protocol. Specifically, comparing to Fig. 12.1(a), subsystem A consists of the qubits supporting the input state $|\psi\rangle$, while subsystem C consists of the coupled qubits.

The diagrammatic notation is particularly useful when working with EPR states. The EPR state on two qubits is defined as $|\text{EPR}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$; for a system of $N$ $d$-dimensional qudits, this is generalized to $\frac{1}{\sqrt{d^N}}\sum_{i=1}^{d^N}|i\rangle_l|i\rangle_r^*$. Here $\{i\}$ is an arbitrary $d^N$-dimensional basis, $*$ denotes time-reversal (i.e. complex conjugation), and $l$ and $r$ denote the left and right system, respectively. In the diagrammatic notation, we represent this as:

$$|\text{EPR}\rangle = \frac{1}{\sqrt{d^N}}\sum_{i=1}^{d^N}|i\rangle_l|i\rangle_r^* \equiv \quad\overset{\text{A}}{\Big|}\quad\overset{\bar{\text{A}}}{\underset{\bullet}{\Big\lfloor}}\quad\overset{\bar{\text{A}}}{\underset{}{\Big\rfloor}}\quad\overset{\text{A}}{\Big|} \tag{9.7}$$

We have again decomposed each system into two subsystems, A and its complement, $\bar{\text{A}}$, for convenience (subsystem A is chosen to be identical between the left and right sides). Each dot represents a normalization factor given by the inverse square root of the subsystem's dimension.

To see the utility of the diagrammatic notation, recall that a fundamental property of the EPR state is that an operator acting on the left side is equivalent to its transpose acting

on the right:

$$O_l \left|\text{EPR}\right\rangle = \frac{1}{\sqrt{d^N}} \sum_{i,j} O_{ij} \left|i\right\rangle_l \left|j^*\right\rangle_r$$

$$= \frac{1}{\sqrt{d^N}} \sum_{i,j} O_{ij}^T \left|j\right\rangle_l \left|i^*\right\rangle_r = O_r^T \left|\text{EPR}\right\rangle \tag{9.8}$$

where the middle equality swaps the $i$, $j$ indices of the sum. In diagrammatic notation, this becomes simply



$$\tag{9.9}$$

i.e. the operator $O$ "slides" from the left to right side of the EPR pairs, with its input and output indices correspondingly transposed. Similarly, expectation values in the EPR state can be easily computed in terms of the trace of (one-sided) operators, e.g.



$$\tag{9.10}$$

where the final equality follows from $\left\langle\text{EPR}\right| B_l A_l^T \left|\text{EPR}\right\rangle = (1/d^N) \sum_{ij} \left\langle i^*|j^*\right\rangle \left\langle i\right| B A^T \left|j\right\rangle = (1/d^N) \sum_i \left\langle i\right| B A^T \left|i\right\rangle$.

The EPR state is closely related to the thermofield double (TFD) state,

$$\left|\text{TFD}\right\rangle \equiv \sum_i e^{-\beta E_i/2} \left|E_i\right\rangle_l \left|E_i^*\right\rangle_r / \text{tr}(e^{-\beta H})^{1/2}. \tag{9.11}$$

Here $H$ is a time-reversal symmetric Hamiltonian, $H = H^*$, with eigenstates $\left|E_i\right\rangle$, and eigenvalues $E_i$. The TFD state is parameterized by an effective "temperature" $1/\beta$. At infinite effective temperature ($\beta = 0$), the TFD and EPR states are equal. At finite temperature, the TFD state is obtained by applying the square root of the density matrix,

$\rho^{1/2} \equiv e^{-\beta H/2}/\mathrm{tr}(e^{-\beta H})^{1/2}$, to either side of the EPR state, which we represent as:

$$|\mathrm{TFD}\rangle = \sqrt{d^N \rho_l} \, |\mathrm{EPR}\rangle = \sqrt{d^N \rho_r^*} \, |\mathrm{EPR}\rangle$$



$$(9.12)$$

For the finite temperature TFD state, the analog of Eq. (9.9) holds *only* for operators that commute with the Hamiltonian. Most notably, such operators include the time-evolution operator, $U = e^{-iHt}$, which thus obeys:



$$(9.13)$$

Eq. (9.13) also holds for backwards time-evolution, replacing $U \to U^\dagger, U^T \to U^*$. We note that for time-reversal symmetric Hamiltonians, $U = U^T$. In this case, combining Eqs. (9.13, 9.9), we have the useful identity:

$$O_l(t) \, |\mathrm{TFD}\rangle = O_r^T(-t) \, |\mathrm{TFD}\rangle . \tag{9.14}$$

Applying Eq. (9.10), we can again express 'two-sided' expectation values in the TFD state in terms of 'one-sided' correlation functions, e.g.

$$\langle \mathrm{TFD}| \, A_l(t) B_r(t') \, |\mathrm{TFD}\rangle = \mathrm{tr}(\rho^{1/2} A^T(-t) \rho^{1/2} B(t')). \tag{9.15}$$

Let us now re-draw the full teleportation protocol in Fig. 12.1(a) using the diagrammatic notation:



$$(9.16)$$

This circuit proceeds as follows: ($i$) prepare the TFD state, ($ii$) insert the state $|\psi\rangle$ on subsystem A of the left side, ($iii$) time-evolve the two sides by $U$, $U^*$, ($iv$) couple the two sides via the unitary operator $e^{igV}$, with $V$ as in Eq. (9.1), ($v$) evolve the right side by $U^T$, ($vi$) apply a 'decoding' operator $D$, and ($vii$) measure the output state of subsystem A on the right side. Compared to Fig. 12.1(a), we have made two modifications. First, we have replaced the measurement and classical communication with a quantum coupling $e^{igV}$, as described in Section 9.1. Second, we now include a simple decoding operator, $D$, applied at the end of the circuit before state recovery. We will find that $D = Y \otimes \ldots \otimes Y$ for peaked-size teleportation of a multi-qubit subsystem, where $Y$ is the single-qubit Pauli $Y$ operator (Section 9.5).

Finally, we note that a straightforward application of Eq. (9.13) allows us to re-express the circuit as



$$(9.17)$$
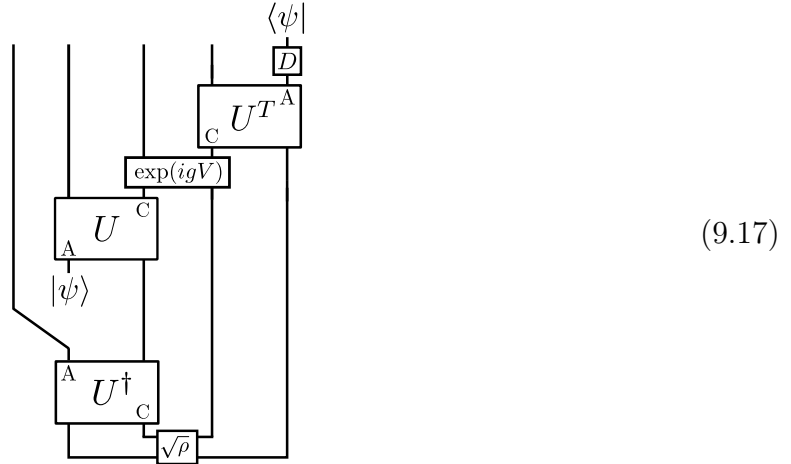
This equivalent version of the protocol was introduced in Refs. [71, 159] and will be more convenient for analysis from here on.

## 9.3 General requirements for successful teleportation

We now introduce heuristic arguments for when teleportation succeeds in this protocol. This will culminate in the two requirements for teleportation listed in Section 9.1. In Section 9.5, we derive these conditions more formally by providing exact relations between the two-sided correlators in Eq. (9.2) and the teleportation fidelity.

We begin with the protocol in Eq. (9.17). To proceed, we insert a resolution of the

identity $\mathbb{1} = \sum_\phi |\phi\rangle\langle\phi|$ on the "swapped out" subsystem A (the output of $U_l^\dagger)^2$:

$$\text{(9.18)}$$

This reformulation makes it clear that teleportation depends on the action of the coupling on states of the form $Q_{A,l}(t)|\text{TFD}\rangle$, where $Q_A = |\psi\rangle\langle\phi|$ and[3] $Q_A(t) \equiv UQ_AU^\dagger$.

Teleportation succeeds when the coupling "transfers" $|\psi(t)\rangle\langle\phi(t)|$ from the left to right side of the TFD state. More precisely, the following identity, if true for *all* operators $Q_A$ on A, would guarantee successful teleportation for all states:

$$= e^{i\theta_Q} \qquad \text{(9.19)}$$

Here $\theta_Q$ is an overall phase and we represent conjugation by the decoding operator as $\tilde{Q}_A \equiv D^\dagger Q_A D$. One can verify this explicitly by plugging the RHS of the above equality into Eq. (9.18): the topmost applications of $DU^T$ and $U^*D^\dagger$ cancel, leaving $Q_A \to |\psi\rangle\langle\phi|$ as the topmost operator on the right side, i.e. subsystem A is in the state $|\psi\rangle$.

To quantify whether this equality holds, we measure the inner product between the two

---

[2]At infinite temperature, using Eq. (9.9), $|\phi\rangle$ can be understood as the counterpart of $|\psi\rangle$, to be teleported from right to left instead of left to right. To see this, use Eqs. (9.9, 9.13) to re-expess $\langle\phi|_l U_l^\dagger \to U_l^\dagger |\phi\rangle_r$, and apply $D_l U_l^\dagger$ after coupling to recover $|\phi\rangle$.

[3]Traditionally, this would be considered reverse time-evolution, and denoted $Q_A(-t)$. For brevity, we have flipped the sign of $t$ throughout the text.

states[4]:

$$C_Q = \langle \text{TFD}| \, \tilde{Q}_{A,r}^\dagger(-t) e^{igV} Q_{A,l}(t) \, |\text{TFD}\rangle$$



(9.20)

This is precisely the two-sided correlation function introduced in Eq. (9.2), now modified to include the decoding operator. In particular, if the correlation function is maximal for all operators $Q_A$, then Eq. (9.19) holds and teleportation succeeds with perfect fidelity for all initial states.

In practice, it is sufficient to evaluate the correlators for a complete basis of operators on subsystem A (e.g. the Pauli operators). In this case, we now have *two* requirements on the operator correlators, as listed in Section 9.1: ($i$) all correlators must have maximal magnitude, i.e. equal to 1, and ($ii$) all correlators must have the *same phase*—if two operators both individually obey Eq. (9.19) but with different phases, their sum will not.

At infinite temperature, owing to Eq. (9.9), we will see that the first requirement is satisfied even in the absence of the coupling, for any symmetric or antisymmetric operator. To satisfy the second requirement, the role of the coupling $e^{igV}$ must be to apply a $Q_A$-dependent overall phase. In the following section, we analyze the action of the coupling and show precisely when such an overall phase occurs.

## 9.4 Connection to operator size

In this section, we outline the connection between the coupling $V$ and the operator size when $V$ is acted on states of the form:

$$Q_{A,l}(t) \, |\text{TFD}\rangle = Q_{A,l}(t) \rho_l^{1/2} \, |\text{EPR}\rangle . \tag{9.21}$$

This connection was discovered in a number of previous works, focusing primarily on a specific bilinear coupling in fermionic systems [71, 294, 295, 350, 379, 397, 445]. In the

---

[4]For simplicity of notation and consistency with previous works [159, 160, 317], from here on we have assumed that the unitary is symmetric, $U^T = U$, $U^\dagger = U^*$.

following, we introduce this connection in the context of bosonic systems and argue that it applies to a good approximation for any generic, local couplings. From this, we then show that the action of the exponentiated coupling, $e^{igV}$, is particularly simple—it applies an overall phase—whenever operator size distributions are tightly peaked.

## Coupling measures size

In bosonic qudit systems, we define the size of a Pauli string as its number of non-identity elements [397]. For instance, the Pauli string

$$\mathbb{1} \otimes X \otimes \mathbb{1} \otimes \mathbb{1} \otimes Z \otimes X \otimes \mathbb{1} \tag{9.22}$$

has size 3. A more general operator can be written as a sum of Pauli strings, $R$:

$$Q_A(t)\rho^{1/2} = \sum_R c_R(t)R, \tag{9.23}$$

and possesses a corresponding *size distribution* [379, 397][5]:

$$P(\mathcal{S}) = \sum_{\mathcal{S}[R]=\mathcal{S}} |c_R(t)|^2. \tag{9.24}$$

The distribution is normalized to 1 if $Q_A$ is unitary,

$$\sum_{\mathcal{S}} P(\mathcal{S}) = \sum_R |c_R(t)|^2 = \mathrm{tr}(Q_A^\dagger Q_A \rho) = 1. \tag{9.25}$$

One can naturally characterize the size distribution via its moments—for instance, the average size, $\mathcal{S}[Q_A(t)\rho^{1/2}] \equiv \sum_{\mathcal{S}} P(\mathcal{S})\mathcal{S}$ (when context is clear, we denote this simply as $\mathcal{S}$), and the size width, $\delta\mathcal{S}$.

We will now show that the coupling $V$ approximately measures the operator size, in the sense that it acts on states of the form Eq. (9.21) as:

$$VQ_{A,l}(t)\,|\mathrm{TFD}\rangle \approx d^{N/2} \sum_R \left(1 - \eta_d \frac{\mathcal{S}[R]}{N}\right) c_R(t) R_l\,|\mathrm{EPR}\rangle, \tag{9.26}$$

where $\eta_d \equiv 1/(1 - 1/d^2)$ is an order one constant determined by the local qudit dimension, $d$. Expectation values of $V$ thus measure the average size, while higher powers of $V$ measure higher moments of the size distribution [379, 397]. In particular, the exponentiated coupling in the teleportation protocol applies a *size-dependent phase* to each Pauli string of $Q_A(t)\rho^{1/2}$:

$$e^{igV} Q_{A,l}(t)\,|\mathrm{TFD}\rangle \approx d^{N/2}$$
$$e^{ig} \sum_R e^{-i\eta_d g \mathcal{S}[R]/N} c_R(t) R_l\,|\mathrm{EPR}\rangle, \tag{9.27}$$

---

[5]We note that, at finite temperature, the coefficients $c_R(t)$ will generally be complex. Their phases thus carry information beyond that captured by the size distribution, which we discuss in Section 9.8.

We derive this connection by first introducing an *exact* measure of operator size in bosonic qudit systems, generalizing previous measures for Majorana fermionic systems [379, 397]. We then argue that successively more generic couplings display approximately the same behavior, when acted on time-evolved operators in generic many-body scrambling dynamics.

In bosonic qudit systems, we find that the operator size is precisely measured by a sum of individual EPR projectors on each qudit $i$:

$$V_s = \frac{1}{N} \sum_{i=1}^{N} P_{\text{EPR},i} = \frac{1}{Nd^2} \sum_{i=1}^{N} \sum_{P_i} P_{i,l} P_{i,r}^*, \tag{9.28}$$

where $d$ is the local qudit dimension, $N$ is the number of qudits, and $P_i$ form a complete basis of single-qudit operators (e.g. for qubits $P_i \in \{\mathbb{1}, X, Y, Z\}$). To see this, let us first analyze the action of a single EPR projector, $P_{\text{EPR},i}$. Writing a given Pauli string as a tensor product of single-qudit Paulis, $R = \bigotimes_{j=1}^{N} R_j$, we find

$$P_{\text{EPR},i} R_l \left| \text{EPR} \right\rangle = \delta_{R_i,\mathbb{1}} R_l \left| \text{EPR} \right\rangle, \tag{9.29}$$

using Eq. (9.10) and $\text{tr}_i(R_i)/d_i = \delta_{R_i,\mathbb{1}}$. A single EPR projector thus acts as a binary variable, giving eigenvalue 1 or 0 if a given Pauli string is, or is not, the identity on the designated qudit. The full coupling is a sum of these binary variables over all qudits and therefore counts the total number of non-identity elements in the Pauli string, i.e. the operator size. Its eigenvectors are the states $R_l \left| \text{EPR} \right\rangle$ with eigenvalues $1 - \mathcal{S}[R]/N$, as in Eq. (9.26).

We now turn to more general local couplings. First, as a trivial but useful modification, we can remove the identity operators from $V_s$, since these are not included our original definition of the coupling, $V$ [Eq. (9.1)]. These constitute a fraction $1/d^2$ of the complete basis, $P_i$, summed in Eq. (9.28). Removing these terms renormalizes the eigenvalues of the coupling:

$$\left( \frac{1}{N(d^2-1)} \sum_{i=1}^{N} \sum_{P_i \neq \mathbb{1}} P_{i,l} P_{i,r}^* \right) R_l \left| \text{EPR} \right\rangle$$
$$= \left[ 1 - \eta_d \frac{\mathcal{S}[R]}{N} \right] R_l \left| \text{EPR} \right\rangle, \tag{9.30}$$

which now match those quoted in Eq. (9.26). Note that the left side sum is now over $N(d^2-1)$ non-identity operators and normalized accordingly.

Second, we consider omitting some of the *non-identity* $P_i$ at each site. Intuitively, under thermalizing dynamics, if an operator has spread to some qudit $i$ it should not matter which Pauli operator we use to probe the operator's presence. For example, for qubits, we could omit the $O_j = X_i, Y_i$ couplings and keep only $O_j = Z_i$. A random Pauli string has equal probability to commute with $Z_i$ as it would with $X_i$ and $Y_i$; thus, coupling using only $Z_i$ operators is sufficient for measuring a thermalized operator's size.

Third, we expect even more general couplings, composed of $O_i$ that are local but not necessarily Pauli operators, to behave similarly. Specifically, each individual coupling, $O_{i,l}O_{i,r}$, will asymptote to two different expectation values before and after the time-evolved operator has spread to the support of $O_i$. Before, the coupling will maintain its expectation value in the unperturbed TFD state, $\mathrm{tr}(O_i\rho^{1/2}O_i^\dagger\rho^{1/2})$. After, the spread of $Q_A(t)$ will disrupt the two-sided correlations in the TFD state that give rise to this initial expectation value, and the coupling will instead asymptote to its value in two thermal states, $\mathrm{tr}(O_i\rho)\cdot\mathrm{tr}(O_i\rho)$. As before, the sum of many terms, each behaving as above, leads to an approximate measure of operator size.

Lastly, we consider the case where the coupling is restricted to act only on some subsystem C, consisting of $K$ qudits[6]. The coupling now measures the number of non-identity elements of a Pauli string within C—we denote this as the $K$-*size*, $\mathcal{S}_K$, of the Pauli string. The eigenvalues of the coupling are the same as those in Eq. (9.30), with the replacement $\mathcal{S}/N \to \mathcal{S}_K/K$. For a typical Pauli operator, we expect the $K$-size distribution of an operator to be similar to its full size distribution when $K$ is large and the coupled qubits are distributed randomly. In particular, in this scenario we expect the average $K$-size, $\mathcal{S}_K$, to be proportional to the average size, $\mathcal{S}$,

$$\frac{\mathcal{S}_K}{K} \approx \frac{\mathcal{S}}{N}. \tag{9.31}$$

For simplicity, we will make this substitution in the remainder of the work. However, if $C$ is a spatially local subsystem (instead of a random subsystem), then this replacement will be modified depending on the spatial extent of the operator.

As a final remark, we note that the operator size distribution is directly related to out-of-time-order correlators (OTOCs), a more familiar quantity for probing operator growth [277, 315, 436]. In particular, the average size is equal to a sum of OTOCs between $Q_A$ and $O_i$ [379, 397],

$$\mathrm{OTOC}_1 = \langle V \rangle_Q \equiv \langle \mathrm{TFD}|\, Q_{A,l}^\dagger(t)\, V\, Q_{A,l}(t)\, |\mathrm{TFD}\rangle$$



$$= \frac{1}{K}\sum_{i=1}^{K} \tag{9.32}$$

$$= \frac{1}{K}\sum_{i=1}^{K} \mathrm{tr}\left(\rho^{1/2}\, Q_A(t)\, O_i\, Q_A^\dagger(t)\, \rho^{1/2}\, O_i^\dagger\right)$$

using Eqs. (9.9-9.15). Higher moments of the size distribution can also be probed by OTOCs, now between $Q_A$ and various products of the $O_i$, e.g. $O_iO_j$ for the size width. We discuss

---

[6]For simplicity, this assumes that there is a single coupling per qudit in C.

these relations further, paying particular attention to subtleties that arise at finite temperature, in Section 9.8.

## Peaked-size distributions

The exponentiated coupling [Eq. (9.27)] has a particularly action when the size distribution of $Q_A(t)\rho^{1/2}$ is *tightly peaked* about its average size. In this regime, each Pauli string gains approximately the same phase, and so the action of the coupling reduces to applying a $Q_A$-dependent overall phase,

$$e^{igV}Q_{A,l}(t)\left|\text{TFD}\right\rangle \approx e^{ig\langle V\rangle_Q}Q_{A,l}(t)\left|\text{TFD}\right\rangle, \tag{9.33}$$

where the applied phase is proportional to the average $K$-size [see Eq. (9.30, 9.31)],

$$\begin{aligned}
g\left\langle V\right\rangle_Q &= g\left\langle\text{TFD}\right|Q_{A,l}^\dagger(t)VQ_{A,l}(t)\left|\text{TFD}\right\rangle \\
&\approx g - \eta_d g\frac{\mathcal{S}_K[Q_A(t)\rho^{1/2}]}{K},
\end{aligned} \tag{9.34}$$

defining $\eta_d \equiv 1/(1 - 1/d^2)$ for convenience.

Corrections to this behavior are controlled by higher moments of the size distribution. Focusing on the overlap of the coupled and uncoupled states, the leading order correction is equal to the $K$-size variance, $\delta S_K^2/K^2 = \langle V^2\rangle_Q - \langle V\rangle_Q^2$, multiplied by $g^2$:

$$\begin{aligned}
\left\langle e^{igV}\right\rangle_Q &= \left\langle 1 + igV - \frac{1}{2}g^2V^2 + \dots\right\rangle_Q \\
&= \left(1 + ig\left\langle V\right\rangle_Q - \frac{1}{2}g^2\left\langle V\right\rangle_Q^2 + \dots\right) \\
&\quad - \frac{1}{2}g^2\left(\left\langle V^2\right\rangle_Q - \left\langle V\right\rangle_Q^2\right) + \dots \\
&= \exp\left(ig\left\langle V\right\rangle_Q\right) - \frac{1}{2}(\eta_d g)^2\delta\mathcal{S}_K^2/K^2 + \dots
\end{aligned} \tag{9.35}$$

The $K$-size variance receives contributions from two sources: the variance of the full size distribution, $\delta\mathcal{S}^2$, and a statistical error from sampling only $K$ of $N$ qubits for the $K$-size. If the $K$ qubits are distributed randomly, these errors scale as $\delta\mathcal{S}_K \sim \delta\mathcal{S} \cdot (K/N)$ and $\delta\mathcal{S}_K \sim \sqrt{\mathcal{S}_K} \approx \sqrt{\mathcal{S}K/N}$, respectively (see Appendix D.6 for a detailed derivation of the latter). These are small compared to the average $K$-size whenever $\delta\mathcal{S} \ll \mathcal{S}$ and $1 \ll \mathcal{S}_K$.

In Appendix D.1, we go beyond these leading order corrections and provide quantitative bounds on when the peaked-size approximation in Eq. (9.33) is valid. In general, we can strictly prove that this approximation holds whenever there is a parametric separation between an asymptotic size width, defined in the appendix, and the average size.

## 9.5 Peaked-size teleportation

Having established general conditions for successful teleportation (Section 9.3) as well as the connection between the coupling in the TW protocol and operator size distributions (Section 9.4), we are now ready to introduce the peaked-size mechanism for teleportation. In this section, we first demonstrate peaked-size teleportation in its simplest context: teleportation of a single qubit at infinite temperature. We then show that the fidelity of peaked-size teleportation is necessarily suppressed at finite temperature. For ease of reading, we relegate rigorous results supporting each of the above arguments to the end of the section. We turn to specific physical systems realizing peaked-size teleportation in the following sections: in Section 9.6 we show that peaked-size teleportation of a single qubit occurs in all scrambling systems at late times, while in Section 9.7 we show that peaked-size teleportation of multiple qubits occurs in certain systems at intermediate times.

### Single-qubit teleportation

To analyze teleportation of a single qubit, we turn to the two-sided correlators in Eq. (9.20), with $Q_A \in \{\mathbb{1}, X, Y, Z\}$ running over the single-qubit Pauli operators. We recall that the requirements for teleportation are for all $C_Q$ to have ($i$) maximal magnitude and ($ii$) the same phase.

The first requirement is naturally satisfied at infinite temperature even *before coupling and decoding* but the second requirement is not. In particular, the four correlators with $D = \mathbb{1}$, $g = 0$ are:

| $Q_A$ | $C_Q$ |
|:---:|:---:|
| $\mathbb{1}$ | $+1$ |
| $X$ | $+1$ |
| $Y$ | $-1$ |
| $Z$ | $+1$ |

$$(D = \mathbb{1})$$
$$(g = 0)$$

where the left entries are qubit operators, $Q_A$, and the right entries are the correlators, $C_Q$. The correlators have maximal magnitude because each operator can be transferred perfectly from left to right using Eq. (9.9). However, the $Y$ operator picks up an overall minus sign during this process, since $Y^T = -Y$, and so the correlator phases are not aligned. One can verify the resulting teleportation fidelity is indeed trivial. Our goal will be to show that the action of the coupling in Eq. (9.33), as well as a simple decoding operation, are sufficient to align the four phases.

To begin, we assume that all time-evolved Pauli operators have a tightly peaked size distribution and that the average size $\mathcal{S}$ is the same for all non-identity operators. From Eqs. (9.33-9.34), we have that the coupling applies a total phase difference $\eta_d g \mathcal{S}/N$ between

the thermofield double state (the identity operator; size zero) and all perturbed states (time-evolved Pauli operators; size $\mathcal{S}$). Our table of correlator phases is thus modified to:

| $Q_A$ | $C_Q$ |
|-------|-------|
| $\mathbb{1}$ | $+1$ |
| $X$ | $+1$ |
| $Y$ | $-1$ |
| $Z$ | $+1$ |

$(D = \mathbb{1})$
$(g = 0)$

$\longrightarrow$

| $Q_A$ | $C_Q$ |
|-------|-------|
| $\mathbb{1}$ | $e^{-i\eta_d g \mathcal{S}/N}$ |
| $X$ | $+1$ |
| $Y$ | $-1$ |
| $Z$ | $+1$ |

$(D = \mathbb{1})$
$(g \neq 0)$

We again do not achieve perfect phase alignment. However, we can now correct the misaligned phases using the decoding operator, $D = Y$. This applies an additional minus sign to the $X$ and $Z$ correlators:

| $Q_A$ | $C_Q$ |
|-------|-------|
| $\mathbb{1}$ | $+1$ |
| $X$ | $+1$ |
| $Y$ | $-1$ |
| $Z$ | $+1$ |

$(D = \mathbb{1})$
$(g = 0)$

$\longrightarrow$

| $Q_A$ | $C_Q$ |
|-------|-------|
| $\mathbb{1}$ | $e^{-i\eta_d g \mathcal{S}/N}$ |
| $X$ | $+1$ |
| $Y$ | $-1$ |
| $Z$ | $+1$ |

$(D = \mathbb{1})$
$(g \neq 0)$

$\longrightarrow$

| $Q_A$ | $C_Q$ |
|-------|-------|
| $\mathbb{1}$ | $e^{-i\eta_d g \mathcal{S}/N}$ |
| $X$ | $-1$ |
| $Y$ | $-1$ |
| $Z$ | $-1$ |

$(D = Y)$
$(g \neq 0)$

The correlator phases are now aligned whenever

$$\eta_d g \frac{\mathcal{S}}{N} = \pi \bmod 2\pi, \tag{9.36}$$

leading to perfect teleportation at these values.

## Peaked-size teleportation at finite temperature

There are two important modifications to peaked-size teleportation at finite temperature. First, the relevant notion of operator size is modified [379]. In particular, in the peaked-size regime, the difference in phase applied between the identity and non-identity Pauli operators is modified to

$$\mathcal{S}[Q_A(t)] \to \mathcal{S}[Q_A(t)\rho^{1/2}] - \mathcal{S}[\rho^{1/2}]. \tag{9.37}$$

Second, the maximal fidelity of peaked-size teleportation is reduced at finite temperature. In particular, when sizes are tightly peaked, the two-sided correlators factorize into a constant magnitude multiplied by an overall phase:

$$\begin{aligned} C_Q &= \langle \mathrm{TFD} | \, \tilde{Q}_{A,r}^\dagger Q_{A,l} \, | \mathrm{TFD} \rangle \, e^{i(g - \eta_d g \mathcal{S}_K [Q_A(t)\rho^{1/2}]/K)} \\ &= G_\beta(Q_A) \cdot e^{i\theta_Q} \end{aligned} \tag{9.38}$$

where $\theta_Q$ combines the effects of transposition, coupling, and decoding, and the correlator magnitude corresponds to an imaginary-time Green's function,

$$G_\beta(Q_A) \equiv \mathrm{tr}(Q_A^\dagger \, \rho^{1/2} Q_A \, \rho^{1/2}) \leq 1. \tag{9.39}$$

This Green's function is unity at infinite temperature and generically decreases at finite temperatures, due to the reduced entanglement of the TFD state. This violates the maximal magnitude requirement for teleportation, and therefore leads to a corresponding decrease in the teleportation fidelity.

The astute reader will recall that finite temperature teleportation is known to succeed with $\mathcal{O}(1)$ fidelities (i.e. higher than $G_\beta$) in theories with a gravity dual [159, 160, 317]; this is a signature of physics outside the peaked-size regime, which we connect to in Section 9.8.

## Rigorous expressions for teleportation fidelity

We now derive formal expressions of the teleportation fidelity for $n$ teleported qubits as a function of the correlator phases. To do so, we consider a variant of the protocol where instead of teleporting a quantum state we attempt to distill an EPR pair:



$$\tag{9.40}$$
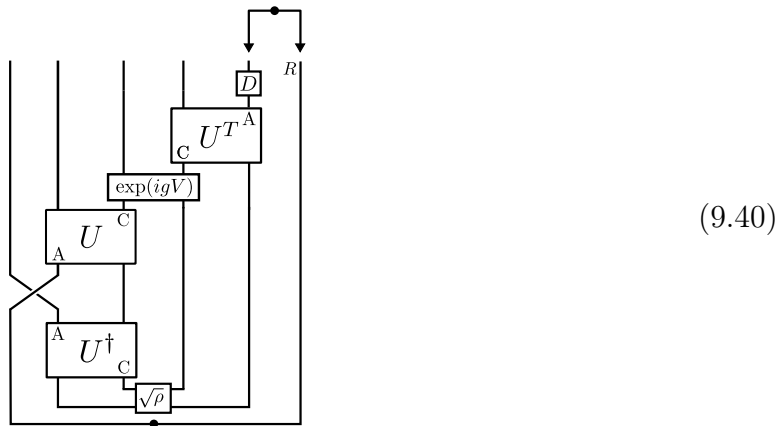
Here state insertion is replaced by swapping in one "half" of an EPR pair with a reference subsystem R (far right) into subsystem A of the left side. When subsystem A is teleported from left to right, the circuit results in an EPR pair between the reference subsystem R and subsystem A of the right (top arrows). The fidelity of EPR distillation is precisely related to the average fidelity of state teleportation [518], $F_{\mathrm{EPR}} = [(d_A + 1)\langle F_\psi \rangle - 1]/d_A$, where $d_A = 2^n$ is the dimension of subsystem A when teleporting $n$ qubits.

We calculate the teleportation fidelity by Pauli decomposing the SWAP operator as

$\text{SWAP} = \sum_{Q_A} Q_A \otimes Q_A^\dagger / d_A$. This gives:



$$\tag{9.41}$$

where the third equality utilizes the diagrammatic identities Eqs. (9.9, 9.10), and the fourth equality inserts the identity, $\mathbb{1} = D_r U_r U_r^\dagger D_r^\dagger$, in the center of the right side (recall our notation $\tilde{Q}_{1/2} = D^\dagger Q_{1/2} D$). Writing the rightmost diagram as an equation, we have:

$$
\begin{aligned}
F_{\text{EPR}} = \frac{1}{d_A^4} \sum_{Q_1, Q_2} \langle \text{TFD} | \, Q_{2,l}^\dagger(t) \, e^{-igV} \, \tilde{Q}_{2,r}(-t) \\
\times \, \tilde{Q}_{1,r}^\dagger(-t) \, e^{igV} \, Q_{1,l}(t) \, | \text{TFD} \rangle .
\end{aligned}
\tag{9.42}
$$

Similar expressions for teleportation of quantum states are contained in Appendix D.3.

In general, the teleportation fidelity and two-sided correlators are related only by a lower bound,[7]

$$
F_{\text{EPR}} \geq \left| \frac{1}{d_A^2} \sum_{Q_A} C_Q \right|^2 .
\tag{9.43}
$$

This is obtained diagrammatically by inserting the projector, $|\text{TFD}\rangle\langle\text{TFD}|$, into the center

---

[7]Under special circumstances, namely large-$N$ models, one may be able to factorize the above expression in terms of correlators of the form Eq. (9.20) [159].

of Eq. (9.41):



$$\frac{1}{d_A^4} \sum_{Q_1,Q_2} \geq \frac{1}{d_A^4} \sum_{Q_1,Q_2} \tag{9.44}$$

A similar bound was obtained in Ref. [71, 350], conditional on certain assumptions about operators' size distributions.

At infinite temperature *in the peaked-size regime*, we have $C_Q = e^{i\theta_Q}$ and the fidelity is equal to the lower bound:

$$F_{\mathrm{EPR}} = \frac{1}{d_A^4} \sum_{Q_1,Q_2} e^{i(\theta_{Q_1}-\theta_{Q_2})} = \left| \frac{1}{d_A^2} \sum_{Q_A} e^{i\theta_Q} \right|^2. \tag{9.45}$$

The sum is over $d_A^2$ terms, and is unity only when all the operators' phases are the same. In the case of a single-qubit teleportation at infinite temperature in the peaked-size regime, plugging the final table of Section 9.5 into the above equation gives a fidelity:

$$F_{\mathrm{EPR}} = \frac{5}{8} - \frac{3}{8} \cos(\eta_d g \mathcal{S}/N), \tag{9.46}$$

which oscillates between trivial fidelity ($F_{\mathrm{EPR}} = 1/4$) and unity as a function of the operators'

size. At finite temperature in the peaked-size regime, we instead find

$$
\begin{aligned}
F_{\mathrm{EPR}} &= \frac{1}{d_A^4} \sum_{Q_1, Q_2} e^{i(\theta_{Q_1} - \theta_{Q_2})} \mathrm{tr}(Q_2^\dagger Q_1 \, \rho^{1/2} Q_1^\dagger Q_2 \, \rho^{1/2}) \\
&\leq \frac{1}{d_A^4} \sum_{Q_1, Q_2} \mathrm{tr}(Q_2^\dagger Q_1 \, \rho^{1/2} Q_1^\dagger Q_2 \, \rho^{1/2}) \\
&= \frac{1}{d_A^2} \sum_{Q_A} \mathrm{tr}(Q_A \, \rho^{1/2} Q_A^\dagger \, \rho^{1/2}) \\
&= \frac{1}{d_A^2} \sum_{Q_A} G_\beta(Q_A).
\end{aligned}
\tag{9.47}
$$

where the maximum fidelity is again achieved when the correlator phases align. However, its value is now less than unity, and instead is equal to a sum of various imaginary time Green's functions, i.e. the correlator magnitudes [Section 9.5, Eq. (9.39)].

## 9.6 Peaked-size teleportation at late times

We now introduce the simplest physical example of peaked-size teleportation: teleportation in any scrambling system at late times (after the scrambling time). There are two distinguishing features of this regime: (*i*) the circuit can only teleport a single qubit, i.e. the channel capacity is one, and (*ii*) as for all peaked-size teleportation, the teleportation fidelity is suppressed at low temperatures. We also demonstrate that this regime of peaked-size teleportation, as well as the full quantum circuit implementing the TW protocol, are equivalent to HPR teleportation of a single qubit. In Section 9.7, we will demonstrate that the single-qubit late time channel capacity can be overcome at intermediate times in many scrambling systems.

### Teleportation at late times

At late times, the dynamics of a scrambling system can be approximated by a Haar random unitary[8] [202, 399]. In this case, each time-evolved operator, $Q_A(t)$, becomes a sum of random Pauli strings, each with probability $1/d^2$ to be the identity at any individual site. As a result, time-evolved operators have an average size,

$$
\mathcal{S} \approx (1 - 1/d^2)N,
\tag{9.48}
$$

---

[8]This approximation is modified in systems with a conserved quantity. Size distributions in such systems have been considered in Refs. [246, 384, 385]; at late times (after conserved quantities have diffused across the entire system), they are expected to be similar to size distributions without a conserved quantity, up to corrections $\sim 1/N$.

and a size width,

$$\delta\mathcal{S} \sim \sqrt{N}, \tag{9.49}$$

where the scaling is based on the central limit theorem. The $K$-size distribution takes the same form, replacing $N$ with $K$, and is tightly peaked as long as $K$ is large (specifically, $g\delta\mathcal{S}_K/K \approx g/\sqrt{K} \ll 1$).

For simplicity, we will focus on late time teleportation at infinite temperature; finite temperature modifications follow according to Section 9.5. Using Eqs. (9.33-9.34), we find that the coupling applies a relative phase $e^{ig}$ between the identity operator (size zero) and all non-identity Pauli operators (size above) [317]:

$$e^{igV}\left|\text{EPR}\right\rangle = e^{ig}\left|\text{EPR}\right\rangle$$
$$e^{igV}Q_{A,l}(t)\left|\text{EPR}\right\rangle = Q_{A,l}(t)\left|\text{EPR}\right\rangle. \tag{9.50}$$

The lack of an applied phase for non-identity Pauli operators corresponds to the vanishing of $\langle V\rangle_Q$ at late times, when OTOCs have decayed to zero [see Eq. (9.34)]. From Section 9.5, we see that whenever

$$g = \pi \bmod 2\pi, \tag{9.51}$$

single-qubit teleportation succeeds.

A brief argument shows that late time teleportion of higher dimensional quantum states is not possible. Consider teleportation of a $d$-dimensional qudit, with a basis of states $|i\rangle$, $i = 0, \ldots, d-1$. The qudit Pauli operators are generated by the 'clock' and 'shift' operators: $Z\left|i\right\rangle = e^{i\omega}\left|i\right\rangle$, with $\omega = 2\pi/d$, and $X\left|i\right\rangle = \left|i+1\right\rangle$. The two generators obey the commutation relation, $XZ = e^{-i\omega}ZX$. After transposition, each Pauli operator, $X^pZ^q$, becomes

$$(X^pZ^q)^T = Z^{T,q}X^{T,p} = Z^qX^{-p} = e^{-ipq\omega}X^{-p}Z^q. \tag{9.52}$$

Meanwhile, late time dynamics ensure that the coupling applies an overall phase only to the identity operator. For teleportation to be successful, we would therefore require a decoding operation, $D$, that acts as $DX^{-p}Z^qD^\dagger \sim X^pZ^q$. Suppose there was such a unitary operator[9], and consider its action on the generators: $DXD^\dagger = X^{-1}$ and $DZD^\dagger = Z$. The above action implies that commuting the two generators gives a different phase before and after decoding: $DXZD^\dagger = e^{-i\omega}DZXD^\dagger = e^{-i\omega}ZX^{-1}$ and $DXZD^\dagger = X^{-1}Z = e^{+i\omega}ZX^{-1}$. This is a contradiction whenever $e^{+i\omega} \neq e^{-i\omega}$, i.e. whenever $d > 2$.

## Equivalence to HPR protocol

We now turn to the equivalence between peaked-size teleportation and teleportation in the HPR protocol. The latter was originally introduced to recover information in the Hayden-Preskill thought experiment [202, 517], and is reviewed in detail in Appendix D.2.

---

[9]The astute reader may note that this operation is in fact implemented by the *anti-unitary* operator, $D\left|i\right\rangle = \left|-i \bmod d\right\rangle^*$. However, if one decomposes state insertion in terms of Pauli operators as $|\psi\rangle\langle\phi| = \sum_{Q_A} c_Q Q_A$ (see Section 9.3), one desires that the entire operator $|\psi\rangle\langle\phi|$ be transferred from left to right for all possible $\langle\phi|$. The preceding anti-unitary operator will complex conjugate the coefficients $c_Q$, thus spoiling teleportation for any $\langle\phi|$ where these are complex.

Here, we restrict our attention to teleportation in the deterministic variant of the protocol, of a single qubit at infinite temperature [274, 517]. The protocol takes the form:



$$\tag{9.53}$$
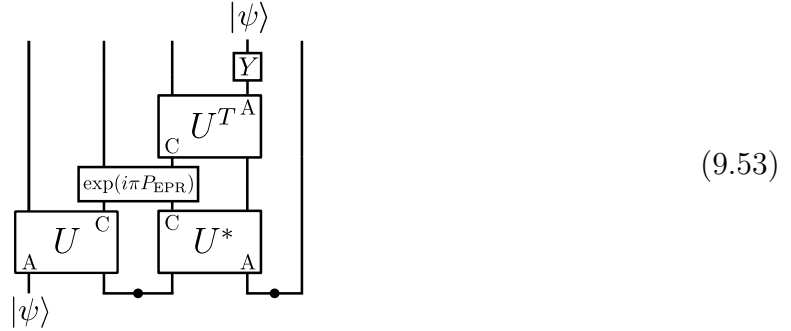
where $P_{\mathrm{EPR}}$ projects onto an EPR pair between subsystems C on the left and right sides.

The equivalence between this protocol and the TW protocol [Eq. (9.16)] is manifest, with the only difference being the locality of the coupling. Specifically, the HPR coupling is of the same general form as the TW coupling [Eq. (9.1)]:

$$gV \equiv \pi P_{\mathrm{EPR}} = \frac{\pi}{d_C^2} \sum_{P_C} P_{C,l}\, P_{C,r}^*, \tag{9.54}$$

where the sum is over of a complete basis of $d_C^2$ Pauli operators on C. However, the operators $P_C$ are typically non-local across C, whereas the coupling considered in the TW protocol was restricted to local operators. As a consequence, the HPR coupling functions as a binary variable measuring whether or not an operator has support on subsystem C (see Section 9.4). In contrast, the TW coupling measures the operator size within C, which takes an approximately continuous range of values when C is large. Crucially, at late times under scrambling dynamics, the effect of both couplings will be the same: to apply an overall phase to non-identity operators.

A few additional remarks are in order. First, while the leading order effect of the HPR and TW couplings is the same, they lead to different finite-size corrections. In particular, in a fully scrambled system, the variance in the phases applied by the HPR coupling is equal to the probability of a random Pauli string not having support on C, which is suppressed exponentially in the size of C, i.e. $1/d_C^2$. On the other hand, the variance in phases applied by the TW coupling is suppressed only polynomially, by $\sim g^2 \delta S_K^2/K^2 \sim g^2 S_K/K^2 \sim g^2/K$ [see Eq. (9.49) and the discussion below Eq. (9.35)]. These enhanced phase fluctuations are relevant for finite-size implementations of the TW protocol, as discussed further in Section 9.9.

Second, it has previously been shown that an extended version of the HPR protocol allows for teleportation of *multiple* qubits at late times [517]. Because of the equivalence between the protocols, this extension would also allow for multi-qubit teleportation via the peaked-size mechanism. However, the enhanced channel capacity comes with a trade-off: the circuit complexity (measured by the number of applications of the unitary $U$) grows exponentially in the number of qubits to be teleported. As we will see in the following
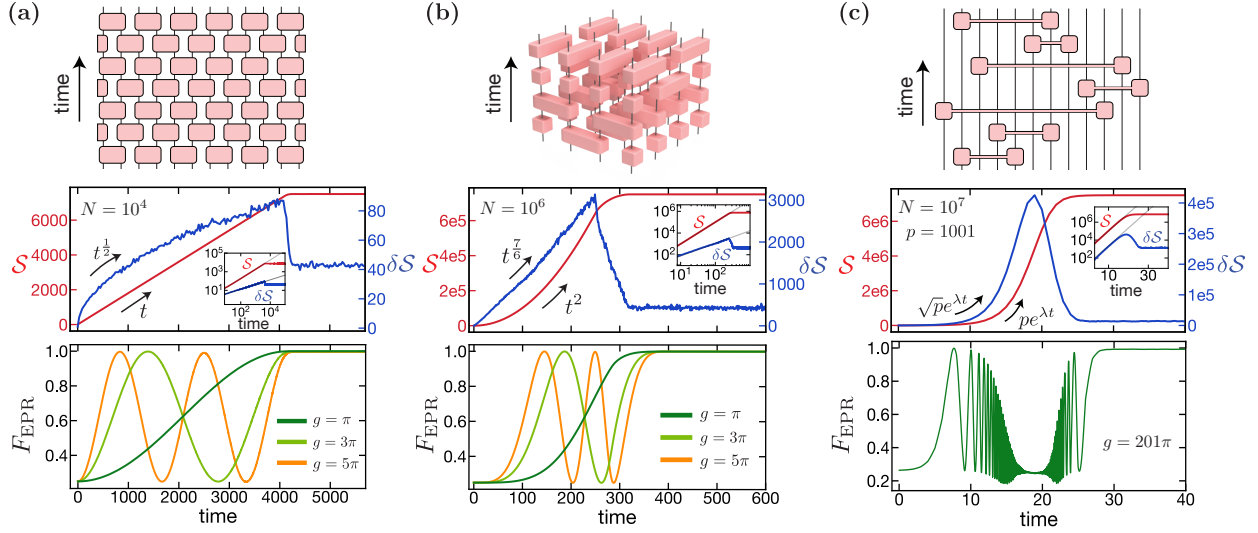
Figure 9.2:   Numerical results for averaged operator size and teleportation fidelity of 1D, 2D, and 0D RUCs. **(a-b)** In 1D and 2D, sizes grow ballistically in time, while the size width grows with a slower power of $t$ and matches predictions from the KPZ universality class (Section 9.7). Because of the separation between the size and size width, the teleportation fidelity for a single qubit exhibits an oscillatory behavior at intermediate times, with nearly perfect maximum fidelity. At late times, the teleportation fidelity saturates close to 1 for odd values of $g/\pi$, as expected for any scrambling system (Section 9.6). **(c)** In 0D all-to-all coupled RUCs, both the size and size width grow exponentially in time and obtaining a large separation between them requires encoding the initial state into $p$-body operators. With this encoding, the teleportation fidelity displays a distinct three-regime profile for $g \gg 1$. In particular, as in 1D and 2D, peaked-size teleportation succeeds $(i)$ at early times, with an oscillating fidelity, and $(ii)$ at late times, where the fidelity saturates close to 1 (for odd $g/\pi$). Between these regimes, no teleportation occurs because the size width has grown too large, $g\delta\mathcal{S}/N \gtrsim 1$.

section, this limitation can be overcome by peaked-size teleportation in the TW protocol at intermediate times, owing to the locality of the TW coupling.

## 9.7    Peaked-size teleportation at intermediate times

We now turn to analyzing the behavior of peaked-size teleportation at intermediate times, i.e. before the scrambling time. In this regime, multiple qubits can be teleported given a certain condition on the growth of time-evolved operators, namely when the overlap of the operators' support is sufficiently small.

We explicitly demonstrate that this condition is satisfied, and multi-qubit teleporation is possible, in a wide variety of physical systems at infinite temperature. These include random unitary circuits (RUCs) in $\geq$1D, for which peaked sizes naturally occur due to local thermalization within each operator's light cone, and time-evolved operators are non-overlapping due to spatial locality. More surprisingly, we show that multi-qubit peaked-size teleportation can also be achieved in 'fast scrambling', all-to-all coupled systems, including 0D random unitary circuits and the SYK model (at infinite temperature) [202, 431]. In this case, operators are *not* spatially separated at any nonzero time; nonetheless, the overlap of their size distributions remains *probabilistically* small at sufficiently early times. Furthermore, we demonstrate that while size distributions of local operators are generically *not* tightly peaked in all-to-all systems, peaked size distributions can be engineered in the TW protocol by *encoding* one's initial state into large $p$-body operators.

Finally, we consider the channel capacity—i.e. the maximum number of qubits that can be teleported (allowing both $g$ and $t$ to vary)—of peaked-size teleportation in all-to-all coupled systems. This is an essential question for comparing the capabilities of peaked-size teleportation with those of gravitational teleportation in traversable wormholes [317]. Remarkably, we provide analytic and numerical evidence that the channel capacity of peaked-size teleportation in 0D RUCs, a quite simple microscopic system, is asymptotically equivalent to that of the gravitational mechanism! Namely, the number of qubits $n$ that can be teleported scales with the number of couplings in the protocol, $n \sim K$.

## Multi-qubit teleportation: additive operator sizes

We begin with a few simple examples of multi-qubit teleportation to build intuition. First, consider a unitary $U$ that factorizes as $U = U_1 \otimes \cdots \otimes U_n$, where each $U_i$ acts on a disjoint subsystem. If we insert $n$ qubits individually into the $n$ different subsystems, then the entire protocol decouples into $n$ independent channels and there is no restriction on sending multiple qubits. This trivial example relies on the fact that $U$ does not scramble information across the entire system but only within each disjoint subsystem. We see that full scrambling of information by $U$ in fact *inhibits* the teleportation protocol's channel capacity (considered for a fixed set of qubits and dynamics).

A similar situation occurs even when the dynamics are not factorizable, as long as the teleported qubits are in causally separated regions. For example, consider a $(\mathcal{D} \geq 1)$-dimensional system with short-range interactions, where the inserted qubits are spatially separated. At intermediate times, the time-evolved qubit operators will have support within a local 'light cone' about their initial location, but will continue to act on disjoint subsystems. This scenario is therefore no different from the previous example and multi-qubit teleportation remains possible, as long as ($i$) the size distributions of each operator is tightly peaked, ($ii$) the coupling $V$ has support within each qubit's light cone, and ($iii$) the light cones of each qubit are non-overlapping. This final requirement constrains the number of qubits that can be sent at a given time $t$. In particular, the light cone of each operator will have a radius $v_B t$ where $v_B$ is the butterfly velocity. The maximum number of non-overlapping light cones—

equal to the total number of qubits $n$ that can be teleported—is therefore $n \lesssim N/(v_B t)^{\mathcal{D}}$, where $N$ is the total system volume.

More formally, we can analyze the success of $n$-qubit teleportation using the two-sided correlators, $C_Q$. We are concerned with $n$-qubit operators $Q(t) = Q_1(t) \dots Q_n(t)$, where each $Q_i \in \{I, X, Y, Z\}$ is a single-qubit Pauli on the $i^{\text{th}}$ teleported qubit. We work at infinite temperature and assume that sizes are tightly peaked. Teleportation therefore succeeds whenever all correlators have the same phase.

Inspired by the example of $n$ decoupled protocols, we will take the decoding operator to be the tensor product, $D = Y \otimes \dots \otimes Y$. The combination of transposition and conjugation by $D$ thus applies a minus sign to every single-qubit non-identity Pauli operator. An additional phase is applied by coupling proportional to the size of each operator. For example, for $n = 2$ qubits, we have:

| $Q_A$ | $C_Q$ |
|---|---|
| $\mathbb{1} \otimes \mathbb{1}$ | $1$ |
| $Q_1 \otimes \mathbb{1}$ | $-1 \times e^{-i\eta_d g \mathcal{S}_1/N}$ |
| $\mathbb{1} \otimes Q_2$ | $-1 \times e^{-i\eta_d g \mathcal{S}_2/N}$ |
| $Q_1 \otimes Q_2$ | $(-1)^2 \times e^{-i\eta_d g \mathcal{S}_{12}/N}$ |

where $\mathcal{S}_i$ and $\mathcal{S}_{ij}$ are shorthand for $\mathcal{S}[Q_i(t)]$ and $\mathcal{S}[Q_i(t)Q_j(t)]$. In order for all correlators to have the same phase, we require that $\eta_d g \mathcal{S}_1/N = \eta_d g \mathcal{S}_2/N = \pi$ mod $2\pi$, and that the operator sizes *add*, such that $e^{-i\eta_d g \mathcal{S}_{12}/N} \approx e^{-i\eta_d g(\mathcal{S}_1+\mathcal{S}_2)/N} = e^{i(\pi+\pi)} = 1$.

This requirements generalize straightforwardly to $n$ qubits. Specifically, teleportation succeeds whenever the single-qubit operator sizes obey $\eta_d g \mathcal{S}_i/N = \pi$ mod $2\pi$ and the multiqubit operator sizes *add* under operator multiplication:

$$\begin{aligned} \mathcal{S}[Q_1(t)Q_2(t)\dots Q_n(t)] \\ \approx \mathcal{S}[Q_1(t)] + \mathcal{S}[Q_2(t)] + \dots + \mathcal{S}[Q_n(t)]. \end{aligned} \tag{9.55}$$

This latter requirement implies that the phases applied by the coupling, $e^{igV}$, factorize, and allows the $n$ qubits to be teleported 'in parallel' as in the previous simple examples.

The size addition requirement naturally bounds the channel capacity in terms of the number of couplings, $K$. Specifically, the $K$-size takes integer values between 1 and $K$. However, the requirement that all three single-qubit Pauli operators have the same $K$-size increases the minimum $K$-size to 2. From Eq. (9.55), this implies that an $n$-qubit operator has a $K$-size of at least $2n$, which is only possible if

$$2n \leq K. \tag{9.56}$$

Indeed, this strict upper bound can also be understood from an information theoretic perspective: teleporting $n$ qubits requires an increase of $2n$ in the mutual information between the left and right sides of the system. Each of the $K$ classical bits sent from left to right in Fig. 12.1(a) increases the mutual information by at most 1, so at least $2n$ bits are required.

## ≥1D random unitary circuits

As a first concrete example of intermediate time peaked-size teleportation, we consider a random unitary circuit (RUC) applied to a lattice of $N$ qubits in one or higher dimensions. At each time step, pairs of neighboring qubits are evolved via independent Haar random unitaries arranged in a 'brick-layer' fashion, with periodic boundary conditions [Fig. 9.2(a,b)]. Operator growth in such systems has been studied at great length, and is believed to be a good model for many aspects of information scrambling under Hamiltonian dynamics [246, 292, 340, 385, 439, 476]. We extend these previous studies by demonstrating new results on the behavior of the operator size *width*—i.e. power-law scaling at intermediate times and suppression at late times—which we show can be detected by the teleportation fidelity (Fig. 9.3).

A key property of Haar random unitary circuits is that the expectation values of many circuit quantities can be computed by replacing the Haar random unitaries with randomly chosen *Clifford* unitaries, thereby enabling efficient classical simulation [122, 340]. Generally, this equivalence holds for any quantity that contains no more than *two* copies each of $U$ and $U^\dagger$ (e.g. the Renyi-2 entropy, or the OTOC); however, for systems of qubits, this property holds for up to three copies [271, 484, 534]. From Eq. (9.42), we see that the teleportation fidelity contains three copies of $U$ and $U^\dagger$, so the average fidelity is efficiently simulable[10]. Moreover, by definition, the size distributions of operators under Clifford dynamics are perfectly tightly-peaked, since a Pauli operator $Q_A$ evolved under a Clifford unitary remains a single Pauli string. Hence, the teleportation fidelity can be computed using the simplified expression given in Eq. (9.45).

In more detail, we calculate the average EPR fidelity for teleporting $n$ qubits through the following procedure. First, we choose a particular realization of $U$ by sampling each 2-qubit unitary from a uniform distribution of 2-qubit Clifford unitaries. Second, we determine the $K$-size of $UQ_AU^\dagger$ for each $n$-qubit Pauli operator, $Q_A$, or, if $n$ is large, for a random subset of these operators; such simulations can be performed efficiently with a time cost that scales linearly with the circuit depth. Third, we compute the fidelity for a given coupling $g$ using Eq. (9.45), with the phases $\theta_Q = \eta_d g \mathcal{S}_K / K + \pi \mathcal{S}[Q_A(0)]$, where the latter term captures the fact that decoding and transposition apply a minus sign for each non-identity element of the initial $Q_A$. Finally, we average the EPR fidelity over multiple realizations of $U$.

The results of these simulations for $n = 1$ qubit in 1D and 2D are shown in Fig. 9.2(a,b). As expected, the average operator size grows ballistically, $\mathcal{S} \propto t^{\mathcal{D}}$, until the operator's light cone reaches the edge of the system, at which point the size saturates to $3/4N$. While the behavior of the size width is more complex, in both dimensionalities it grows more slowly than the average size. This implies that the size distribution is tightly-peaked and the teleportation fidelity can be approximated by $F = \frac{5}{8} - \frac{3}{8}\cos(\eta_d g \mathcal{S}/N)$ [Eq. (9.46)]. We verify that the time profile of the fidelity follows this prediction, and nearly perfect fidelity is achieved when $\eta_d g \mathcal{S}/N = \pi \mod 2\pi$. In Appendix D.5, we also demonstrate that

---

[10]For higher-dimensional qudits, while we cannot efficiently simulate the teleportation fidelity, we can still calculate the correlators Eq. (9.20), which lower bound the fidelity via Eq. (9.43).
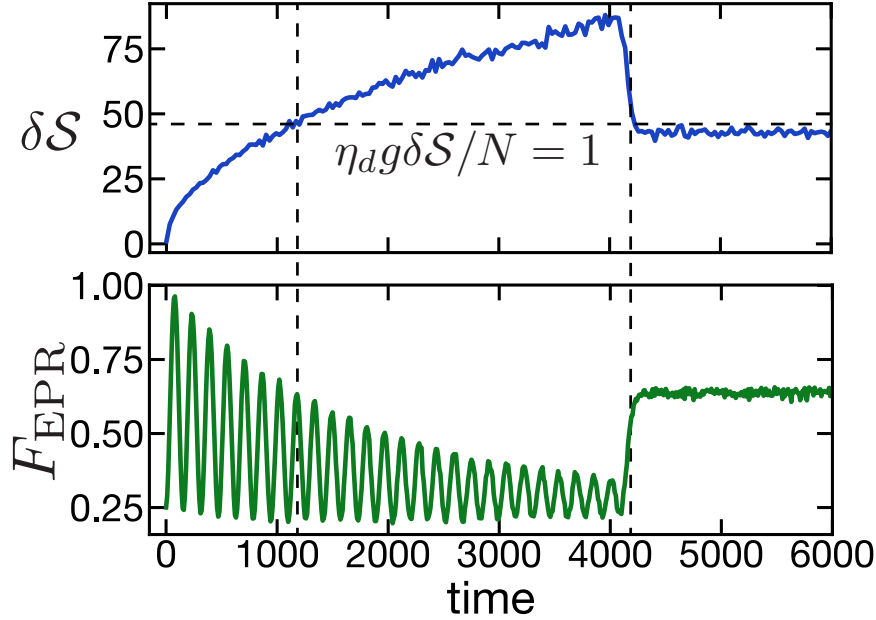
Figure 9.3:   Probing operator size width in a 1D RUC. (top) The size width initially grows
as $t^{1/2}$ and reaches a peak at the scrambling time $t^* \sim N = 10000$. (bottom) We probe
this behavior by measuring the teleportation fidelity of a single qubit with a large coupling
$g = 57\pi \sim \sqrt{N}$. The fidelity exhibits a distinct decay-revival profile, controlled by whether
the size width has exceeded the threshold $g\delta\mathcal{S}/N \approx 1$: nearly perfect fidelity initially, power
law decay towards a trivial fidelity at intermediate times, and partial revival at late times.

teleportation of $n > 1$ qubits is also possible at intermediate times, as long as their light
cones do not overlap.

   *Probing the size width*—Let us now turn to the time profile of the size width, which
exhibits a *peak* near the scrambling time in both 1D and 2D. Qualitatively, this behavior
arises from fact that the size width receives contributions from two sources: the interior of
the light cone, and the boundary of the light cone. Within the light cone, we expect a $\geq$1D
system with a small local Hilbert space to 'locally thermalize' as the operator spreads. This
implies that the bulk's contribution to the size width scales as $\delta\mathcal{S}_{\mathrm{bulk}} \propto \sqrt{\mathcal{S}} \propto t^{\mathcal{D}/2}$ and
saturates at the scrambling time. Second, the size width also receives contributions from the
light cone's boundary, which has not yet thermalized. At late times, the boundary of the
light cone reaches the edge of the system and these additional contributions subside, leading
to the peak in the size width at the scrambling time.

   To quantify these effects, we note that the growth of operators in $\geq$1D RUCs is predicted
to fall in the KardarâĂŞParisiâĂŞZhang (KPZ) universality class [238, 340]. In 1D, fluctu-
ations in the light cone boundary have been verified numerically to have a growing width

$\sim t^\alpha$ with the KPZ growth exponent $\alpha = 1/2$ [340]. This implies that the contribution of the boundary to the size width is $\delta S_{\text{boundary}} \propto t^{1/2}$, and the full width is

$$\delta S = \begin{cases} (\alpha_{\text{bulk}} + \alpha_{\text{boundary}})t^{1/2}, & t \lesssim t_{\text{scr}} \\ \alpha_{\text{bulk}} t_{\text{scr}}^{1/2}, & t \gtrsim t_{\text{scr}} \end{cases} \tag{9.57}$$

We note that the maximum size width relative to the late-time size width is a constant set by $(\alpha_{\text{bulk}} + \alpha_{\text{boundary}})/\alpha_{\text{bulk}}$. Comparing the size width of multiple system sizes, we observe excellent agreement with predicted scalings over a wide range of system sizes (Appendix D.5).

The time profile of the size width is directly observable in the peaked-size teleportation fidelity if we scale $g \sim t_{\text{scr}}^{1/2} \sim N^{1/2}$. In particular, by setting $N/g$ to lie between the maximum size width and the late time size width, we observe a distinct decay-revival profile for the teleportation fidelity (Fig. 9.3). At early times, we observe successful teleportation with an oscillating fidelity. The fidelity decays slowly, as a power law in time, as it receives corrections proportional to the growing size variance $\sim g^2 \delta S^2/N^2$. After the scrambling time, we see a revival in the teleportation fidelity as the size width narrows. The lack of a parametric separation between the maximum and late time size widths means that late time teleportation will also have some finite error for this value of $g$.

In 2D, we find that the scaling of the size width also matches predictions from the KPZ universality class. In this case, the width of the boundary scales as $\sim t^\alpha$, with $\alpha = 1/3$ [340]. However, to calculate the boundary's contribution to the size width, one must take into account two additional considerations. First, the boundary is 1-dimensional, so its length trivially grows in time as $\sim t$. Second, fluctuations of the boundary are expected to have a finite correlation length, $\xi \sim t^{1/z}$, where $z = 3/2$ is the KPZ dynamic exponent [109]. Thus, the boundary can be modeled as $n_\xi \sim t/\xi = t^{1/3}$ uncorrelated regions, each of length $\xi$. Each region contributes $\sim \xi t^\alpha$ to the size width; adding the uncorrelated contributions from all regions yields a total size width $\delta S \sim \sqrt{n_\xi}\, \xi\, t^\alpha = t^{1/6+2/3+1/3} = t^{7/6}$.

The time profile of the size width in 2D is thus given by

$$\delta S = \begin{cases} \beta_{\text{bulk}}t + \beta_{\text{boundary}}t^{7/6}, & t \lesssim t_{\text{scr}} \\ \beta_{\text{bulk}}t_{\text{scr}}, & t \gtrsim t_{\text{scr}} \end{cases} \tag{9.58}$$

We confirm these scalings in our numerics (Fig. 9.2(b) and Appendix D.5). Notably, the size width is now dominated by the boundary contribution at intermediate times, such that the ratio of the maximum size width to the late time size width scales as $t_{\text{scr}}^{1/6} \sim N^{1/12}$. As in 1D, one can probe this behavior using the peaked-size teleportation fidelity, now with $g \sim N/t_{\text{scr}}^{7/6} \sim N^{5/12}$. We emphasize that in 2D, the scaling of the size width is determined by *correlations* between different points on the light-cone boundary. This goes beyond the behavior studied in previous works on RUCs, which focus on quantities probed by local OTOCs.

## 0D random unitary circuits

We now turn to random unitary circuits in zero dimensions, a prototypical model for 'fast scramblers' [202, 431]. These circuits are constructued as follows: at each time-step, we partition the $N$ qubits into randomly chosen pairs, and apply independent Haar random 2-qubit unitaries to each pair.

Below we analyze such circuits using theoretical arguments, in combination with numerical simulations via Clifford circuits. As the later parts of our analysis are rather technical, we briefly summarize the main results: ($i$) peaked size teleportation remains possible but only if the input state is initially encoded in non-local, $p$-body operators; ($ii$) even though there is no complete separation of operator light cones, size addition still occurs at intermediate times in a probabilistic sense and enables mutli-qubit teleportation; and ($iii$) the maximum channel capacity is linear in the number of coupled qubits, $K$. These results are depicted numerically in Fig. 9.2(c) and 9.4.

*Peaked sizes*—In all-to-all coupled systems, operators are generally expected to grow exponentially in time, $\mathcal{S} \sim e^{\lambda t}$, where $\lambda$ is the Lyapunov exponent [431]. The reason is simple: at each time step, every term in an operator—rather than just those on a 'light-cone' boundary—has a fixed probability of spreading under random pairwise unitaries. A somewhat less intuitive expectation is that the size width also generally grows exponentially [379]. One way of understanding this is by imagining two realizations of the dynamics: in one realization the initial operator doubles at the first time and in the other it does not. In effect, the latter system now lags behind the former by one time step, $\Delta t$, and the difference in their sizes at later times will be exponentially magnified, to $e^{\lambda t}(1 - e^{-\lambda \Delta t})$.

The lack of separation between the size and size width seems to preclude the possibility of peaked-size teleportation at intermediate times. Nevertheless, we can engineer such a separation by *encoding* the information of each input qubit into $p$-body operators, with $p \gg 1$ [159]. As an example, consider encoding a single qubit into $p = 5$ qubit operators via

$$
\begin{aligned}
E(X \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1})E^\dagger &= Z \otimes X \otimes X \otimes Y \otimes Z \\
E(Y \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1})E^\dagger &= Y \otimes Z \otimes Z \otimes X \otimes Y \\
E(Z \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1})E^\dagger &= X \otimes Y \otimes Y \otimes Z \otimes X,
\end{aligned}
\tag{9.59}
$$

Here, $E$ is a Clifford unitary encoding operation that conjugates state insertion and decoding [explicitly, replacing $U \to UE, U^* \to U^*E^*$, and $U^T \to E^T U^T$ in Fig. 12.1(a)]. The success of teleportation is now dependent on the size distributions of time-evolved $p$-body operators, $Q_A(t) = UEPE^\dagger U^\dagger$, where $P$ runs over the initial unencoded single-qubit Pauli operators. As we will soon verify explicitly, before the scrambling time the support of each of the $p$ operators composing $Q_A$ will be approximately non-overlapping, so that their size distributions will convolve. Thus, the total operator size is multiplied by a factor of $p$ but, through the central limit theorem, the size width is multiplied only by $\sqrt{p}$.

In more detail, consider the size growth of an operator, $Q_A$, with initial size $\mathcal{S}_0 = p$. During a single time step, each qubit $i$ in the support of $Q_A(t)$ is paired with another

random qubit; for simplicity, we assume the second qubit is outside the support of $Q_A(t)$, which should be valid at times well before the scrambling time. Under random two-qubit Clifford time-evolution, $Q_A(t)$ grows to have support on both qubits with probability $\nu = 1 - 2(d^2 - 1)/(d^4 - 1)$ (9/15 for qubits). The operator size, $\mathcal{S}_t$, therefore grows stochastically in time, according to

$$
\begin{aligned}
\mathcal{S}_{t+1} &= \mathcal{S}_t + \sum_{i=0}^{\mathcal{S}_t} s_i \\
&= \mathcal{S}_t + \text{Bi}_t(\mathcal{S}_t, \nu) \\
&\approx (1 + \nu)\mathcal{S}_t + \sqrt{\mathcal{S}_t \nu(1 - \nu)}\,\mathcal{N}_t(0, 1)
\end{aligned}
\tag{9.60}
$$

where each $s_i$ is a binary random variable that increases the size by 1 with probability $\nu$ and 0 with probability $1 - \nu$, and $\text{Bi}_t(\mathcal{S}_t, \nu)$ denotes the binomial distribution with $\mathcal{S}_t$ trials and probability $\nu$, which we can approximate as a normal distribution, $\mathcal{N}_t(\nu\mathcal{S}_t, \sqrt{\mathcal{S}_t\nu(1 - \nu)})$. The size at time $t$ can thus be written as a sum of random variables drawn at each time step:

$$
\begin{aligned}
\mathcal{S}_t \approx &(1 + \nu)^t p \\
&+ \sqrt{\nu(1 - \nu)} \sum_{t'=0}^{t-1} (1 + \nu)^{t - t' - 1} \sqrt{\mathcal{S}_{t'}}\,\mathcal{N}_{t'}(0, 1)
\end{aligned}
\tag{9.61}
$$

from which we see that the average size grows exponentially in time with Lyapunov exponent $e^\lambda = 1 + \nu$. Deviations arise at each time step $t'$, with typical magnitude $(1 + \nu)^{t - t' - 1}\sqrt{\mathcal{S}_{t'}} \approx (1 + \nu)^{t - 1 - t'/2}\sqrt{p}$. Since this decays exponentially in $t'$, we can approximate the total variation, $\delta\mathcal{S}_t$, as the largest term in the sum ($t' = 0$), which has magnitude

$$
\delta\mathcal{S}_t \sim (1 + \nu)^{t - 1}\sqrt{p} \approx \frac{\mathcal{S}_t}{\sqrt{p}}.
\tag{9.62}
$$

As anticipated, the size width is dominated by early time errors that have exponentially grown in time, so that the ratio of the size width to the size remains constant at $\sim 1/\sqrt{p}$ (after some period of growth from its initial value, 0).

To support these claims, we numerically simulate the time-evolved size distribution of operators with an initial size $p \approx 1000$ [Fig. 9.2(c)]. As expected, we observe that the average size grows exponentially as $\sim pe^{\lambda t}$ and saturates at a timescale $t^* \sim \log(N/p)$. Moreover, the size width grows at the same exponential rate but its magnitude is suppressed by a factor of $\sqrt{p}$ compared to the average size.

To verify that this allows for teleportation, we next compute the fidelity for teleporting a single qubit, in the regime $g \gg 1$. As shown in Fig. 9.2(c), teleportation occurs with near perfect fidelity beginning at $t \approx t^* - \log(gp)$, corresponding $g\mathcal{S}/N \approx 1$. Thereafter, the teleportation fidelity decreases exponentially in time, consistent with the increase of the size width. At time $t \approx t^* - \log(g\sqrt{p})$, teleportation stops succeeding entirely, since the size width

has reached the limit $\delta\mathcal{S}/N \sim 1$. Finally, at late times $t \approx t^* - \log(p)$, the fidelity revives as the system becomes fully scrambled and the operator size width narrows to $\delta\mathcal{S} \sim \sqrt{\mathcal{S}}$.

*Size addition*—We now turn to the possibility of teleporting multiple qubits in 0D RUCs. Within the peaked-size regime, this reduces to the question of whether operator sizes add according to Eq. (9.55). Satisfying this requirement in all-to-all coupled systems is not as trivial as in $\geq 1$D, since time-evolved operators typically act on overlapping subsystems at any finite time. Nevertheless, we now provide a simple argument for why size addition holds despite this.

To do so, we model each time-evolved Pauli operator $Q_i(t)$ as an independent random Pauli string of size $\mathcal{S}[Q_i]$. Consider two such strings, $P_1$ and $P_2$, with support on regions $A_1$ and $A_2$ and sizes $\mathcal{S}[P_1] = |A_1|$ and $\mathcal{S}[P_2] = |A_2|$. The size of the product, $P_1 P_2$, is the size of the union $A_1 \cup A_2$, minus the number of sites where the two strings overlap and have the same single-qubit Pauli operator. This occurs with probability $1/(d^2 - 1) = 1/3$ at each site in the region $A_1 \cap A_2$, giving

$$\begin{aligned} \mathcal{S}[P_1 P_2] &\approx |A_1 \cup A_2| - \frac{1}{3}|A_1 \cap A_2| \\ &= \mathcal{S}[P_1] + \mathcal{S}[P_2] - \frac{4}{3}|A_1 \cap A_2|. \end{aligned} \tag{9.63}$$

The deviation from the simple additive rule $\mathcal{S}[P_1 P_2] = \mathcal{S}[P_1] + \mathcal{S}[P_2]$ is thus controlled by $|A_1 \cap A_2|$. If the Pauli strings $P_1, P_2$ have independently random areas of support, the size of this intersection scales as:

$$|A_1 \cap A_2| \sim \mathcal{S}[P_1]\mathcal{S}[P_2]/N, \tag{9.64}$$

which is subleading to $\mathcal{S}[P_i]$ at intermediate times (when $\mathcal{S}/N \ll 1$). To derive this, note that the probability for *both* strings to have support on a given qubit is $\sim (\mathcal{S}[P_1]/N)(\mathcal{S}[P_2]/N)$; summing over $N$ qubits gives the above result.

For $n$-qubit teleportation, one must consider the combined size, $\mathcal{S}[P_1 \ldots P_m]$, of $m$ independent Pauli strings, where $m$ takes a typical value $m \approx 3n/4$ (a typical $n$-qubit operator has non-identity support on $3n/4$ qubits). In general, this quantity will receive corrections from $\binom{m}{k}$ different $k$-way intersections of the strings, for all $2 \leq k \leq m$. For random Pauli strings, the expected size of these intersections scales as $N\overline{|A_1 \cap \ldots \cap A_k|} = \prod_{i=1}^k \frac{|A_i|}{N} \sim \mathcal{S}^k/N^{k-1}$, where $\mathcal{S} \sim |A_i|$ is the typical size of a single Pauli string [see Eq. (9.64) above]. For a given $k$, the correction to size addition will be the sum of $\binom{m}{k} \sim m^k$ different intersections and therefore scales as $m\mathcal{S}(m\mathcal{S}/N)^{k-1}$. These corrections can be neglected if they are small compared to the total size; this occurs when $m\mathcal{S} \ll N$, which corresponds to a timescale much less than the scrambling time.

To demonstrate this claim, we numerically simulate the teleportation protocol with $n > 1$ qubits in the regime $1 \ll p, np \ll K$ [Fig. 9.4]. Analogous to single-qubit teleportation, the teleportation fidelity exhibits oscillations beginning at $t \approx t^* - \log(gp)$, and vanishes at $t \approx t^* - \log(g\sqrt{pn})$ due to the growth of the combined size width. However, in contrast

to the single-qubit case, teleportation of multiple qubits is not possible at late times, $t \gtrsim t^* - \log(gpn)$, as predicted in Section 9.6. Interestingly, between these two regimes, we observe a partial revival of the fidelity: this indicates that the operator size widths begin to narrow before the additive condition is completely invalidated.

*Error analysis*—While we have confirmed that multi-qubit teleportation can be achieved in certain ideal limits, a key question remains: how does the maximum number of qubits that can be teleported scale as a function of $K$, i.e. what is the protocol's channel capacity? To answer this question, we now estimate how deviations from these ideal limits lead to errors in peaked-size teleportation and ultimately constrain the channel capacity. Throughout this discussion, we assume that the size, $\mathcal{S}$, is extensive, but $K$ is not; this is the natural regime for probing the channel capacity of the protocol at intermediate times, and is the physical scenario in the context of traversable wormholes [317]. The details of this and the following subsection are quite technical in nature, and may be skipped by most readers.

In summary, we identify four distinct sources of error in the multi-qubit teleportation fidelity, $F = 1 - \epsilon$:

1. Errors due to finite $p$: $\epsilon \sim ng^2\mathcal{S}_K^2/K^2p$

2. Errors due to finite $K$: $\epsilon \sim ng^2\mathcal{S}_K/K^2$

3. Errors due to imperfect size addition: $\epsilon \sim \left[n^2g^2\mathcal{S}_K^4/K^4 + \dots\right]$, where ellipses indicate higher orders in $(n\mathcal{S}_K/K)^2$

4. Errors due to fluctuations in size addition: $\epsilon \sim \left[n^2g^2\mathcal{S}_K^2/K^3 + \dots\right]$, where ellipses indicate higher orders in $n\mathcal{S}_K/K$

We discuss each of these errors in detail below.

The first and second sources of error are due to imperfectly peaked $K$-size distributions. The $K$-size width receives contributions from finite-$p$ corrections, $\sim \mathcal{S}_K/\sqrt{p}$, and finite-$K$ corrections, $\sim \sqrt{\mathcal{S}_K}$ [see the discussion below Eq. (9.35)]. To translate these into errors in the teleportation fidelity, we multiply the size width by $g/K$ and take the square. This gives fidelity errors $\sim g^2\mathcal{S}_K^2/pK^2$ and $\sim g^2\mathcal{S}_K/K^2$ per teleported qubit.

The third and fourth sources of error arise from imperfect size addition. This leads both to 'systematic' errors, due to the average overlap of operators, as well as 'sampling' errors, due to random fluctuations in this overlap. We begin with the systematic errors: as we recall, the size addition of $m$ time-evolved operators receives corrections from $k$-way overlaps of the operators, each scaling as $\sim m\mathcal{S}_K(m\mathcal{S}_K/K)^{k-1}$, for $2 \leq k \leq m$ (rescaling our previous results to the $K$-size instead of the size). The nonlinear dependence on $m$ indicates that sizes do not add perfectly. Nevertheless, when teleporting an $n$-qubit initial state for large $n$, we can correct for the above effect at leading order by using a linear approximation for $m^k$ about its typical value, $(3n/4)^k$. This leads to an effectively smaller operator size, which can be observed in the reduced frequency of the fidelity oscillations for 10-qubit teleportation compared to 1-,3-qubit teleportation in Fig. 9.4(a). The leading errors after this shift are

quadratic in $\delta m \equiv m - 3n/4$, which has a typical magnitude $\delta m \sim \sqrt{n}$. Multiplying by $g/K$ and taking the square, we therefore find multi-qubit fidelity errors $\sim (g\mathcal{S}_K/K)^2(n\mathcal{S}_K/K)^{2k-2}$; at leading order $k = 2$, this gives $\sim n^2 g^2 \mathcal{S}_K^4/K^4$.

Finally, each intersection above is subject to additional random fluctuations about its average value. When operator sizes are much smaller than the system size, we can treat each intersection as arising from a binomial process, in which case fluctuations are proportional to the square root of the intersection's average size (see Appendix D.6 for a detailed accounting). These add in quadrature for $\sim n^k$ overlaps, producing a total fidelity error $\sim (g^2/K)(n\mathcal{S}_K/K)^k$.

*Channel capacity*— To define the channel capacity of the teleportation protocol, we fix a per qubit error threshold $\epsilon_{\text{th}}$, and determine the maximum number of qubits that can be sent while maintaining a multi-qubit fidelity above this threshold[11], i.e. $F \geq 1 - n\,\epsilon_{\text{th}}$. We are interested in how the channel capacity scales with the number of couplings, $K$, while allowing both $g$ and $\mathcal{S}_K$ (determined by the evolution time) to vary.

In 0D RUCs, all errors increase with $g$, so it is optimal to set $g$ to its minimal value, $\eta_d g \mathcal{S}/N = \pi$. This gives a per qubit error

$$\frac{\epsilon}{n} \sim \frac{1}{p} + \frac{1}{\mathcal{S}_K} + \left[\frac{n\mathcal{S}_K^2}{K^2} + \dots\right] + \left[\frac{n^2}{K} + \dots\right]. \tag{9.65}$$

The first term is negligible in the large $p$ limit and so we will neglect it from here on.

We minimize the remaining terms with respect to $\mathcal{S}_K$. There are two relevant regimes. For $n \lesssim \sqrt{K}$, the minimum is determined entirely by the leading order contributions in $n\mathcal{S}_K/K$ to the error (i.e. neglecting the ellipses). Taking the derivative and setting to zero, we have the minimum at $\mathcal{S}_K^{(1)} \sim K^{2/3}/n^{1/3}$. As we increase $n$, the optimal size approaches the value $\mathcal{S}_K^{(2)} \sim K/n$. At this point, size addition errors of all orders (i.e. the ellipses) become large, and so the true minimum becomes fixed just below $\mathcal{S}_K^{(2)}$. This crossover between these two minima occurs at $n \sim \sqrt{K}$, at which $\mathcal{S}_K^{(1)} \sim \mathcal{S}_K^{(2)}$.

The above minima give two distinct scalings for the per qubit error and thus the channel capacity. The first minimum has a per qubit error $\epsilon^{(1)}/n \sim (n/K^2)^{1/3}$, which gives rise to a superlinear channel capacity, $n \lesssim \epsilon_{\text{th}}^3 K^2$. However, as we increase $K$, this capacity eventually surpasses the value $\sqrt{K}$. Above this, the optimal size is given by the second minimum, which has an error $\epsilon^{(2)}/n \sim n/K$, and thus the channel features an asymptotically linear capacity,

$$n \lesssim \epsilon_{\text{th}} K. \tag{9.66}$$

This is a stronger instance of the strict general bound Eq. (9.56). Intuitively, this channel capacity arises because the individual $K$-sizes must be large, $\mathcal{S}_K \gg 1$, for the $K$-size to be tightly peaked, while at same time the combined $K$-size must be much smaller than $K$, $n\mathcal{S}_K \ll K$, for the $K$-sizes to add; hence $n \ll K$.

---

[11]We note that this definition of channel capacity differs from more conventional definitions [353]; we do not expect this difference to qualitatively affect the scaling of the channel capacity with $K$, as the fidelity drops off steeply above the capacity [Fig. 9.4(b)].

We test this scaling numerically by simulating the teleportation protocol and measuring the per qubit fidelity, $F_{\text{EPR}}^{(1)}$, as a function of $n$ and $K$. Specifically, for each value of $K$, we sweep the number of qubits $n$ and determine the maximum qubits that can be sent before the infidelity exceeds a threshold, $1 - F_{\text{EPR}}^{(1)} = \epsilon_{\text{th}}$. These results are shown in Fig. 9.4(b) and exhibit a clear linear trend across two orders of magnitude, confirming our prediction of a linear channel capacity.

A few final remarks are in order. First, while in principle the per qubit fidelity can be calculated by taking the $n^{\text{th}}$ root of the full $n$-body fidelity, this approach is numerically unstable for large $n$. Thus, we instead compute the fidelity of a *single* qubit, while trying to send multiple qubits, using an approach derived in Appendix D.5. This amounts to performing a sum analogous to Eq. (9.45), but only including pairs of $Q_1$ and $Q_2$ that are equal on all sites except for one.

Second, the range of system parameters that lie within the linear scaling regime is ultimately constrained by the finite total system size, $N = 10^8$. In particular, to maximize the linear scaling regime, we choose $p = 101$ and $\epsilon_{\text{th}} = 0.07$. The former ensures that finite-$p$ errors are negligible, while the latter allows the number of qubits at the threshold to be large enough to access the $n \gtrsim \sqrt{K}$ regime but small enough that the operators are initially dilute, i.e. $n \ll N/p$.

## Large-$q$ SYK model: infinite temperature

We now demonstrate peaked-size teleportation in a 0D Hamiltonian system, the large-$q$ SYK model, at infinite temperature. While teleportation at low temperatures in the SYK model is known to succeed via the gravitational mechanism, teleportation at infinite temperature was discovered only recently [159]. In addition to showing that this mechanism is in fact peaked-size teleportation, we also find that, remarkably, *all* qualitative aspects of this teleportation match those of 0D RUCs.

The large-$q$ SYK model is defined by the Hamiltonian [316, 379]:

$$H = i^{q/2} \sum_{1 \leq j_1 \leq \ldots \leq j_q} J_{j_1,\ldots,j_q} \psi_{j_1} \ldots \psi_{j_q}, \tag{9.67}$$

where $\psi_i$ are Majorana fermions, $\{\psi_i, \psi_j\} = 2\delta_{ij}$, and the couplings are drawn independently from a Gaussian distribution with zero mean and a variance $\langle J_{j_1,\ldots,j_q}^2 \rangle = J^2/2q\binom{N-1}{q-1}$. This model is exactly solvable at all temperatures in the large-$q$, large-$N$ limit [316, 379].

To construct the teleportation protocol for the SYK model, we first define the $N$-fermion EPR state,

$$\psi_{j,l} \,|\text{FEPR}\rangle \equiv -i\psi_{j,r} \,|\text{FEPR}\rangle, \quad \forall\, j = 1, \ldots, N \tag{9.68}$$

From this, the TFD state is obtained as before,

$$|\text{TFD}\rangle \equiv e^{-\beta H_l/2} \,|\text{FEPR}\rangle. \tag{9.69}$$
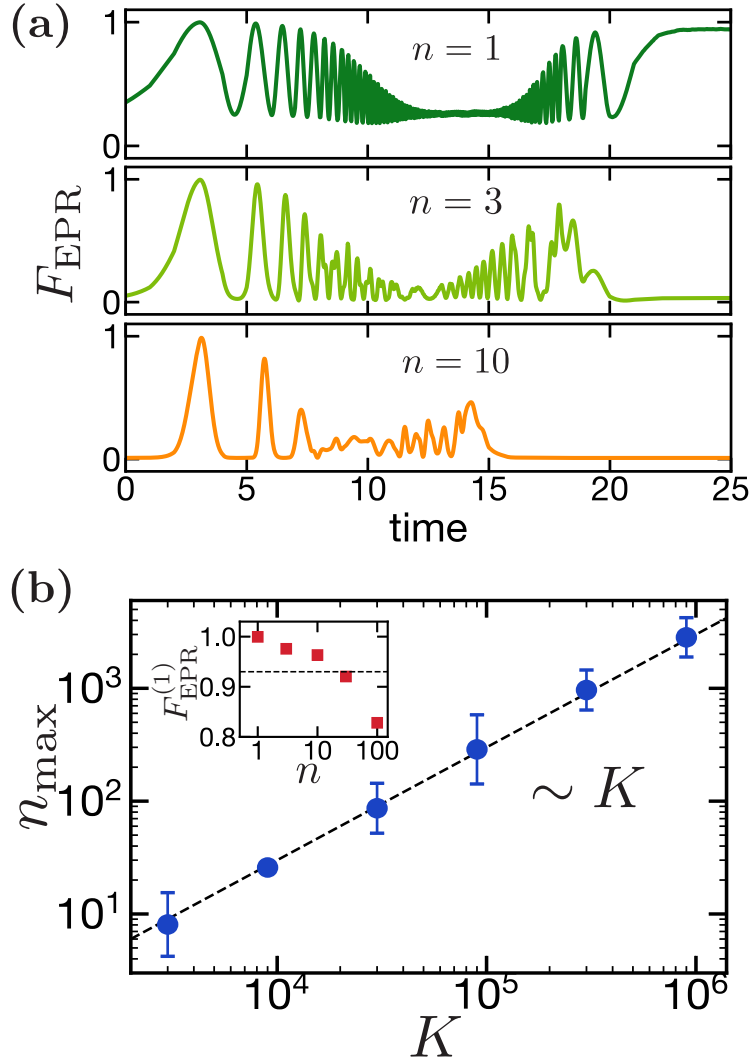
Figure 9.4:   Teleportation of multiple qubits in 0D RUCs. **(a)** Many-body teleportation
fidelity, $F_{\mathrm{EPR}}$, as a function of time for teleporting $n = 1, 3, 10$ qubits with fixed coupling
strength ($g = 177\pi$).   Compared to a single qubit, the decay-revival profile for multiple
qubits is shifted to earlier times, since multi-qubit operators both have a larger size width
and saturate the system size earlier.   Moreover, multi-qubit teleportation is not possible
at late times, resulting in a trivial late-time fidelity (Sec. 9.6).   **(b)** Numerical results for
the channel capacity $n_{\mathrm{max}}$ as function of the number of coupled qubits $K$, which exhibit a
clear linear scaling. To determine the channel capacity, we compute the maximum *per qubit*
fidelity $F_{\mathrm{EPR}}^{(1)}$ for a fixed number of qubits, $n$, and couplings, $K$, while allowing the coupling
strength, $g$, and evolution time to vary. For fixed $K$, $F_{\mathrm{EPR}}^{(1)}$ decreases as the number of qubits
$n$ is increased, as depicted in the inset for $K = 9000$. The channel capacity $n_{\mathrm{max}}$ is defined
as the maximum number of qubits for which the fidelity is above a fixed threshold (dashed
line).

For the two-sided coupling, we consider the simple bilinear interaction,

$$V = \frac{1}{2qN} \sum_{j=0}^{N} i\psi_{j,l}\psi_{j,r}, \tag{9.70}$$

which measures the size of operators in the Majorana string basis, divided by $qN$ [379, 397].

As in 0D RUCs, the size and size width of time-evolved operators in the SYK model increase exponentially in time, and exhibit a large separation only when initially encoded in $p$-body operators. To see this, we can generalize previous computations of size distributions in the large-$q$ SYK model [379] to initial $p$-body operators, $\psi = \psi_1\psi_2\dots\psi_p$; this relies on the factorization of SYK correlation functions in the large-$N$ limit [159]. After the relaxation time $(t \gtrsim 1/J)$, but before the scrambling time $(t \lesssim \log(N/p)/J)$, the size and size width are:

$$\mathcal{S} \approx \frac{p}{2}e^{2Jt}, \quad \delta\mathcal{S} \approx \frac{\sqrt{2qp}}{4}e^{2Jt}. \tag{9.71}$$

The scaling $\delta\mathcal{S} \sim \mathcal{S}/\sqrt{p}$ matches that found for 0D RUCs; in particular, ensuring a large separation between the size and size width requires $p \gg q$. Note that our condition for peaked size distributions depends on the (large) parameter $q$, through the size width.

This large separation suggests that peaked-size teleportation is possible at early times in the large-$p$ limit. To verify this, we analyze the two-sided correlator, which is given by [160]

$$C_\psi(t) = \langle e^{-igV}\psi_r(-t)e^{igV}\psi_l(t)\rangle$$

$$= \left(\frac{1}{1 + i\frac{g}{N}\frac{1}{4}e^{2Jt}}\right)^{2p/q} \tag{9.72}$$

at infinite temperature before the scrambling time[12]. For large $p$ and early times, we can approximate the correlator as

$$C_\psi(t) \approx \exp\left(-i\frac{g}{qN}\frac{p}{2}e^{2Jt}\right), \tag{9.73}$$

using $(1 + ix)^m \approx e^{imx}$, valid when $mx^2 \equiv \frac{2p}{q}\left(\frac{g}{N}\frac{1}{4}e^{2Jt}\right)^2 \ll 1$. We refer to this regime as the "early time regime", and analyze its analog in large-$N$ systems at finite temperature in Section 9.8.

Crucially, as expected for peaked-size teleportation, the early time correlator consists of an overall phase equal the average operator size, Eq. (9.71), multiplied by $g/qN$. This indicates that teleportation succeeds with nearly maximal fidelity beginning when $g\mathcal{S}/qN \approx 1$. Based on its similarity with 0D RUCs, we expect that teleportation in this regime is capable of teleporting $\mathcal{O}(K)$ qubits (Table 9.1); however, we do not calculate this explicitly.

---

[12]The inclusion of $e^{-igV}$ in the correlator applies a phase $e^{-ig}$ to the bra on the left side, which conveniently subtracts off the constant term in $V$'s relation to operator size [Eq. (9.26)].

Teleportation continues to succeed until the above approximation breaks down, which occurs when the size width, $\delta\mathcal{S}$, becomes of order $(g/qN)^{-1}$. As for all scrambling systems, the two-sided correlator is expected to revive at late times, $t \gtrsim \log(N/p)/J$, at which point the sizes saturate the entire system [160, 317] (see Section 9.6); this is not reflected in Eq. (9.72), which is valid only before the scrambling time.

## 9.8 Interplay between peaked-size and gravitational teleportation

In this section, we seek to understand the interplay between peaked-size and gravitational teleportation. A central theme in this understanding is a comparison between the size distribution introduced in Section 9.4, and the *winding size distribution* introduced in Ref. [71, 350].

To illustrate the distinction between these distributions, consider a time-evolved Majorana fermion operator, decomposed in a basis of Majorana strings, $\chi$ [379, 397]:

$$\psi(t)\rho^{1/2} = \sum_\chi c_\chi \chi. \tag{9.74}$$

From this decomposition, one defines the size distribution [379, 397],

$$P(\mathcal{S}) = \sum_{\chi\,:\,\mathcal{S}[\chi]=\mathcal{S}} |c_\chi|^2, \tag{9.75}$$

and the winding size distribution [71, 350],

$$f(\mathcal{S}) = \sum_{\chi\,:\,\mathcal{S}[\chi]=\mathcal{S}} c_\chi^2, \tag{9.76}$$

where $\mathcal{S}[\chi]$ is the size of the string $\chi$. Note that the size distribution is real-valued, while the winding size distribution may be complex.

The teleportation correlators [under coupling Eq. (9.70)] are, in fact, directly related to the winding size distribution [71, 350]:

$$C_\psi(t) = -i \sum_{\mathcal{S}=0}^{\infty} e^{-ig\mathcal{S}/qN} f(\mathcal{S}), \tag{9.77}$$

which can be derived by explicitly plugging Eq. (9.74) into the teleportation correlator. The size distribution, by contrast, is related to "one-sided" correlation functions, e.g. Eq. (9.32), where both instances of the time-evolved operator appear on the same side of the TFD state [379].

Despite this distinction, we have so far been able to analyze teleportation using the size distribution, as opposed to the winding size distribution, because the two are equal in two

circumstances. The first is at infinite temperature, where the coefficients $c_\chi$ are real because $\psi(t)$ is Hermitian. The second has been precisely our focus: when size distributions are perfectly tightly peaked, in which case both distributions approach a delta function.

In what follows, we describe several scenarios in which the distinction between the two distributions becomes relevant. First we begin in large-$N$ systems, where large-$N$ factorization provides a precise relation between the teleportation correlator and the OTOC at early times. We find that, even in the presence of the large-$p$ encoding, the correlator deviates from the peaked-size prediction whenever the OTOC contains an imaginary part. Large-$N$ systems encompass both peaked-size and gravitational teleportation—our results suggest that the former occurs in systems where the OTOC is real (e.g. at infinite temperature with large-$p$ encoding, see Section 9.7), while the latter occurs where the OTOC is imaginary (e.g. at low temperature in SYK) [181, 252]. Second, we review recent results showing that this deviation eventually leads an $\mathcal{O}(1)$ correlator magnitude when the winding size distribution takes a particular form, thereby enabling teleportation with unit fidelity (see Section 9.3). This is conjectured to be the microscopic origin of gravitational teleportation [71, 350], and so we expect it to occur only in low temperature models with a gravity dual. Third, we return to teleportation in the large-$q$ SYK model and show that this model interpolates between gravitational teleportation at low temperatures and peaked-size teleportation at high temperatures. Surprisingly, this interpolation occurs despite the fact that the large-$p$ encoding ensures a large separation between the size and size width, i.e. the size distribution naively appears tightly peaked, even at low temperatures. Finally, motivated by this smooth interpolation, we conclude this section by searching for a 'dual' description of peaked-size teleportation in a bulk gravitational theory. In particular, we argue that strong stringy effects lead to the same qualitative features as peaked-size teleportation.

## Early time teleportation in large-$N$ systems

In Section 9.4, we saw that for peaked-size operators the teleportation correlator depends only on the first moment of the size distribution, i.e. the average size [Eq. (9.32)]. We will now show that a more general relationship holds for large-$N$ systems at early times, where we substitute the average size with the first moment of the *winding* size distribution. Specifically,

using Eqs. (9.9-9.15), the first moment of the winding size is given by a two-sided OTOC:

$$\mathrm{OTOC}_2 = i\langle\mathrm{TFD}|\,\psi_{1,r}(-t)\,V\,\psi_{1,l}(t)\,|\mathrm{TFD}\rangle$$

$$= \frac{-1}{K}\sum_{i=1}^{K}\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\;\; \tag{9.78}$$

$$= \frac{1}{K}\sum_{i=1}^{K}\mathrm{tr}\!\left(\rho^{1/2}\,\psi_1(t)\,\psi_i\,\rho^{1/2}\psi_1(t)\,\psi_i\right)$$

using Eqs. (9.9-9.15). This differs from the one-sided OTOC, for probing the average size [Eq. (9.32)], in terms of the placement of the thermal density matrix.

To relate the OTOC and the teleportation fidelity, we consider two simplifying assumptions. First, we focus on 0D large-$N$ systems, e.g. the SYK model, with a $p$-body initial encoding. In such systems, the teleportation correlator in fact factorizes into a product of single-body correlators (up to $1/N$ corrections) [159]:

$$
\begin{aligned}
C_\psi(t) &= \left\langle e^{-igV}\psi_r(-t)e^{igV}\psi_l(t)\right\rangle \\
&\approx \left[\left\langle e^{-igV}\psi_{1,r}(-t)e^{igV}\psi_{1,l}(t)\right\rangle\right]^p,
\end{aligned}
\tag{9.79}
$$

where $\psi_1$ is a single-body operator.

Second, generalizing Eqs. (9.35, 9.73), we consider sufficiently early times to work at leading order in $g$[13]:

$$
\begin{aligned}
C_\psi(t) &\approx e^{-igp\langle V\rangle}\left[\langle\psi_{1,r}\psi_{1,l}\rangle + ig\langle\psi_{1,r}\,V\,\psi_{1,l}\rangle + \ldots\right]^p \\
&\approx e^{-igp\langle V\rangle}\langle\psi_{1,r}\psi_{1,l}\rangle^p\left[\exp\!\left(igp\frac{\langle\psi_{1,r}\,V\,\psi_{1,l}\rangle}{\langle\psi_{1,r}\psi_{1,l}\rangle}\right)+\ldots\right] \\
&= (-iG_\beta)^p\exp\!\left(-i\frac{gp}{2q}\left[\frac{\mathrm{OTOC}_2}{G_\beta}-G_\beta\right]\right)+\ldots \\
&= (-iG_\beta)^p\exp\!\left(-i\frac{gp}{2qN}G_\beta\mathcal{F}_2(t)\right)+\ldots
\end{aligned}
\tag{9.80}
$$

---

[13]In the first line, we use the fact that the thermofield double state has peaked size [379] to pull $e^{-igV}$ outside the correlator. In the second line we use the expansion $(1+ix)^m\approx e^{imx}$. In the third line we use $\langle V\rangle = i\langle\psi_{1,l}\psi_{1,r}\rangle/2q = -G_\beta/2q$.

where $G_\beta = i\langle\psi_{1,r}\psi_{1,l}\rangle = \mathrm{tr}(\rho^{1/2}\psi_1\rho^{1/2}\psi_1)$ is the imaginary time Green's function, and $\mathcal{F}_2(t)$ is the first-order, connected component of the two-sided OTOC [Eq. (9.78)],

$$\mathrm{OTOC}_2 \approx G_\beta^2 \left(1 + \frac{1}{N}\mathcal{F}_2(t) + \cdots\right). \tag{9.81}$$

Similar to Eq. (9.35), the leading correction to Eq. (9.80) is $\sim pg^2[\langle V^2\rangle_\psi - \langle V\rangle_\psi^2/G_\beta]$, and the approximation holds when this is small.

Let us now consider the behavior of the teleportation correlator, Eq. (9.80), under different physical scenarios. We focus on chaotic systems during the so-called Lyapunov regime, which occurs between the thermalization time, $t \sim \mathcal{O}(1)$, and the scrambling time, $t \sim \mathcal{O}(\log N)$. In this regime, the connected OTOC is characterized by a simple exponential $\mathcal{F}_2(t) \sim e^{\lambda t}$ with a prefactor that is generally complex. As a result, we expect the teleportation correlator to exhibit two distinct effects: ($i$) the real part of $\mathcal{F}_2(t)$ causes rapid phase oscillations in the teleportation correlator, while ($ii$) the imaginary part increases/decreases the teleportation correlator magnitude, depending on the sign of the coupling $g$.

At infinite temperature, $\mathcal{F}_2(t)$ is strictly real and thus only effect ($i$) can occur. Indeed, in this case, the two-sided OTOC directly measures the operator size and Eq. (9.80) is equivalent to Eq. (9.35). It follows that peaked-size teleportation can be achieved with perfect fidelity: the teleportation correlator magnitudes are equal to one due to the infinite temperature, and their phases can be aligned by tuning $g$ or $t$. More generally, at finite temperature, $\mathcal{F}_2(t)$ contains both a real and imaginary part, and the real part—which leads to effect ($i$)—is formally distinct from the first moment of the size distribution. Rather, recent work has shown that $\mathrm{Re}\{\mathcal{F}_2(t)\}$ is computable via a ladder diagram identity and is physically interpreted as a 'branching time' [181, 528]. Here teleportation is similarly possible by tuning $g$ or $t$ to align the correlator phases, however the teleportation fidelity is bounded from above if the correlators do not have magnitude one (Section 9.3).

At the opposite extreme, effect ($ii$) is dominant in systems with a gravity dual [181, 252] (as well as other maximally chaotic systems, e.g. maximally chaotic 2D CFTs with a large central charge [161]). In such cases, $\mathcal{F}_2(t)$ is mostly imaginary and leads to the growth (or decay) in the magnitude of the correlator. This opens the door to magnitudes *greater* than the two-point function, $|C_\psi(t)| > G_\beta$, which is not possible in peaked-size teleportation (Section 9.5). Interpolating between the two above limits, it has been conjectured that the prefactor of $\mathcal{F}_2(t)$ is proportional to $e^{i\lambda\beta/4\pi}$ [181, 252]. This would imply that the imaginary part is dominant if and only if $\lambda \approx 2\pi\beta$, i.e. the system approaches the bound on chaos [315].

## Gravitational teleportation and the size-winding mechanism

We now move beyond early times and provide a brief review of how the correlator can achieve its maximal magnitude, 1, at finite temperatures. This occurs via the 'size winding' phenomenon introduced in Ref. [71, 350] as the microscopic mechanism for gravitational teleportation [160, 317]. We refer the reader to Ref. [350] for a complete discussion of this

mechanism, including its connection to physical quantities in the bulk gravity theory. As we emphasize in Section 9.3, maximizing the magnitude of the correlators is necessary for high fidelity teleportation, but it is not sufficient: we must also align the correlator phases, for every operator on the subspace to be teleported.

To begin, note that the winding size distribution is normalized to the two-point function, $G_\beta \leq 1$, in contrast to the size distribution, which is normalized to 1. From Eq. (9.76), we see that this norm being less than one implies that the phases of the coefficients $c_\chi$ are not perfectly aligned for different strings $\chi$. It is convenient to separate this misalignment into two classes: first, when coefficients of strings of the same size $\mathcal{S}$ are misaligned, which manifests in the magnitude of $f(\mathcal{S})$ being less than maximal for a given $\mathcal{S}$, and second, when the phases of $f(\mathcal{S})$ for different sizes $\mathcal{S}$ do not align with each other.

We focus on the latter case and, more specifically, consider an ansatz in which the coefficients' phases *wind* with the size [71, 350]:

$$c_\chi = e^{-i\alpha \mathcal{S}[\chi]/q}|c_\chi|, \tag{9.82}$$

In this case, the coupling of the teleportation protocol, by applying a phase that is also proportional to the size, can serve to unwind the phases of $f(\mathcal{S})$ at the value $g/N = -2\alpha$ [see Eq. (9.77)]. This increases the teleportation correlator magnitude from its initial value, $G_\beta$, to unity. Although seemingly artificial, in the following subsection we show that this ansatz holds exactly for the SYK model at low temperatures.

## Large-$q$ SYK model: finite temperature

We now turn to explore the interplay between peaked-size and gravitational teleportation in an explicit example: the large-$q$ SYK model at finite temperature and large-$p$ encoding [379]. Despite the fact that this model features a large separation between the size and size width, we show that teleportation is *not* governed by the peaked-size mechanism at low temperatures, due to the presence of strong size winding.

To begin, let us consider the finite-temperature teleportation correlator, given by [159]:

$$C_\psi(t) = (-iG_\beta)^p \left( \frac{1}{1 - \frac{g}{N}\frac{J}{2\lambda}e^{\lambda t}\sin(\lambda\beta/4) + i\frac{g}{N}\frac{1}{4}e^{\lambda t}} \right)^{2p/q}, \tag{9.83}$$

where $(G_\beta)^p = i^p \langle \psi_r \psi_l \rangle = (\lambda/2J)^{2p/q}$ is the $p$-body two-point function, and the Lyapunov exponent $\lambda$ corresponds to the solution of

$$\beta\lambda = 2\beta J \cos(\lambda\beta/4) \tag{9.84}$$

and interpolates between $2\pi/\beta$ at low temperatures and $2J$ at high temperatures. At infinite temperature, the correlator reduces to Eq. (9.72), and follows our expectations for peaked-size teleportation (see Section 9.7). At low temperatures, where the model is known to possess a gravitational dual [249, 252, 316], the correlator behaves substantially differently;

most notably, its magnitude increases from $G_\beta^p$ at time zero to unity when $gJe^{\lambda t}/2\lambda N = 1$
[illustrated in Fig. 12.1(c)].

From this correlator, we can verify the two predictions made in Sections 9.8 and 9.8:
($i$) the early time behavior is governed by the two-sided OTOC, and ($ii$) the size winding
mechanism is responsible for the $\mathcal{O}(1)$ peak in the correlator magnitude at low temperatures.
To see the former, we expand the correlator in the early time regime:

$$C_\psi(t) \approx (-iG_\beta)^p \exp\left(-\frac{igp}{2qN}\left[i\frac{2J}{\lambda}e^{\lambda t}\sin(\lambda\beta/4) + e^{\lambda t}\right]\right). \tag{9.85}$$

Indeed, the term in the exponent is directly proportional to the connected piece of the
two-sided OTOC [181],

$$\mathcal{F}_2(t) = i\frac{2J}{\lambda}e^{\lambda t}\sin(\lambda\beta/4) + e^{\lambda t}, \tag{9.86}$$

matching Eq. (9.80)[14]. At high temperatures this OTOC is equal to two times the operator
size [Eq. (9.71)], resulting in phase oscillations, whereas at low temperatures the OTOC ro-
tates to become predominantly imaginary, leading to an exponential growth in the correlator
magnitude.

Next, to understand the role of size winding, we must analyze the full winding size
distribution. We can derive this distribution by expanding the teleportation correlator in
powers of $e^{-ig/qN}$ to match Eq. (9.77) [71, 350, 379]. To do so, it is convenient to consider
the exact correlator (before a $g/N \ll 1$ approximation) [159, 379]:

$$C_\psi(t) =$$
$$(-iG_\beta)^p \left(\frac{e^{-ig/2N}}{1 + i(1 - e^{-ig/N})[\frac{J}{2\lambda}\sin(\lambda\beta/4) - \frac{i}{4}]e^{\lambda t}}\right)^{2p/q} \tag{9.87}$$

Rewriting this correlator using Eq. (9.84) and the Taylor expansion,

$$\left(\frac{1}{1 + (1 - e^{-\mu})x}\right)^{2p/q}$$
$$= \frac{1}{(1+x)^{2p/q}}\sum_{n=0}^{\infty} e^{-n\mu}\binom{n + \frac{2p}{q} - 1}{n}\frac{1}{(1 + 1/x)^n}, \tag{9.88}$$

and identifying the $n^{\text{th}}$ coefficient with the winding size distribution, we have:

$$f(qn + p) = -\frac{(-iG_\beta)^p}{(1 + \frac{J}{2\lambda}e^{\lambda t}e^{i\lambda\beta/4})^{2p/q}}$$
$$\times \binom{n + \frac{2p}{q} - 1}{n}\frac{1}{(1 + \frac{2\lambda}{J}e^{-\lambda t}e^{-i\lambda\beta/4})^n}. \tag{9.89}$$

---

[14]More precisely, the correlator in Eq. (9.85) is missing a factor of $G_\beta^p$ compared to Eq. (9.80). This same
mismatch is noted in Ref. [379], and is attributed to the large-$q$ limit utilized for the calculation, since in
this limit $G_\beta$ approaches 1.

At intermediate times and large $p$, the distribution takes a particularly simple form,

$$f(qn + p) \approx (-iG_\beta)^p \frac{(\gamma + i2\alpha)^{2p/q}}{\Gamma(\frac{2p}{q})} n^{\frac{2p}{q}-1} e^{-\gamma n} e^{-i2\alpha n} \tag{9.90}$$

where we define the size decay rate, $\gamma$, as

$$\gamma = \frac{2\lambda}{J} e^{-\lambda t} \cos(\lambda\beta/4) = \left(\frac{\lambda}{J}\right)^2 e^{-\lambda t}, \tag{9.91}$$

using Eq. (9.84), and the size winding coefficient, $\alpha$, as

$$2\alpha = -\frac{2\lambda}{J} e^{-\lambda t} \sin(\lambda\beta/4). \tag{9.92}$$

The above expression holds when $(2p/q)^2 \ll n \ll 1/\gamma^2, 1/\alpha^2$. Crucially, the distribution follows the size winding ansatz, $f(n) = |f(n)|e^{-i2\alpha n}$. Thus, we recognize that the maximum in the correlator magnitude occurs when the coupling has unwound the phases of $f(n)$, at $g/N = -2\alpha$, as expected from Section 9.8 [71, 350].

The fact that the correlator magnitude increases in time, and moreover reaches an $\mathcal{O}(1)$ value at low temperatures, is a hallmark of gravitational teleportation and signals physics outside the peaked-size regime. Naively, this result is surprising, as we expect the $p$-body encoding to ensure a peaked size distribution. Indeed, the average size and size width remain separated by $\sqrt{p}$ at all temperatures [379]:

$$\mathcal{S}[\psi(t)\rho^{1/2}] - \mathcal{S}[\rho^{1/2}] \approx \frac{p}{2} \left(\frac{2J}{\lambda}\right)^2 e^{\lambda t} = \frac{2p}{\gamma}, \tag{9.93}$$

$$\delta\mathcal{S}[\psi(t)\rho^{1/2}] \approx \frac{\sqrt{2qp}}{4} \left(\frac{2J}{\lambda}\right)^2 e^{\lambda t} = \frac{\sqrt{2qp}}{\gamma}. \tag{9.94}$$

This demonstrates that our simple intuition, of judging a size distribution to be tightly peaked if the ratio between the size width and average size is small, is not always correct. Rather, in Appendix D.1, we provide a more precise condition for when peaked-size teleportation holds, and explicitly show that this condition breaks down for the SYK model at finite temperature (but remains satisfied at infinite temperature).

Let us now provide intuition for *how* peaked-size teleportation is modified by size winding at low temperatures. To this end, we express the SYK correlator in terms of the winding size distribution parameters:

$$C_\psi(t) \approx (-iG_\beta)^p \frac{(\gamma + i2\alpha)^{2p/q}}{\Gamma(\frac{2p}{q})}$$

$$\times \int_0^\infty dn\, n^{\frac{2p}{q}-1} \exp(-\gamma n) \exp(-i[g/N + 2\alpha]n). \tag{9.95}$$

$$= (-iG_\beta)^p \left[\frac{\gamma + i2\alpha}{\gamma + i2\alpha + ig/N}\right]^{2p/q}$$
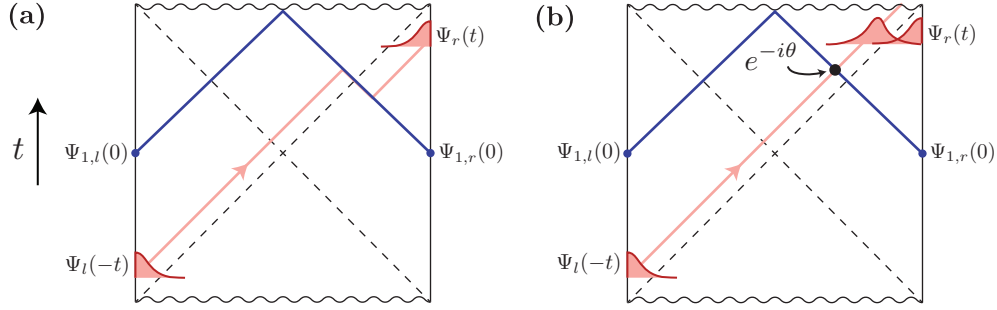
Figure 9.5: Schematic of the teleportation protocol from the bulk gravitational perspective
in $AdS_2$, under both **(a)** semiclassical gravity, and **(b)** strong stringy corrections. The TFD
state corresponds to a two-sided black hole. Local quantum mechanical operators, $\psi_{l/r}$,
create or annihilate particles near the two boundaries, with wavefunctions $\Psi_{l/r}$ (red). The
protocol begins by inserting a particle on the left side, with wavefunction $\Psi_l$ (red, bottom
left), at time $-t$, which then falls towards the interior of the geometry during time-evolution
(red line). The two-sided coupling, $\frac{g}{N} \sum_i \psi_{i,l}\psi_{i,r}$, is then applied, producing a shock wave
(blue) that interacts with the in-falling particle [160, 317]. **(a)** In the semiclassical limit, the
shock wave shifts the position the in-falling particle outside of the right horizon (dashed),
which enables the particle to reemerge near the right boundary (red, top right) [160, 317].
**(b)** When stringy effects are present, the scattering amplitude between the in-falling particle
and the shock wave is modified according to Eq. (9.100) [317, 437]. In the highly stringy
limit and at early times, the interaction results in an overall phase shift, $\theta = gG_N A_\varepsilon (\Delta/2)^\varepsilon e^{\varepsilon t}$
[Eq. (9.102)]. The overlap between the in-falling particle and a particle at the right boundary
is nevertheless non-zero (red, top right), and is given by the unperturbed two-point function,
$G_\beta = i \langle \psi_l \psi_r \rangle$. [Note that stringy effects may also modify the initial wavefunctions of $\Psi_{l/r}$,
as we discuss in the context of Eq. (9.104).]

At early times, this integral can be solved using a saddle-point approximation. At infinite
temperature, the saddle point, $n_s$, occurs precisely at the average size, $n_s = (2p/q)/\gamma = \mathcal{S}/q$,
giving the peaked-size correlator, $C_\psi = (-iG_\beta)^p \cdot \exp(-ig\mathcal{S}/qN)$. In contrast, at finite
temperature, the size winding $\alpha$ shifts the saddle point in the imaginary direction of the
complex plane, giving $n_s = (2p/q)/(\gamma + 2i\alpha)$ and a correlator $C_\psi = (-iG_\beta)^p \cdot \exp(-ign_s/qN)$.
From this, we recognize the saddle point as precisely the two-sided OTOC, $n_s = \frac{p}{2q}\mathcal{F}_2(t)$.

The inclusion of the size winding in the low temperature saddle point thus has two
effects. First, it contributes an imaginary part to the OTOC and thereby increases the
magnitude of the teleportation correlator. More subtly, it also alters the *real* part of the
OTOC. At low temperatures, $\alpha/\gamma \approx \beta J \gg 1$, and we can approximate the saddle as
$n_s \approx (2p/q)/(2i\alpha) + (2p/q\gamma)(\gamma/2\alpha)^2$. Recognizing $\mathcal{S} = 2p/\gamma$, we see that the real part of the
OTOC now corresponds to the average size suppressed by two factors of the ratio $(\alpha/\gamma)^2$.

## Gravity with stringy effects

While the bulk of this paper approaches teleportation firmly through the lens of quantum mechanics, we would be remiss not to explore the analog of peaked-size teleportation in gravitational physics. Specifically, we would like to ask: is there a teleportation mechanism in gravitational systems that shares the same features as peaked-size teleportation? Such a connection might seem surprising, given the prevalence of peaked-size teleportation in quantum mechanical models with no apparent connection to gravity. Nonetheless, the smooth blending between gravitational teleportation and peaked-size teleportation in the SYK model suggests a positive answer.

Here, we demonstrate—in a particular gravitational geometry, $AdS_2$—that an analog of peaked-size teleportation indeed occurs when strong stringy corrections [317, 437] are included in the gravitational theory[15]. Intuitively, our results are consistent with our previous analysis of the SYK model, where, in the dual gravitational theory, increasing the temperature is known to add stringy effects [252].

Our derivation closely follows that of Ref. [317] and assumes a background familiarity with the gravitational description of teleportation in $AdS_2$ (a thorough summary of which can be found in the seminal works of Refs. [160, 317]). In this setting, the teleportation correlator can be calculated explicitly by considering gravitational scattering in a wormhole geometry [Fig. 9.5]. We will maintain our SYK notation, so that $V$ consists of $K$ single-body fermion operators, $\psi_i$, and our input operator is a $p$-body fermion, $\psi$. The correlator can be solved for by decomposing the fermion operators in a momentum basis and applying the scattering matrix:

$$
\begin{aligned}
C_\psi^{\mathrm{sc}}(t) = e^{-ig\langle V \rangle} \int dk \Psi_r(k,t) \Psi_l^*(k,-t) \\
\times \exp\left( ig \int ds e^{i\delta(k,s)} i \Psi_{1,r}(s,0) \Psi_{1,l}^*(s,0) \right)
\end{aligned}
\tag{9.96}
$$

where $\Psi_{l/r}(k,t)$ is the wavefunction for the $p$-body operator inserted on the left/right boundary with in-falling momentum $k$ (and similarly $\Psi_{1,l/r}(s,0)$ for any single-body operator in $V$), and $e^{i\delta(k,s)}$ is the scattering matrix element between $\psi(t)$ and $\psi_1(0)$. In pure gravity, i.e. in the absence of stringy effects, these quantities take the form [317]:

$$
\Psi_r(k,t) \Psi_l^*(k,-t) = \frac{(2ike^{-t})^{2\Delta} e^{-4ike^{-t}}}{i\Gamma(2\Delta)(-k)} \Theta(-k)
\tag{9.97}
$$

$$
\delta(k,s) = G_N k s
\tag{9.98}
$$

where we have set $\beta = 2\pi$ for convenience, $\Theta(x)$ is the Heavyside function, and $\Delta = p/q$ is the conformal weight of $\psi$. The single-body wavefunction, $\Psi_1(s,0)$, is obtained by setting $t = 0$ and replacing $\Delta \to \Delta_1 = 1/q$ (i.e. the conformal weight of a single fermion).

---

[15]We are grateful to Zhenbin Yang and Douglas Stanford for discussions leading to this connection.

In the semiclassical limit, we can evaluate the correlator by expanding $e^{i\delta}$ to linear order in $G_N$ [317]. We find:

$$C_\psi^{\mathrm{sc}}(t) = \langle \psi_l \psi_r \rangle \, \frac{(-i)4^{2\Delta}}{\Gamma(2\Delta)}$$
$$\times \int_0^\infty dk \, (-ik)^{2\Delta-1} \exp\left(-i(\tilde{g}G_N e^t - 4)k\right),$$
(9.99)

where $\tilde{g} \equiv g4^{-\Delta_1}\Delta_1/2$. This expression is almost identical to the large-$q$ SYK correlator of Eq. (9.95), setting the size decay rate to zero, $\gamma = 0$, and identifying the momentum $k$ in the gravitational calculation with the size $n$ in the SYK model [448]. Notably, the correlator diverges at the teleportation time, $4 = \tilde{g}G_N e^t$. In bulk gravity, this divergence is exactly the light-cone pole between the left and right sides of the traversable wormhole, and is regulated by including higher order terms in $G_N$ or stringy corrections [317].

While the full effects of stringy scattering in an AdS background are not known, we will take a phenomenological treatment as in Ref. [317, 437]. Here, the total effect of stringy corrections is to change the scattering amplitude to

$$\delta(k, s) = iG_N(-iks)^\varepsilon, \ 0 \leq \varepsilon \leq 1,$$
(9.100)

where $\varepsilon$ controls the strength of stringy effects, and varies from 1 in pure gravity to 0 in the highly stringy limit.

Again expanding $e^{i\delta}$ to leading order in $G_N$, and Wick rotating $k \to -ik$, we can write the correlator as

$$C_\psi^{\mathrm{stringy}}(t) = \langle \psi_l \psi_r \rangle \, \frac{4^{2\Delta}}{\Gamma(2\Delta)}$$
$$\times \int dk \, k^{2\Delta-1} e^{-4k} \exp\left(-i^{1+\varepsilon} g G_N A_\varepsilon k^\varepsilon e^{\varepsilon t}\right)$$
(9.101)

where $A_\varepsilon$ is a constant of order 1. Note that the $k$-dependence in front of exponential is a Poisson distribution with a saddle point at $k_s \approx \Delta/2$ in the heavy particle limit, $\Delta = p/q \gg 1$. At early times, $e^{\varepsilon t}G_N \ll 1$, and for strong stringy effects, $\varepsilon \to 0$, the change in this saddle point from the scattering, $g$, is negligible. In these limits, the saddle point approximation thus gives the correlator:

$$C_\psi^{\mathrm{stringy}}(t) \approx \langle \psi_l \psi_r \rangle \exp\left(-ig G_N A_\varepsilon (\Delta/2)^\varepsilon e^{\varepsilon t}\right),$$
(9.102)

which has exactly the same form as in peaked-size teleportation [Eq. (9.38)][16]! Specifically, the correlator is equal to the two-point function, $G_\beta = i\langle \psi_l \psi_r \rangle$, multiplied by a pure phase.

---

[16]Note that the phase in Eq. (9.102) becomes order-one within the Lyapunov regime, i.e. $t \lesssim 1/\epsilon \log(1/G_N)$, but at sufficiently early times to satisfy $G_N e^{\varepsilon t} \ll 1$. These conditions are consistent as long as $\Delta = p/q$ is sufficiently large to ensure $A_\varepsilon(\Delta/2)^\varepsilon \gg 1$.

Tentatively, this suggests interpreting the phase as the operator size in a dual boundary theory. This size,

$$\mathcal{S}/N \sim G_N A_\varepsilon (\Delta/2)^\varepsilon e^{\varepsilon t}, \qquad (9.103)$$

grows exponentially in time with a non-maximal Lyapunov exponent, $2\pi\varepsilon/\beta$.

A few remarks are in order. First, while in the above treatment the strength of stringy effects depends on a 'free' parameter $\varepsilon$, we expect that in a UV complete theory $\varepsilon$ would in turn depend on the temperature (and other physical parameters). In particular, we expect $\varepsilon \to 1$ at low temperature in theories that are dual to pure gravity, and $\varepsilon \to 0$ at high temperature, where stringy, UV effects should play an important role. This statement also follows from the point of view of the boundary field theory, since the scattering matrix is proportional to an OTOC of the boundary theory, which is real at infinite temperature.

Second, if we would like to recover the infinite temperature SYK correlator, Eq. (9.72), from the scattering computation, choosing a proper $\varepsilon$ as a function of $\beta$ is not enough. One also needs to modify the *wavefunction* of $\psi$, to:

$$\Psi_r(k,t)\Psi_l^*(k,-t) = \frac{\varepsilon(2ik^\varepsilon e^{-\varepsilon t})^{2\Delta} e^{-4ik^\varepsilon e^{-\varepsilon t}}}{i\Gamma(2\Delta)(-k)}\Theta(-k) \qquad (9.104)$$

Such a wavefunction modification due to UV data should be model dependent, and it would be interesting to understand how to derive this 'stringy-corrected' wavefunction from the bulk point of view. Nevertheless, one particular feature of the modified wavefunction has a clear motivation from the boundary perspective. Specifically, Wick rotating Eq. (9.104), $k \to -ik$, leads to a distribution whose width, $\delta k \sim \Delta^{1/\varepsilon}$, *broadens* as $\varepsilon \to 0$. This broadening increases the phase variations in the exponential of Eq. (9.101) and results in the decay of the correlator at the timescale $e^{\varepsilon t}G_N/\sqrt{\Delta} \approx 1$ for small $\varepsilon$. From the boundary point of view, this decay corresponds to the requirement that the size width must be small, $g\delta\mathcal{S}/N \lesssim 1$, for peaked-size teleportation, as we saw for 0D RUCs and infinite temperature SYK (Section 9.7). We expect this decay to be common to many 0D quantum systems at high temperatures, which suggests that the broadening of the bulk stringy wavefunction as $\varepsilon \to 0$ might also be a general feature.

Finally, the most obvious effect of a non-unity $\varepsilon$ is to change the scattering phase, $\delta(k,s)$, from being real-valued to complex. Indeed, in the strong stringy limit, $\delta(k,s)$ becomes purely imaginary. In general scattering theory, a complex $\delta$ means that the scattering matrix, $e^{i\delta}$, is no longer normalized, and implies the existence of inelastic scattering [437]. Since peaked-size teleportation is replicated in the limit $\varepsilon \to 0$, this suggests a more general relationship between peaked sizes and inelastic scattering. In Appendix D.8, we demonstrate that these two phenomena also coincide at infinite temperature, for arbitrary wavefunctions and scattering amplitudes.

## 9.9 Experimental proposals

Having illustrated the wide breadth of physics that enters into the TW protocol, in this section we outline explicitly how one can probe this physics in the laboratory. We begin with a summary of the key signatures of teleportation, and how they can be applied towards ($i$) characterizing operator size distributions in generic scrambling dynamics, and ($ii$) distinguishing generic vs. gravitational scrambling dynamics. For ($i$), we show that the TW protocol can be simplified dramatically at infinite temperature, where an equivalent 'one-sided' protocol eliminates the need to experimentally prepare the thermofield double state. We next present two near-term experimental realizations of the protocol: first with neutral atoms and second with trapped ions. The fundamental requirement is the ability to time-evolve forwards and backwards under many-body scrambling dynamics; recent experimental progress has demonstrated this in a number of quantum simulation platforms [30, 163, 289, 325, 487]. We conclude with a discussion of the effect of experimental error, and a comparison of the TW protocol with other diagnostics of scrambling physics.

### Signatures of the TW protocol

We begin by reviewing the key signatures of the TW protocol, as discussed in the previous sections and summarized in Table 9.1. We first recall that the simplest experimental signal— that is, any non-trivial teleportation fidelity of a single qubit—has already been demonstrated experimentally in the closely-related HPR protocol [61, 274]. As discussed in Section 9.6, this signifies that the implemented unitary is scrambling but does not distinguish between peaked-size or gravitational teleportation. In what follows, we discuss two more refined applications of the TW protocol.

*Characterizing size distributions in generic scrambling dynamics*—The dynamics of the teleportation fidelity within the TW protocol can be used to probe the size distributions of time-evolved operators. This approach relies on the peaked-size teleportation mechanism and thus applies to generic scrambling systems, including the examples analyzed in Section 9.7 (e.g. RUCs, spin chains, high T SYK).

Specifically, the teleportation fidelity as a function of time exhibits three relevant features. First, since peaked-size teleportation relies on the *width* of the size distribution being small, $g\delta\mathcal{S}/N \lesssim 1$, its success or failure indicates whether the width has surpassed the tunable value, $N/g$. Depending on the model and the value of $g$, this leads to a temporal profile that exhibits three regimes: initial teleportation when the size width is small, no teleportation when $\delta\mathcal{S} \gtrsim N/g$, and late time teleportation once the size width converges to its small final value in a finite-size system [as depicted schematically in Fig. 12.1(c) and observed numerically in 0D RUCs in Fig. 9.2(c)].

Second, within the peaked-size regime, oscillations in the teleportation fidelity as a function of time, $F = \frac{5}{8} - \frac{3}{8}\cos(\eta_d g\mathcal{S}(t)/N)$ [Eq. (9.46)], provide a direct measurement of the growth in operator size. In particular, setting $g = 2\pi n + \pi$, one expects to see $n$ oscillations in the teleportation fidelity before it reaches its late time plateau. The peaks in these os-

cillations give the operator size as a function of time: $\mathcal{S} = (m/n)(1 - 1/d^2)N$ at the $m^{\text{th}}$ peak.

Third, the teleportation of multiple qubits demonstrates the equivalent channel capacities of peaked-size and gravitational teleportation (Section 9.7). Formally, multi-qubit teleportation probes whether the sizes of time-evolved operators *add* under operator composition. While this is trivial when the operators are causally separated, determining the requirements for size addition under more general dynamics—e.g. all-to-all or power-law interactions— remains an open question[17].

*Distinguishing gravitational scrambling dynamics*—The TW protocol can also be used as an experimental litmus test for gravitational dynamics. To this end, we propose to use two experimental signatures that distinguish between gravitational and peaked-size teleportation: (*i*) the teleportation fidelity at low temperature, and (*ii*) the behavior of the teleportation fidelity as a function of time, $t$, and the coupling strength, $g$. For (*i*), the observation of a high teleportation fidelity, $\sim \mathcal{O}(1)$, at low temperatures strongly suggests the occurrence of gravitational teleportation, since the fidelity of peaked-size teleportation is limited at such temperatures by the (small) two-point function, $G_\beta$. For (*ii*), one observes that the qualitative profile of the teleportation fidelity as a function of time differs between the two mechanisms (see Fig. 12.1(c) for a comparison between the two, and Figs. 9.2, 9.3 for additional examples of peaked-size teleportation). Namely, keeping $g$ fixed, the fidelity of gravitational teleportation is expected to display a single peak as a function of time, whereas the fidelity of peaked-size teleportation is highly oscillatory in time. Furthermore, gravitational teleportation works only for a specific *sign* of the coupling, $g > 0$, while the peaked-size teleportation fidelity is an even function of $g$ [71, 160, 317, 350].

*Contrasting with finite-size effects*—Finally, we would like to distinguish many-body teleportation from spurious effects that may be seen in the TW protocol at small-size systems. The most effective way to avoid such signals is by utilizing a coupling $gV$ [Eq. (9.1)] whose individual terms have a small magnitude, i.e. $g/K \ll 1$; this is most naturally achieved by including many couplings—which requires a sufficiently large system—and setting $g \sim \mathcal{O}(1)$. In this limit, the action of the coupling is negligible unless local operators have grown significantly under many-body dynamics, i.e. $\mathcal{S} \sim K/g \gg 1$ (see Section 9.4); any teleportation signal is thus necessarily a result of scrambling dynamics. Furthermore, we expect large-size operators to generically exhibit smooth size distributions, justifying our approximation (Section 9.4) that the teleportation fidelity is governed by the distributions' first few moments.

Away from this limit, our general framework relating the teleportation fidelity to operator size distributions remains valid [e.g. Eq.(9.27)]. However, for $g/K \lesssim 1$, we expect the fidelity to be sensitive to the discrete nature of the size distributions, and our predictions based on the first few moments may no longer apply. Fortunately, as we show in the following subsections, none of these complications are evident for experimentally relevant system sizes (e.g. $K \sim N \sim 20$) and $g \sim \mathcal{O}(1)$ coupling strengths; indeed, our finite-size numerical results

---

[17]Indeed, recent work has indicated that, in theories with a gravitational dual, the lack of size addition is related to a scattering event among infalling particles [188].

agree very well with predictions from the peaked-size teleportation framework [Fig. 9.7(b) and 9.8].

Lastly, in the case where $g/K \sim 1$, operator growth is no longer necessary for the coupling to have a strong effect, leading to the possibility of a teleportation signal unrelated to scrambling. Indeed, for $g/K = \pi$, the coupling effectively 'swaps' the left and right qubits. This is made precise for the coupling $V_s$ [Eq. (9.28)], where $\exp(i\pi N V_s) = (\mathrm{SWAP})Y_l Y_r$. In this case, one would observe perfect teleportation fidelity even without many-body time evolution, i.e. $U = \mathbb{1}$; in fact, if $U$ is perturbed away from the identity via scrambling dynamics, the teleportation fidelity would actually become suppressed. The simplest way to see this is via Fig. 12.1(a)—in particular, any subsequent time-evolution on the right side of the system is in the wrong direction to refocus the time-evolved state (one would want to apply $U^\dagger$ after the coupling, not $U^T$). To achieve a large teleportation fidelity, the combined time-evolution, $U^T U$, would therefore need to preserve the "teleported" state, $\langle \psi | U^T U | \psi \rangle \sim 1$, a situation that is only likely to occur if the dynamics are non-scrambling ($U = \mathbb{1}$ is a special case of this) or undergo a late-time, fine-tuned, Poincare-type recurrence.

## One-sided implementation of teleportation circuit

Before proceeding to the experimental blueprints, we first introduce a simpler implementation of the teleportation protocol that works at infinite temperature (Fig. 9.6). The outcome of this protocol is equivalent to that of the two-sided protocol (up to experimental errors), yet it eliminates the need to prepare EPR pairs and requires half as many degrees of freedom. The cost of this simplification is two-fold: ($i$) it is restricted to simulating an infinite temperature TFD state, and ($ii$) it requires a higher depth quantum circuit.

We derive the one-sided implementation from the 'two-sided' implementation [copied in Fig. 9.6 from Fig. 12.1(a)] by sliding all operations from the left side of the many-body EPR pairs to the right side, using Eq. (9.9). The initial state of the one-sided circuit thus corresponds to the top left of the two-sided implementation. Namely, we initialize the $K$ 'measured' qubits of subsystem C in a definite outcome state, $|o_1 \cdots o_K\rangle$ (purple). These states should be drawn from the distribution of measurement outcomes, but when teleporting an EPR pair at infinite temperature they will be uniformly distributed. For the $N - K$ 'unmeasured' qubits, we use the resolution of the identity $\mathbb{1} \propto \sum_s |s\rangle \langle s|$ to replace the unterminated legs with an initial product state in the computational basis, $|o_{K+1} \cdots o_N\rangle$ (gray). This state should be sampled from shot-to-shot over all $2^{N-K}$ basis states, in effect preparing a maximally mixed state on these qubits. Finally, we include one ancillary qubit for each qubit to be teleported, whose initial state is sampled over a complete basis $|\phi\rangle$ for the teleported subsystem (i.e. subsystem A in Section 9.2). Similar to the unmeasured qubits, this corresponds to the unterminated leg of the thermofield double state when we insert the teleported qubit $|\psi\rangle$ in the two-sided implementation.

Having defined an initial pure state, we now implement the circuit starting from the top left of the two-sided implementation and proceeding counter-clockwise (Fig. 9.6). The circuit consists of three successive applications of $U$ or $U^\dagger$, interspersed with a swap gate exchanging
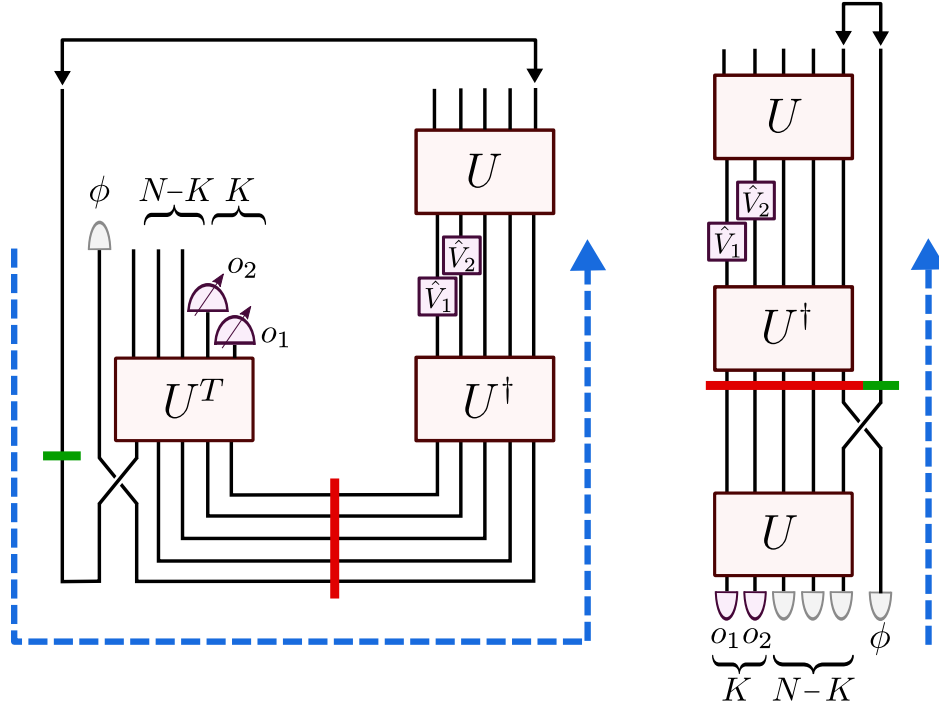
Figure 9.6:   One-sided implementation (right) of the original two-sided teleportation protocol
(left), derived using repeated applications of Eq. (9.9) [replacing $U \to U^T$ for convenience,
compared to Fig. 12.1(a)]. Blue arrows denote the sequence of operations in the one-sided
protocol, the green band marks the teleported qubit and its corresponding component in
the one-sided protocol, and the red band marks the initial EPR state and its corresponding
component.

subsystem A with the ancillary qubit(s), and operations $\hat{V}_i = e^{igo_i\hat{O}_i/K}$ determined by the
initial state of the 'measured' qubits. The outcome of the circuit is an EPR measurement
between the ancilla qubit and subsystem A (black arrows).

As one can see in Fig. 9.6, the one-sided implementation no longer performs teleporta-
tion, but rather prepares an EPR pair from an otherwise scrambled, many-body system.
Specifically, we know that upon swapping out, subsystem A is maximally entangled with the
remaining qubits whenever the unitary, $U$, is scrambling; the one-sided circuit distills this
entanglement into an output EPR pair. This connection has been noted in gravity, where
similar one-sided protocols can be interpreted as distilling the partner operators of emitted
Hawking radiation [515, 516] or observing behind the horizon in the SYK model [266].

## Preparing the thermofield double state

In the previous subsection, we introduced a one-sided protocol that obviates the need to prepare the highly entangled TFD state. However, this approach was restricted to infinite temperature; at *finite temperature*, one must implement the original two-sided protocol, which necessitates preparing a finite temperature TFD state. A number of recent works have explored the preparation of TFD states variationally using quantum approximate optimization algorithms (QAOA) [444, 502, 533]; we note that these preparation strategies require no additional experimental capabilities beyond those already necessary for the TW protocol. The optimization step within a QAOA-based TFD preparation relies on a cost function that requires one to measure the entanglement entropy between the two sides [502, 533]. While challenging, this can in principle be experimentally realized by either using several copies of the system [6, 121, 231] or via randomized measurements [142], both of which have been demonstrated in small-scale trapped ion experiments [77, 298].

## Implementation with neutral Rydberg atoms

One particularly promising platform for implementing the traversable wormhole protocol is a quantum simulator based on neutral alkali or alkaline-earth atoms held in a reconfigurable and controllable array of optical dipole traps. Recent experiments have already achieved near-deterministic trapping and loading of atoms into arbitrary geometries in one, two, and three dimensions [43, 319, 503]. By leveraging the strong dipole coupling between atomic Rydberg states, high-fidelity analog quantum simulations and digital gates have also recently been demonstrated [52, 174, 273, 311, 319, 496]. These demonstrations have primarily used two natural schemes of encoding qubits into neutral atoms:

1. A qubit can be encoded by choosing an atomic ground state $|g\rangle$ to be the $|0\rangle$ state, and a highly excited Rydberg state $|r\rangle$ with principal quantum number $n \gg 1$ as the $|1\rangle$ state [see Fig. 9.7(a)].

2. Alternatively, the qubit states can also be chosen as two long-lived hyperfine ground states (for alkali atoms or fermionic alkaline earth atoms) or a ground state and a metastable clock state (for bosonic alkaline earth atoms), such that the $|1\rangle$ state can be coupled to a Rydberg state to perform entangling gates [see Fig. 9.7(c)].

We will show how both encodings can be used to realize the teleportation protocol in feasible near-term experiments. We find that the first encoding is naturally suited to 'analog' time-evolution under the native (Ising-type) Hamiltonian for a Rydberg setup, but is limited to system sizes of $\lesssim 30 - 35$ qubits (in one spatial dimension) due to the inability to perfectly time-reverse long-range interactions. On the other hand, the second encoding is more flexible and allows for digital time-evolution including RUCs and Floquet dynamics. This time-evolution can be reversed exactly and is limited only by qubit and gate fidelities. While we will primarily consider realizations of our protocol in experimental setups where the neutral
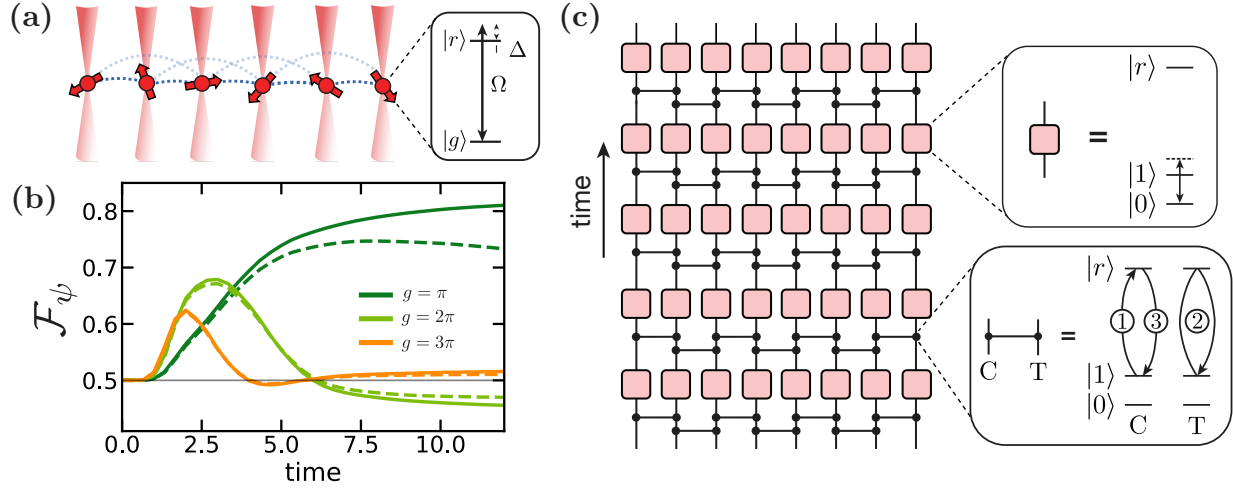
Figure 9.7:   **(a)** In the proposed analog Rydberg teleportation protocol, qubits are encoded
in a ground state $|g\rangle$ and a Rydberg state $|r\rangle$. Nearest-neighbor interactions (dark blue) can
be time-reversed, but next-nearest neighbor interactions (light blue) cannot. **(b)** Numerical
results comparing the average state teleportation fidelity for single-qubit teleportation with
perfectly reversed time-evolution (solid) with the proposed, imperfect time-reversal (dashed).
In particular, we implement the one-sided protocol using $N = 20$ total spins; $K = N - 1$
'measured' spins (i.e. all except the spin encoding $|\psi\rangle$), whose single-qubit rotations are
generated by $\hat{O}_i = \hat{Z}_i$; and time evolution under the analog Rydberg Hamiltonian [Eq.(9.105)]
with parameters $\Omega_i = .9$, $\Delta_i = -1.5$, $J_0 = 1$ (for all $i$). **(c)** Implementation of $U$ or $U^\dagger$ in
the digital protocol, consisting of alternating layers of controlled-phase gates (horizontal
black lines) between nearest neighbor atoms and single-qubit rotations (red boxes). Here,
qubits are encoded in two hyperfine ground states. Insets show possible pulse sequences
to implement the controlled-phase gate and the single-qubit rotations [228]. The full TW
protocol is obtained by inserting this gate sequence (and its Hermitian conjugate) in place
of $U$, $U^\dagger$ in Fig. 9.6.

atoms are individually trapped in optical tweezers and undergo (near-)resonant excitation to Rydberg states, we also conclude by discussing how similar physics can be seen in an optical lattice setup where the atoms are primarily in ground states $|0\rangle$ and $|1\rangle$, but one of these states is 'dressed' by an off-resonant laser field which couples it to a Rydberg state [168, 373, 522].

*Analog implementation*—We first consider the encoding where the qubit states $|0\rangle$ and $|1\rangle$ correspond to a ground state $|g\rangle$ and a highly excited Rydberg state $|r\rangle$. While neutral atoms are effectively non-interacting in their ground states, nearby atoms interact strongly via van der Waals interactions $\propto n^{11}/R^6$ if they are both in the Rydberg state, where $R$ is the distance between the atoms. If one drives the transition $|g_i\rangle \leftrightarrow |r_i\rangle$ at each site $i$ with tunable Rabi frequency $\Omega_i$ and detuning $\Delta_i$ [see Fig. 9.7(b)], the system will undergo analog time evolution under the Hamiltonian

$$H = \sum_i \frac{\Omega_i}{2} X_i + \sum_i \frac{\Delta_i}{2}(1 - Z_i) + \sum_{i \neq j} \frac{J_{ij}}{4}(1 - Z_i)(1 - Z_j) \qquad (9.105)$$

where $X_i = |g_i\rangle\langle r_i| + |r_i\rangle\langle g_i|$, $Z_i = |g_i\rangle\langle g_i| - |r_i\rangle\langle r_i|$, and $J_{ij} = J_0/|i-j|^6$ is the van der Waals interaction strength between two atoms at positions $i$ and $j$.

The Hamiltonian in Eq. (9.105) is scrambling and exhibits a scrambling time limited by the smaller of $J_0$ and $\Omega_i$, $t^* \sim N/\min(J_0, \Omega_i)$. To minimize the total evolution time, we set $|\Omega_i| \sim J_0$, so that evolution under $H$ for a time $\sim N/J_0$ implements a fully scrambling unitary $U$ in the teleportation protocol. To implement $U^\dagger$, we reverse the nearest-neighbor interactions by conjugating time-evolution via Pauli operators $X_i$ (i.e. applying $\pi$-pulses) on every other site. The tunable single-site parameters $\Omega_i$ and $\Delta_i$ are then adjusted to ensure that each single-site term is also reversed. We note that this simple scheme does *not* reverse the (much weaker) next-nearest-neighbor interactions.

In a one-dimensional array, the errors in our implementation will arise from two main sources: $(i)$ the finite lifetime of the Rydberg state, which gives rise to a nonzero decoherence rate at each of the $N$ sites, and $(ii)$ the weak next-nearest neighbor interactions $\sim J_0/2^6 = J_0/64$, which cannot be time-reversed simultaneously with nearest neighbor interactions. To estimate the effect of the former, let us consider the specific case of $^{87}$Rb atoms excited to the $70S$ Rydberg state [52, 273], which has a lifetime $\tau \approx 150$ $\mu$s. Realistically achievable Rabi frequencies and interaction strengths are of order $\sim 2\pi \times 10 - 100$ MHz. The total time to implement the three scrambling unitaries of the teleportation protocol is thus $\sim 3N/|\Omega_i|$; when summed over $N$ qubits and compared to the Rydberg lifetime, this gives an estimated many-body error $\sim 3N^2/|\Omega_i|\tau$.

In order to precisely characterize the effects of imperfect backwards time-evolution, we perform large-scale numerical simulations of the teleportation protocol with the Rydberg Hamiltonian, Eq. (9.105) [1]. Our results are depicted in Fig. 9.7(b) for a one-dimensional chain of $N = 20$ atoms and three values of the coupling $g$. Analogous to our 1D RUC numerics [Fig. 9.2(a)], the fidelity increases monotonically in time for $g = \pi$; while, for $g = 2\pi$ and $g = 3\pi$, the fidelity oscillates in time, reaching a local maximum whenever the
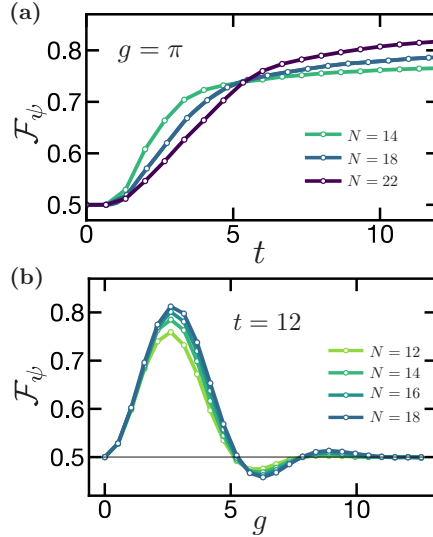
Figure 9.8:   Finite-size scaling of the Rydberg simulations **(a)** as a function of time with
$g = \pi$, and **(b)** as a function of coupling strength $g$ with $t = 12$. The system was evolved
under the Rydberg Hamiltonian, Eq. (9.105), with the same system parameters as in Fig. 9.7.
At late times, the fidelity increases for larger systems but decreases for larger values of $g$.
This is consistent with our error analysis in Section 9.6; in particular, we expect the error
to scale as $g^2 \delta S^2 / N^2$ and the size distribution to approach a binomial distribution for which
$\delta S \sim S / \sqrt{N}$. In contrast, at early times, smaller systems exhibit a larger fidelity not because
of the size width but because the acquired phase is $\eta_d g S(t) / N$, where $\eta_d g$ is fixed and $S(t)$
is initially independent of size. The curves in (a) intersect near the scrambling time due to
the transition between the early and late time regimes.

average size satisfies the phase-matching condition [Eq.(9.36)]. Notably, even with *perfect*
time reversal, the overall fidelity is reduced from unity due to the finite width of the size
distribution. This is a general feature of peaked-size teleportation in finite-size systems, since
the relative size width scales as $\delta S / S \sim 1/\sqrt{N}$ (Section 9.6). Indeed, in Fig. 9.8, we confirm
that the fidelity improves with increasing system size and is consistent with our peaked-size
error analysis [e.g. see Eq. (9.35)].

   With *imperfect* time reversal, we observe an additional $\sim 10\%$ reduction in the fidelity
compared to the ideal case at the scrambling time [Fig. 9.7(b)]. We can estimate the mag-
nitude of this effect by assuming errors due to the next-nearest-neighbor interactions add
coherently over time-intervals $\delta t \sim 1/J_0$ (the local thermalization time), and incoherently at
larger time-scales. Within each $\delta t$, each atom accumulates an error $\sim (\delta t\, J_0/64)^2$; summed
over $N$ atoms and total time $3t^* \approx 3N\delta t$, this gives a total many-body error $\sim 3N^2/64^2$.
Thus, the error due to imperfect time reversal is magnified at larger system sizes and will
eventually outweigh the improvement in fidelity from the narrowing of the size distribution.

Combined with the Rydberg lifetime error, this suggests that near-term experiments should be able to implement peaked-size teleportation in systems of $N \sim 35$ qubits. We note that in higher dimensions, the smaller relative distance of next-nearest neighbor atoms gives rise to a larger error contribution from imperfect time-reversal.

*Digital implementation*—To implement the protocol in larger systems, higher dimensions, and at finite temperature, we propose a digital scheme, using the second type of qubit encoding (i.e. hyperfine ground states) [Fig. 9.7(c)]. In this approach, we envision time-evolution to be formed from alternating layers of nearest-neighbor controlled-phase gates and single-qubit rotations. Here, the controlled-phase gates can be implemented by applying a simple pulse sequence to excite and de-excite qubits from the $|1\rangle$ state to the $|r\rangle$ state, so that the wavefunction acquires a phase of $-1$ if either of the two qubits are in the $|1\rangle$ state, but not if both qubits are in the $|0\rangle$ state [see Fig. 9.7(c) insets] [228]. As demonstrated in recent experiments [288], these Rydberg-mediated controlled-phase gates can be performed in parallel for sufficiently well-separated pairs of qubits, and non-nearest neighbor interactions can be avoided by slightly reducing the parallelism within each layer of controlled-phase gates. Single-qubit rotations can be performed with sufficiently high fidelity such that the overall circuit fidelity is primarily limited by the entangling gates [287, 503].

For a generic choice of gates, the circuit will be fully scrambling when $U$ is composed of $\sim N$ layers of controlled-phase gates. The fidelity of the overall implementation is limited by the finite lifetime of the Rydberg state, which is populated for time $\sim 1/J_0$ during each controlled-phase gate. Assuming the same experimental parameters as in the analog case, one expects to be able to perform approximately $\Omega\tau \sim 10^3 - 10^4$ controlled-phase gates within the decoherence time-scale. Thus, in the digital approach, one expects that the teleportation protocol can naturally be implemented for $N \sim 200$ qubits.

The digital approach can also be adapted to experiments using Rydberg-dressed neutral atoms in an optical lattice [168, 373, 522]. In such a setup, qubits are again encoded in hyperfine ground states and strong Ising-like interactions are generated by coupling the qubit state $|1\rangle$ to a Rydberg state with a far-detuned laser field. In this way, the Rydberg interaction gives rise to an energy shift for two neighboring atoms both in the $|1\rangle$ state. Analogous to our previous discussion, a simple scrambling unitary could consist of alternating layers of Rydberg-dressed interactions and single-qubit rotations. While the total accumulated error in the Rydberg-dressing approach is comparable to the gate-based protocol, one potential advantage is an increased tunability of the interactions [57, 285].

In addition to scrambling time evolution, there are three ingredients to implement the one-sided teleportation circuit (Fig. 9.6): ($i$) the ability to 'swap' in the qubit $|\phi\rangle$, ($ii$) single-qubit rotations, $V_i = e^{\pm igZ_i/K}$, and ($iii$) the final measurement in the EPR basis. In both digital setups, these are easily accomplished by combining controlled-phase gates, arbitrary single-qubit rotations, and local measurements. In the analog setup, we propose to temporarily 'turn off' the Hamiltonian by transferring each Rydberg state $|r\rangle$ to a hyperfine ground state (e.g. the state used as $|1\rangle$ in the *digital* protocol) using a resonant laser pulse. Once this is done, all of the above operations can be performed identically as in the digital setup. Afterwards, an additional resonant laser pulse returns the system to the analog
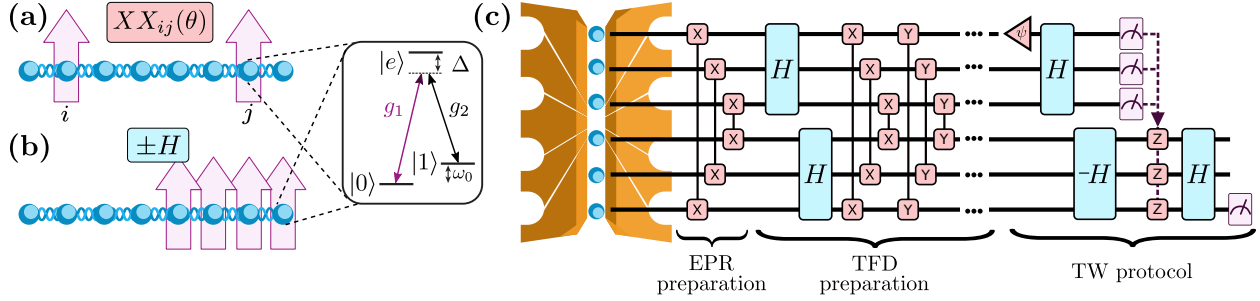
Figure 9.9:   **(a-b)** Chain of atomic ions, with qubit states $|0\rangle$, $|1\rangle$ represented by hyperfine
ground states. The states are coupled by a pair of laser beams, one with individual addressing
(with strength $g_1$, purple) and one applied globally (with strength $g_2$). Each beam is strongly
detuned from an excited state $|e\rangle$ by an amount $\Delta$. The coherent beatnote between the
beams, at frequency $\omega_0$, drives stimulated Raman transitions between the qubit levels with
an effective Rabi frequency $g_1 g_2/2\Delta$, and also modulates the Coulomb interaction between
qubits to give rise to an effective Ising interaction. **(a)** A two-qubit entangling gate, $XX_{ij}(\theta)$,
(red) is performed by addressing only ions $i$ and $j$ with the first beam. **(b)** Half of the qubits
are addressed, which leads to analog time-evolution under the Hamiltonian Eq. (9.106) (blue)
for all addressed spins. **(c)** Quantum circuit implementation of the teleportation protocol
at finite temperature. EPR pairs are formed using two-qubit gates. The TFD state is then
prepared via a QAOA approach by iterating multiple times between two-qubit gates coupling
the sides and analog time-evolution on both sides individually [502, 533]. The state $|\psi\rangle$ is
inserted either by projectively measuring the designated qubit and preparing the state, or by
digitally swapping in an additional qubit (not shown). Finally, teleportation is implemented
using similar ingredients as well as feed-forward measurements (purple dotted lines).


encoding. The ancillary qubit can be decoupled from the system qubits during Hamiltonian
time-evolution in two ways: $(i)$ by physically positioning the ancillary qubit far from the
system, or $(ii)$ by encoding the ancillary qubit in the hyperfine subspace throughout time-
evolution.

The two-sided, finite temperature TW protocol can be achieved by combining the above
techniques with TFD preparation as in Section 9.9. A particularly natural geometry for
such a realization would be two parallel chains of Rydberg atoms, with each chain forming
one side of the TFD state. The coupling between the two sides is naturally realized by the
atoms' Ising interactions. This coupling can be applied independently from the one-sided
Hamiltonian using either full digital control or by manipulating the inter- vs. intra-chain
atomic distance.

## Implementation with trapped ions

A second experimental platform that naturally enables the implementation of the TW protocol is arrays of individual trapped atomic ions [63, 473, 526]. Trapped ion qubits feature near-perfect replicability, negligible idle errors, and the ability to implement both a universal set of reconfigurable quantum gates [91] as well as analog long-range spin Hamiltonians [59, 333]. Entangling quantum gates have been demonstrated between isolated pairs of trapped ions with fidelities exceeding 99.9% [35, 158]. Teleportation protocols—including the HPR protocol [274]—involving gate operations, partial measurement and feedforward operations, have been experimentally realized in a number of contexts [44, 274, 357, 393].

Compared to Rydberg atom arrays, trapped ions offer two new regimes for exploring many-body teleportation. First, trapped ions naturally interact via a long-range analog Hamiltonian, whose time-evolution can be fully reversed within certain experimental regimes [261, 458]. Implementing the TW protocol in this setting would provide a window into operator spreading and size distributions under such long-range dynamics [144, 531]. Second, when operated digitally, the same long-range interaction has already been demonstrated to enable the preparation of thermofield double states [322, 444, 502, 533], a crucial step towards realizing the two-sided TW protocol at finite temperature (see Section 9.9).

We begin by outlining the analog and digital forms of time-evolution that are possible in trapped ion systems. Interactions between qubits typically stem from state-dependent optical dipole forces that off-resonantly drive motional sidebands of the qubit [104, 339]. These sideband operations mediate entanglement and give rise to an effective Ising coupling. When the optical forces are symmetrically detuned far from the upper and lower sidebands, the motion is only virtually excited, resulting in a long-range Ising Hamiltonian [Fig. 9.9(b)]:

$$H = \sum_{i<j} J_{ij} X_i X_j + B_z \sum_i Z_i, \tag{9.106}$$

where $J_{ij} \approx J_0/|i-j|^\alpha$, with $0 < \alpha < 3$ and $J_0 \lesssim 1$ kHz, and the effective magnetic field $B_z$ can be realized by slightly asymmetrically detuning the driving field [332]. The sign of the couplings can be reversed by changing the detuning of the optical forces from the motional sidebands [261, 458].

On the other hand, when the optical dipole forces are closer to resonances of the motional modes, one can mediate interactions significantly faster, allowing for the execution of rapid, entangling quantum gates between pairs of illuminated ion qubits [Fig. 9.9(a)] [126, 536]. The native entangling gates are based upon Ising interactions between any selected pair of ions with a tunable interaction angle; in particular, both $XX_{ij}(\theta) = e^{-i\theta X_i X_j/2}$ and $YY_{ij}(\theta) = e^{-i\theta Y_i Y_j/2}$ gates are available and $\theta = \pi/2$ naturally creates an EPR pair [274, 501]. Typical entangling operations have duration $1/J_{\text{ent}} \sim 100$ $\mu$s, while decoherence time-scales are on the order of $\tau \sim 400$ ms [127]. Following the estimates of Section 9.9 and requiring $3N^2/J_{\text{ent}}\tau \lesssim 1$, we estimate that near-term state-of-the-art experiments can support high-fidelity many-body teleportation for up to $N \sim 35$ qubits.

Let us now describe an implementation of the one-sided TW protocol (Fig. 9.6). We first focus on the ability to implement both $U$ and its inverse $U^\dagger$. For analog time-evolution [Eq. (9.106)], $U^\dagger$ can be implemented by changing the sign of the detuning [163], while for digital time-evolution, one can directly invert and reverse the ordering of the quantum gates.

The one-sided protocol also requires the ability to locally address a sub-extensive number of individual qubits. In particular, a subset $K$ of the qubits must be initially prepared in a product state, $|o_1, \ldots, o_K\rangle$ and later rotated by $\hat{V}_i = e^{igo_i \hat{O}_i / K}$. These rotations can be achieved by taking $\hat{O}_i = \hat{Z}_i$ and individually addressing the target ions using an auxiliary "poke" laser beam [440, 526].

Following the first application of $U$, one must swap out the qubit(s) corresponding to the teleported subsystem. This swap can be implemented either digitally by applying a SWAP-gate, or physically, by exchanging the two ions via a modulation of the ion trap's axial fields [209, 241, 333].

Extending this implementation to the two-sided protocol [Fig. 12.1(a)] is straightforward. Initialization into EPR pairs (for infinite temperature) can be accomplished via simple Ising gates at the input of the circuit [Fig. 9.9(a,c)], while the TFD state (for finite temperature) can be prepared via variational methods (Section 9.9). Time-evolution can again take the form of either digital quantum gates [Fig. 9.9(a)] or analog Hamiltonian dynamics. To separately implement analog dynamics on the two sides of the system, one would illuminate only *half* of the ion chain at any given time [Fig. 9.9(b)]; this has the added benefit of avoiding unwanted coupling between the left and right sides, but implies that the time-evolution must be performed serially [Fig. 9.9(c)].

Finally, in the two-sided protocol, one must perform projective measurements on $K$ qubits that feed-forward to the conditional rotations, $\hat{V}_i$. These partial measurements can be accomplished by using multiple ion species (i.e. different elements or isotopes) [44], or alternatively, this entire procedure can be replaced with a specific interaction, $e^{igV}$, between the two sides; this interaction is naturally realized via an $XX_{ij}(\theta)$ gate with $\theta = 2g/K$.

## Effects of experimental error and relation to quantum error correction

We now turn to the effect of experimental error on the TW protocol. We find that teleportation is robust to nearly all errors that occur on the left side of the TFD state after time-evolution by $U$, but is strongly sensitive to errors at nearly all other locations in the protocol. These two extremes are emblematic of two different relations between scrambling and error: the former corresponds to interpretations of scrambling as an error-correcting code [202], while the latter reflects recent results showing that the effect of errors on scrambling measurements is enhanced proportional to the size, $\mathcal{S}$, of time-evolved operators [427]. In the following discussion, we demonstrate each of these points through simple but representative examples of experimental error.

We begin with the first case: consider errors occurring on the left side of the TFD state

after application of $U$ but before measurement/coupling. Recall that, in the absence of error, one can perform teleportation by using any $K \sim \mathcal{O}(1)$ qubits of the left side. This implies that teleportation is robust to *any* errors that affect only $N - K$ qubits: as long as one has knowledge of at least $K$ qubits that are unaffected, measuring these qubits performs teleportation identically to the error-free case.

This robustness reflects previously noted connections between scrambling and quantum error correction [202]. In particular, we note that many-body teleportation can be understood as an especially generic example of entanglement-assisted quantum error correction (EAQEC) [74]. Indeed, the setup for EAQEC is identical to that of the teleportation protocol: two parties, Alice and Bob, share entanglement (the TFD state), Alice applies an encoding circuit to her share of qubits (the left unitary, $U$), and decoding is achieved by teleporting Alice's quantum state to Bob's share of qubits (via the coupling, $V$, and unitaries on the right). Previous schemes for EAQEC have focused primarily on encodings via Clifford unitaries. In contrast, many-body teleportation, and more specifically peaked-size teleportation, succeeds for a *vastly* broader class of encoding procedures—i.e. scrambling many-body time dynamics—indicating that naturally occurring, strongly interacting systems offer novel methods of EAQEC.

On the other hand, errors that occur *during* encoding or decoding—i.e. during the application of $U$ on the left side or at any point on the right side—strongly inhibit teleportation. As a first example, consider a single local error, $W_1$, occurring with probability $\varepsilon$ on the right side after coupling but before $U^T$ (i.e. just before decoding). If the error, $W_1$, grows to have a size, $\mathcal{S}$, after $U^T$ is applied, one estimates that it will decrease the teleportation fidelity by an amount, $1 - F \sim \varepsilon\mathcal{S}/N$, proportional to the probability that $W_1$ has support on the teleported qubit after time-evolution. If we sum over such errors on all $N$ qubits, we have $1 - F \sim \varepsilon\mathcal{S}$.

As a second example, consider a local error, $W_2$, occurring with probability $\varepsilon$ on the left side simultaneously with state insertion (e.g. a damaged TFD state in Fig. 12.1). In effect, this error shifts the correlator operators [Eq. (9.2)], $Q \rightarrow Q \otimes W_2$; following the arguments of Section 9.7, one then requires that the sizes add for teleportation to succeed, $\mathcal{S}[QW_2] = \mathcal{S}[Q] + \mathcal{S}[W_2]$. In a 1D short-range system (Section 9.7), this condition holds if and only if the light cones of $W_2$ and $Q$ do not overlap. For $\mathcal{O}(\varepsilon N)$ randomly distributed errors, we expect this to hold as long as the spacing between errors, $1/\varepsilon$, is much larger than the size of the light cone, $\varepsilon\mathcal{S} \ll 1$. A similar scaling holds in 0D (Section 9.7). Here, we expect size addition to hold as long as the size of the total error (corresponding to a time-evolved product of $\sim \varepsilon N$ initially local operators), is much smaller than the system size, $N$. Once again, this requires $\varepsilon N\mathcal{S} \ll N$, or $\varepsilon\mathcal{S} \ll 1$.

The two previous examples are straightforwardly generalized to errors that accumulate continuously throughout time-evolution. To do so, we replace the error probability with an error rate, $\varepsilon$ (now with units of inverse time). The total effect of the error is then given by the integral of the error rate multiplied by the size over time, $\varepsilon \int_0^t dt' \, \mathcal{S}(t')$ [427]. In one-dimensional systems evolved up to the scrambling time, i.e. $\mathcal{S} \sim Jt$ and $t_s \sim N/J$ for a local interaction strength $J$, we thus estimate a total error, $\varepsilon \int_0^{t_s} dt' \, Jt' \sim \varepsilon\mathcal{S}t_s \sim \varepsilon N^2/J$, in

agreement with our rough estimates in Sections 9.9 and 9.9.

Finally, we consider a particular form of error that may be relevant for analog time-evolution: mismatches between the evolution times of $U$, $U^*$, and $U^T$. We denote these three evolution times as $t_1, t_2, t_3$, respectively, and their mismatches as $\Delta t_{12} = t_2 - t_1$ and $\Delta t_{13} = t_3 - t_1$. We can characterize the mismatches' effect on the teleportation fidelity using the correlators, $C_Q(t_1, t_2, t_3) = \langle U_r^*(t_3) Q_r U_r^T(t_3) e^{igV} U_l(t_1) Q_l U_l^\dagger(t_2) \rangle$ (Section 9.3). From this, we anticipate that the protocol is relatively insensitive to mismatches between $t_3$ and $t_1, t_2$: teleportation succeeds as long as the mismatch is small compared to the local interaction strength, $J$, i.e. $J\Delta t_{13} \lesssim 1$. To estimate this, we set $g = 0$ and $t_1 = t_2$ in which case the correlator magnitude is given by an autocorrelation function, $C_Q = \langle Q(t_1) Q(t_3) \rangle = G(\Delta t_{13})$. The teleportation fidelity is bounded above by this expression, which decays on a time-scale $\sim 1/J$. On the other hand, teleportation is more strongly sensitive to the mismatch between $t_1$ and $t_2$. To estimate this, we treat the difference in time-evolution between $U$ and $U^*$ as a product of $\sim (J\Delta t_{12})^2 N$ local errors occurring simultaneous with state insertion (to motivate this scaling, note that one can approximate $U(\Delta t_{12})$ as a product of $\sim N$ local unitaries for small $\Delta t_{12}$, and we expect the error to be an even function of $\Delta t_{12}$). Following our previous analysis, teleportation is successful as long as $\mathcal{S}(J\Delta t_{12})^2 N \ll N$, or $\mathcal{S}(J\Delta t_{12})^2 \ll 1$.

## Directly measuring the size distribution

In Section 9.9, we discussed that the time profile of the teleportation fidelity reveals important features of the operators' size distributions, including the average operator size and the size width. We now demonstrate that a more precise characterization of the operator size distribution can be obtained by sweeping the coupling strength, $g$, at a fixed time, $t$.

For simplicity, we restrict to infinite temperature[18] and the coupling $V_s$ in Eq. (9.28), which precisely measures the operator size. In this case, the two-sided correlator [Eq. (9.2)] is equal to the characteristic function, $\Phi_{\mathcal{S}}(g)$, of the size:

$$C_Q(t) = e^{ig} \sum_{\mathcal{S}} P(\mathcal{S}) e^{-ig\mathcal{S}/N} \equiv e^{ig} \Phi_{\mathcal{S}}(g) \tag{9.107}$$

from which the size distribution can be obtained by a Fourier transform in $g$.

More precisely, to measure the *real part* of the characteristic function (i.e. the teleportation correlator), we perform the teleportation protocol with two small modifications: $(i)$ we replace state insertion with the specific projection operator, $(1 + Q)/2$, and $(ii)$ we measure the expectation value of $Q$ applied to the right side, instead of the teleportation fidelity.

---

[18]At finite temperature, a similar procedure to what follows determines the winding size distribution discussed in Section 9.8 [350]. The size distribution can be determined by moving the final measurement of the TW protocol to the left side.

This yields the quantity:

$$
\begin{aligned}
\langle \text{EPR}| &\frac{1+Q_l(t)}{2} e^{-igV_s} Q_r(-t) e^{igV_s} \frac{1+Q_l(t)}{2} |\text{EPR}\rangle \\
&= \text{Re}\big[\, \langle \text{EPR}| \, e^{-igV_s} Q_r(-t) e^{igV_s} Q_l(t) |\text{EPR}\rangle \,\big] \\
&= \text{Re}\big[\varphi_{\mathcal{S}}(g)\big],
\end{aligned}
\tag{9.108}
$$

where in the second line we use that the "diagonal" terms between the two copies of $(1+Q)/2$ vanish at infinite temperature. The imaginary part of the characteristic function can be obtained similarly, by replacing state insertion, $(1+Q)/2$, with an application of the unitary operator, $(1+iQ)/\sqrt{2}$. Analogous to Fig. 9.6, both of these measurement schemes can be adapted into one-sided protocols using Eq. (9.9) whenever the coupling $V$ is classical (i.e. composed of terms $O_{i,l}O_{i,r}^*$, where $\{O_i\}$ mutually commute). While such couplings do not measure the exact size distribution, we expect their behavior to be similar in most cases (Section 9.4).

For completeness, we also note an alternate method to measure the size distribution: one prepares the state $Q_l(t)|\text{EPR}\rangle$ and directly measures the two-sided coupling $V_s$. The probability distribution of the measurement results gives the size distribution [see the discussion below Eq. (9.28)].

Let us now compare these two protocols to other schemes for characterizing the size distribution of operators. First, we recall that a sum of local OTOCs yields the average operator size [Eq. (9.32)]. Hence, many existing protocols for measuring local OTOCs [451, 512] can be straightforwardly adapted to measuring the average size. Higher order moments of the size distribution can similarly be obtained from local OTOCs, using Eq. (9.28):

$$
\begin{aligned}
\langle (1-\mathcal{S}/N)^n \rangle &= \langle V_s^n \rangle_Q \\
&= \frac{1}{N^n} \sum_{P_{i_1},\ldots,P_{i_n}} \text{tr}\big(Q(t) \prod_{k=1}^{n} P_{i_k} Q^\dagger(t) \prod_{k=n}^{1} P_{i_k}^\dagger\big).
\end{aligned}
\tag{9.109}
$$

where the sum is over every possible combination of $n$ single-qubit Pauli operators $P_{i_1},\ldots,P_{i_n}$. Based on this approach, however, the number of measurements required to determine the $n^{\text{th}}$ moment scales as $\mathcal{O}(N^n)$. In certain situations, this scaling may be reduced through sampling, though this depends on the nature of the size distribution and the desired degree of precision. Furthermore, reconstructing the full profile of the size distribution from a finite number of moments is generally a difficult numerical task [230]. In contrast to these limitations, our proposal directly yields the full size distribution, and can recover its moments with a number of measurements independent of the system size[19].

---

[19]This is simplest to see in the protocol which measures the two-sided coupling $V_s$. Here, the error in one's measurement of the $n^{\text{th}}$ moment is equal to the expectation value of the moment's variance divided by the number of measurements, $(\delta\overline{\langle\mathcal{S}^n\rangle})^2 = (\langle\mathcal{S}^{2n}\rangle - \langle\mathcal{S}^n\rangle^2)/M$. If one wishes to resolve the moment to within a relative error $\varepsilon$, i.e. $\delta\overline{\langle\mathcal{S}^n\rangle} < \varepsilon\overline{\langle\mathcal{S}^n\rangle}$, one requires $M \sim \frac{\langle\mathcal{S}^{2n}\rangle - \langle\mathcal{S}^n\rangle^2}{\varepsilon^2\langle\mathcal{S}^n\rangle^2}$ measurements. This number does not scale with $N$ since it contains the same powers of $\mathcal{S}$ in the numerator and denominator.

We can also compare our proposal to an independent protocol for measuring the size distribution introduced in Ref. [381]. The protocol of Ref. [381] is experimentally simpler than our own, and in particular involves only a single application of time-evolution by $U$ (and no backwards time-evolution). However, this simplicity comes at a cost: resolving high-size components of the distribution requires a number of measurements that scales *exponentially* with size.

## 9.10 Discussion and outlook

This Chapter develops a unified framework for understanding many-body teleportation from the perspective of operator growth under scrambling dynamics. The unifying concept within this framework is the size distribution of time-evolved operators [71, 350, 379, 381, 397]: these form the backbone of peaked-size teleportation, and provide a more fine-grained measure of operator growth compared to the average operator size (as given by the expectation value of OTOCs).

Our work suggests several future directions for applying and building upon this framework. First, while we have studied the size distributions in 0D and $\geq$ 1D RUCs, it would be interesting to extend this analysis to a multitude of other physical systems, where one expects to find qualitatively distinct behavior. These include long-range interacting systems [144, 464], interacting and non-interacting integrable systems [381], $\geq$ 1D systems with a large on-site Hilbert space [182], 0D systems with sparse couplings [49], and systems with conserved quantities [246].

Another set of open questions concerns the notion of operator size at finite temperature. In systems with peaked size distributions, we found that the phase of the two-sided teleportation correlator was directly proportional to the conventional definition of operator size [379]. Surprisingly, we observed that this relationship did not hold in the finite temperature SYK model; rather, the phase was given by the real part of the two-sided OTOC. Unlike the conventional size, this OTOC is not UV divergent, and is thus expected to be inherently independent of the microscopic Hilbert space. Recent work has shown that its real part isolates an incoherent component of operator spreading in large-$N$ models [181]; further work is needed to establish and expand this framework. Related to these considerations, one may hope to better understand the bulk analogue of operator size in theories dual to gravity with strong stringy effects. While we have seen that stringy effects can mimic peaked-size teleportation, developing a physical interpretation of this correspondence would be extremely exciting.

Third, we have shown that a promising application of the teleportation protocol is to distinguish between different classes of scrambling dynamics. In particular, we have focused on two classes of scramblers—generic thermalizing systems and those with gravitational duals—and demonstrated that the key distinction between them is their teleportation fidelity at low temperatures. It is intriguing to ask whether the fidelity increase associated with gravitational teleportation may also occur in other systems, without a gravitational dual. For

instance, recently the teleportation correlator magnitude was observed to increase slightly above $G_\beta$ in non-local random Hamiltonian systems [71, 350]; generalizing this to other physical models would be of tremendous interest.

One may also wonder what role an extensive low temperature entropy—a key feature of the SYK model [316]—plays in the teleportation process. In particular, how well can systems with extensive low temperature entropy but no known gravitational dual teleport [18, 415]? We conjecture that an extensive entropy would allow one to *locally* encode each qubit into low-energy degrees of freedom (i.e. operators with an $\mathcal{O}(1)$ two-point function), since one would only require $\mathcal{O}(1)$ qubits on the left side of the TFD in order to have one qubit of mutual information with the right side. Such an encoding would allow low temperature teleportation with perfect fidelity if operator sizes were peaked, naturally motivating the study of operator size distributions in such models.

Finally, we would like to discuss the relation between our results on the TW protocol and the eternal traversable wormhole (ETW) introduced in Ref. [314]. In the latter, the coupling, $V$, has an $\mathcal{O}(1)$ coefficient and, moreover, is applied *simultaneously* with single-sided Hamiltonian evolution (i.e. the full system evolves under a Hamiltonian, $H_l + H_r + g \sum_j O_{j,l} O_{j,r}^*$). Under these conditions, Refs. [314, 371] find that the ETW teleportation fidelity oscillates in time under gravitational dynamics, indicating that information is transmitted back and forth between the two boundaries. Intriguingly, unlike the TW protocol, the ETW oscillations occur at a time-scale given by the single-sided thermalization time ($\sim \beta$, the inverse effective temperature), and *not* the scrambling time. Developing a microscopic understanding of the ETW in terms of operator spreading, as well as exploring analogous physics in more generic many-body systems, remains an exciting open direction.

# Chapter 10

# Operator growth in open quantum systems

Continuing the theme of Chs. 7 and 8, in this Chapter we provide a more complete picture of the behavior of quantum information scrambling [293, 365, 398] under noisy quantum dynamics. Interestingly, our results apply not just to scrambling experiments [61, 77, 111, 138, 139, 163, 224, 233, 274, 289, 326, 416, 418], but also to a broader class of experiments that involve time-reversal operations, such as the Loschmidt echo [173, 303]. In this sense, we can view our results as an application of the framework of quantum information dynamics to questions that long pre-date their modern study.

As we've briefly explored in Chs. 7 and 8, the interplay between scrambling dynamics, extrinsic decoherence, and experimental noise motivates an essential question: What is the nature of quantum information scrambling in *open* quantum systems [8, 21, 28, 30, 37, 55, 197, 274, 326, 450, 463, 473, 518, 529]? Despite the importance of this question, previous studies have been largely empirical and a unified cohesive understanding is lacking.

In this Chapter, we introduce a universal framework—based upon *operator size distributions* [340, 379, 397, 429]—for capturing the effect of local errors on scrambling dynamics. In particular, we conjecture that the propagation of errors in chaotic many-body systems is fundamentally controlled by the size distributions of time-evolved operators and independent of the microscopic error mechanism. Our framework immediately offers predictions for both the Loschmidt echo and out-of-time-ordered correlation (OTOC) functions [277, 506]. In particular, we predict that the decay of the Loschmidt echo, which measures the fidelity associated with backwards time-evolution, occurs at a rate proportional to the operator size. Meanwhile, we predict that the decay of the OTOC, which measures the growth of local operators, is inhibited by open-system dynamics (by an amount proportional to the width of the operator size distribution).

We leverage our framework to characterize operator growth in five distinct classes of open quantum systems, which vary in their dimensionality, range of interaction, conservation laws, and integrability (Table 10.1, Figs. 12.1–10.3). In each class, our framework yields markedly distinct predictions for the Loschmidt echo and OTOCs. We hypothesize that these results
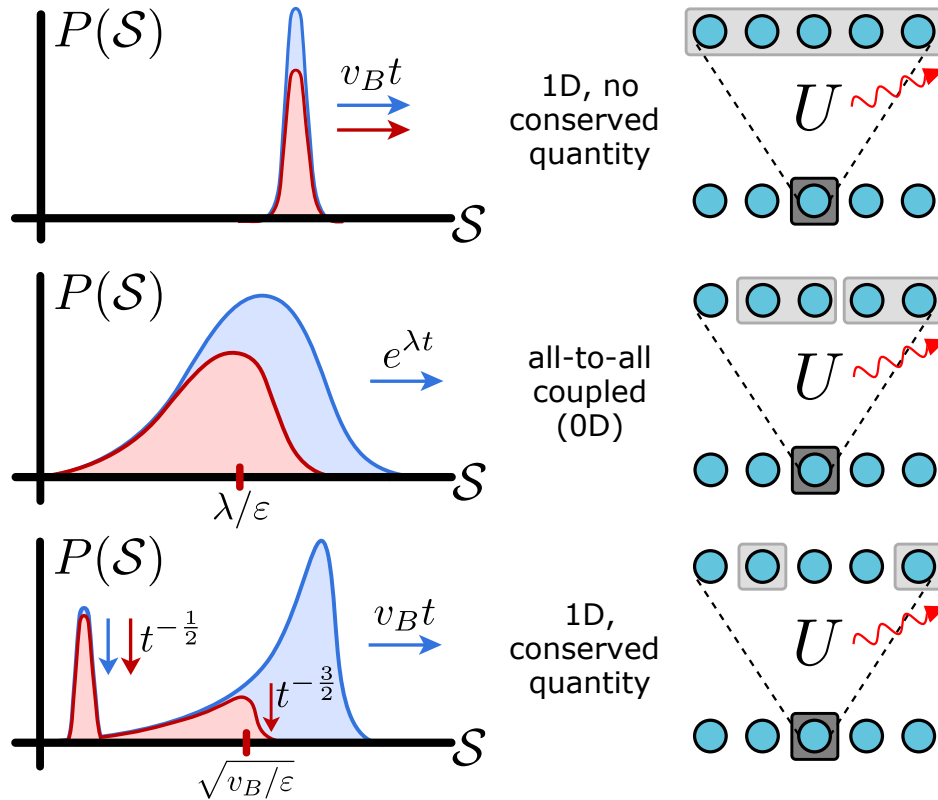
Figure 10.1: Left: Size distributions for three classes of systems under unitary (blue) versus open (red) dynamics. Rightward arrows denote growth in time to larger sizes, ticks denote a fixed size, and downward arrows denote loss of probability at a given size. Right: Qualitative depiction of open-system operator growth. In all cases, operators lose normalization due to open dynamics (dark to light gray boxes). In the latter two classes, operators are dominated by smaller size components compared to unitary evolution (smaller boxes).

provide a theoretical underpinning for recent nuclear magnetic resonance (NMR) experiments [139, 416, 418], and also serve to resolve apparent disagreements between previous empirical studies of open-system scrambling [326, 473, 529]. Finally, we propose and analyze a protocol for measuring operator size distributions via engineered dissipation.

## 10.1 Operator size distributions

We begin with a simple example to build intuition. Consider a lattice of qubits acted on by a series of local quantum gates, each featuring some error $\varepsilon$, before measuring a local operator $\hat{M}$. Noting that the measurement can only be influenced by gates in its past light cone, a naive estimate of the measurement fidelity is $\mathcal{F} \approx (1 - \varepsilon)^{\mathcal{V}_{\text{LC}}}$, where $\mathcal{V}_{\text{LC}}$ is the light cone

Table 10.1: Size distributions in various physical regimes

| System | Unitary dynamics | Open dynamics |
|---|---|---|
| $(d \geq 1)$D, no conservation law | peaked-size $\overline{\mathcal{S}} \sim t^d$ | size unaffected $\overline{\mathcal{S}} \sim t^d$ |
| $(d \geq 1)$D, conserved quantity | bimodal $\overline{\mathcal{S}} \sim t^d$ | size decreases $\overline{\mathcal{S}} \sim 1/t^d$ |
| all-to-all coupled (0D) | broad size $\overline{\mathcal{S}} \sim e^{\lambda t}$ | size plateaus $\overline{\mathcal{S}} \sim \lambda/\varepsilon$ |
| $(d \geq 1)$1D, long-range | broad size super-ballistic | size growth slowed $\overline{\mathcal{S}} \sim t^d$ |
| free fermion integrable | broad size $\overline{\mathcal{S}} \sim t$ | size decreases $\overline{\mathcal{S}} \sim 1/t$ |

volume, i.e. the number of gates it contains [30, 326]. This relation in fact already contains the essential intuition underlying our work: a connection between the measurement fidelity of a local operator and the operator's growth under Heisenberg evolution. By generalizing the light cone volume using operator size distributions, we will show that this connection is significantly richer and more universal than the above example suggests [397, 398].

We now briefly recap the notion of operator size distributions, as introduced in Chapter 2.3. We first define the size of a Pauli string, $\hat{R}$, as its number of non-identity elements; for instance, $\hat{R} = Y \otimes \mathbb{1} \otimes Z \otimes X$ has size $\mathcal{S}_R = 3$. From this, one can define the size superoperator:

$$\mathcal{S}\{\hat{\mathcal{O}}\} \equiv -\sum_{P_i}(P_i\hat{\mathcal{O}}P_i^\dagger - \hat{\mathcal{O}})/4, \tag{10.1}$$

which gives $\mathcal{S}\{\hat{R}\} = \mathcal{S}_R\hat{R}$, where $\hat{P}_i \in \{\hat{\mathbb{1}}_i, \hat{X}_i, \hat{Y}_i, \hat{Z}_i\}$ are single-qubit Pauli operators [429]. More general operators can be expressed as a sum of Pauli strings, $\hat{\mathcal{O}} = \sum_{\hat{R}} c_R\hat{R}$, and thereby possess a *size distribution*, $P(\mathcal{S}) = \sum_{\{\mathcal{S}_R=\mathcal{S}\}}|c_R|^2$, with normalization $\mathcal{N} = \langle\mathcal{O}^\dagger\mathcal{O}\rangle = \sum_R|c_R|^2$; here, $\langle\cdot\rangle \equiv \text{Tr}(\cdot)/\text{Tr}(\mathbb{1})$ represents the infinite temperature expectation value. We note that the operator size distribution is closely related to out-of-time-ordered correlation functions, $\langle\hat{M}(t)\hat{V}_j\hat{M}(t)\hat{V}_j\rangle$. As an operator, $\hat{M}(t)$, grows to have support on site $j$, the OTOC typically decays to zero. From Eq. (10.1), one immediately sees that the average size of $\hat{M}(t)$ is directly proportional to unity minus the OTOC averaged over all single-qubit Pauli operators:

$$\overline{\mathcal{S}} = \frac{\langle\hat{M}(t)\mathcal{S}\{\hat{M}(t)\}\rangle}{\langle\hat{M}(t)\hat{M}(t)\rangle} = \frac{1}{4}\sum_{P_i}\left(1 - \frac{\langle\hat{M}(t)P_i\hat{M}(t)P_i\rangle}{\langle\hat{M}(t)\hat{M}(t)\rangle}\right). \tag{10.2}$$

## 10.2 Open-system operator growth hypothesis

Let us now turn to open quantum systems. Operator evolution is typically governed by the Lindblad master equation:

$$\partial_t \hat{M} = i[\hat{H}, \hat{M}] - \sum_\alpha \varepsilon_\alpha \left( \hat{L}_\alpha^\dagger \hat{M} \hat{L}_\alpha - \frac{1}{2} \{\hat{L}_\alpha^\dagger \hat{L}_\alpha, \hat{M}\} \right). \tag{10.3}$$

The first term describes unitary time-evolution, while the second describes a sum of local error processes, each characterized by a Lindblad operator, $\hat{L}_\alpha$, and an associated error rate, $\varepsilon_\alpha$.

Our central conjecture is that the effects of local errors on operator growth are in fact captured by a much simpler, effective Lindblad equation:

$$\partial_t \hat{M} = i[\hat{H}, \hat{M}] - \varepsilon \mathcal{S}\{\hat{M}\}, \tag{10.4}$$

where $\mathcal{S}$ is the size superoperator. In effect, this model replaces the original Lindblad operators with isotropic decoherence at each qubit [Eq. (10.1)].

This conjecture is rooted in the following intuition—higher size operators are affected by a greater number of local error processes, and thus decohere at a faster rate[1]. More specifically, we expect large-size components of $\hat{M}$ to typically involve exponentially many Pauli strings varying rapidly in time. This serves to "average" the effect of Lindblad operators such that their action depends solely on whether they are in the support of a given size component of $\hat{M}$, independent of their precise microscopic form. The number of Lindblad operators in the support is directly proportional to the size.

Our framework predicts two effects of open-system dynamics on operator growth, which are captured by the behavior of the Loschmidt echo and the average OTOC (i.e. the average operator size), respectively. For the former, we note that the Loschmidt echo fidelity with respect to a local operator is in fact equal to the *normalization* of the operator's size distribution, $\mathcal{N}(t) = \langle \hat{M}(t)\hat{M}(t) \rangle = \int d\mathcal{S} P(\mathcal{S})$. Our framework predicts[2] that the Loschmidt echo decays in time at a rate equal to the average size multiplied by the error rate:

$$\partial_t \log \mathcal{N}(t) = -2\varepsilon \overline{\mathcal{S}}(t). \tag{10.5}$$

Turning to the OTOC, we note that errors decrease the amplitude of large-size components of $\hat{M}(t)$ at a faster rate than small-size components. Thus, compared to purely unitary evolution, open-system dynamics inhibit the growth of operators. More specifically, we predict that the average size, $\overline{\mathcal{S}}$ [related to the OTOC via Eq. (10.2)], evolves according to:

$$\partial_t \overline{\mathcal{S}}(t) = (\text{unitary}) - 2\varepsilon \delta \mathcal{S}(t)^2. \tag{10.6}$$

---

[1]This behavior has been noted qualitatively in previous works [139, 417, 529] and quantitatively in the context of random circuits [326]. It is familiar from, for example, quantum sensing, where many-body correlations in the GHZ state lead both to an increased sensitivity to magnetic fields as well as to increased dephasing [128].

[2]We derive Eqs. (10.5,10.6) by differentiating $\mathcal{N}$ and $\overline{\mathcal{S}}$ with respect to $t$ and applying Eq. (10.4).

Here, the first term captures the specific unitary dynamics of the system, while the second term decreases the size at a rate proportional to the variance of the size distribution, $\delta \mathcal{S}^2$.

## 10.3 Open-system scrambling dynamics

We now apply our framework to five distinct classes of scrambling dynamics (Table 10.1, Figs. 12.1–10.3): local and all-to-all interacting systems without conservation laws, local systems *with* conservation laws, long-range interacting systems, and free fermion integrable systems (see Appendix E for a detailed discussion of the latter two cases).

We begin by demonstrating that operator growth in two paradigmatic scramblers—systems with no conserved quantities under local and all-to-all interactions—are affected by open-system dynamics in drastically different ways. For the former (focusing on 1D systems for specificity), one expects operators to grow ballistically in time under unitary dynamics, with $\overline{\mathcal{S}} \approx \frac{3}{2} v_B t$, where $v_B$ is the butterfly velocity. Meanwhile, the width of the operator size distribution grows "diffusively", $\delta \mathcal{S} \approx c\sqrt{v_B t}$ where $c$ is a constant [340, 429, 476].

Combining these expectations via Eq. (10.6), we arrive at a simple phenomenological equation for operator growth under *open-system* dynamics, $\partial_t \overline{\mathcal{S}} \approx \frac{3}{2} v_B - \varepsilon(c\sqrt{v_B t})^2$, whose solution yields the prediction: $\overline{\mathcal{S}}(t) \approx \frac{3}{2} v_B t - \frac{c}{2}\varepsilon v_B t^2$. From Eq. (10.5), the Loschmidt echo fidelity thus decays as a Gaussian in time, $\mathcal{N}(t) = \exp\left(-\varepsilon \int_0^t dt' \overline{\mathcal{S}}(t')\right) \approx \exp\left(-\frac{3}{4}\varepsilon v_B t^2\right)$, to leading order in $\varepsilon$. To explore these predictions, we numerically simulate the open-system dynamics of a 1D random unitary circuit (RUC) [340, 341] (see Appendix E). As depicted in Fig. 10.2(a), we find that both the operator size and the Loschmidt echo fidelity (solid lines) agree remarkably well with our phenomenological predictions (dashed lines) across multiple orders of magnitude in the error rate.

In contrast, in all-to-all interacting systems, unitary dynamics typically exhibit "fast scrambling" characterized by the exponential growth of operator size in time, $\overline{\mathcal{S}} \sim e^{\lambda t}$, where $\lambda$ is the Lyapunov exponent [49, 249, 256, 316, 410, 429]. Unlike local systems, the size distribution is also extremely broad, $\delta \mathcal{S} \approx b\overline{\mathcal{S}}$ where $b$ is a constant, owing to the exponential growth of early-time fluctuations [379, 397, 429]. Solving Eq. (10.6), i.e. $\partial_t \overline{\mathcal{S}} \approx \lambda \overline{\mathcal{S}} - \varepsilon b^2 \overline{\mathcal{S}}^2$, then yields an intriguing prediction: under open-system dynamics, the average operator size *plateaus* to a system-size independent value, $\overline{\mathcal{S}}_p \approx \lambda/(\varepsilon b^2)$, after a time $t_p \sim \log(\lambda/(\varepsilon b^2))$. This causes the Loschmidt echo to approach a constant rate of decay, $\mathcal{N}(t) \sim \exp(-\lambda t/b^2)$. Notably, the decay rate, $\lambda/b^2$, is *independent* of the microscopic error rate, $\varepsilon$, echoing seminal results in single-particle quantum chaos [229] and tantalizing recent NMR experiments [139, 416, 418]. As shown in Fig. 10.2(b), both of these predictions are indeed born out by RUC simulations.

One can further sharpen the distinction between open-system dynamics for local versus all-to-all interactions, by analyzing their behavior at asymptotically small error rates. Specifically, consider the value of the Loschmidt echo, $\mathcal{N}_*$, at a time when the open-system dynamics have substantially deviated from the unitary dynamics. In all-to-all systems,
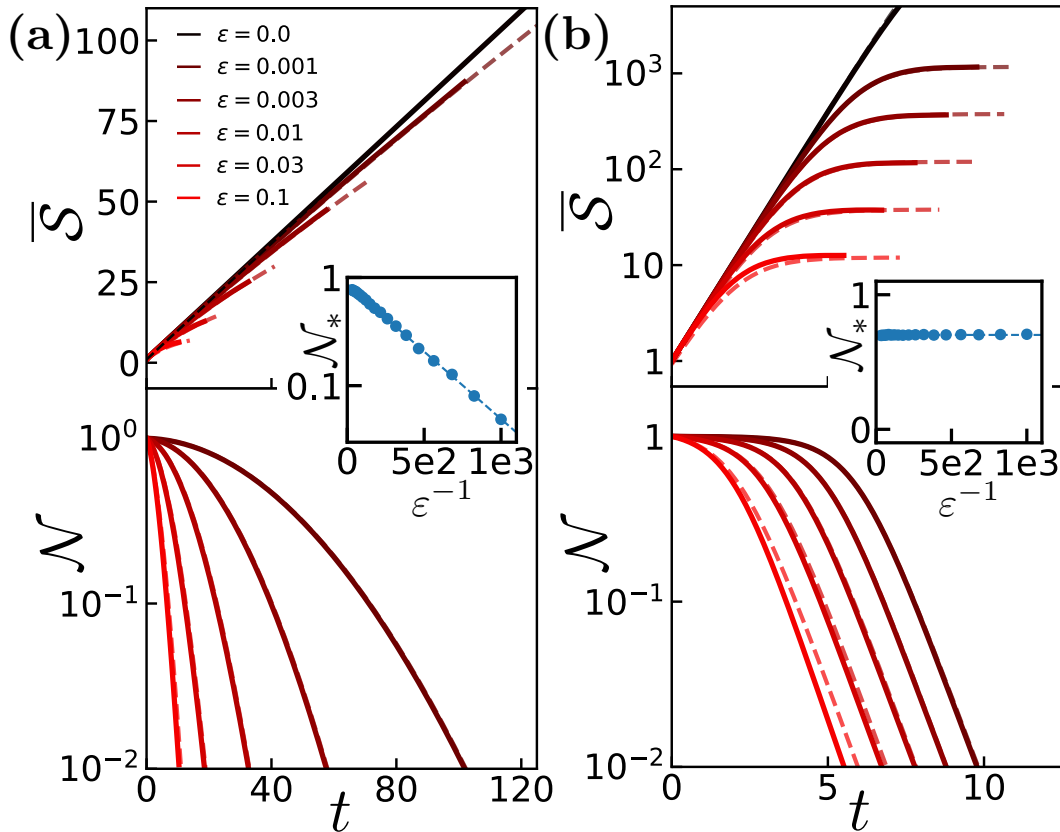
Figure 10.2: **(a)** Average operator size, $\overline{\mathcal{S}}$, and the Loschmidt echo fidelity, $\mathcal{N}$, in a 1D RUC with $N = 200$ (see Appendix E for details). The size grows ballistically with quadratic corrections due to open-system dynamics (solid, data; dashed, theory). Inset: The Loschmidt echo fidelity, $\mathcal{N}_*$, when $\frac{d\overline{\mathcal{S}}}{dt} = 0.9\frac{d\overline{\mathcal{S}}}{dt}\big|_{\varepsilon=0}$, decays exponentially in the inverse error rate, $\varepsilon^{-1}$. **(b)** All-to-all RUC with $N = 1500$ (see Appendix E for details). The size grows exponentially before plateauing to a value which is independent of the system size. The decay rate of the Loschmidt echo is independent of $\varepsilon$ after plateauing (solid, data; dashed, theory). Inset: The Loschmidt echo fidelity, $\mathcal{N}_*$, when $\frac{d\log\overline{\mathcal{S}}}{dt} = 0.9\frac{d\log\overline{\mathcal{S}}}{dt}\big|_{\varepsilon=0}$, is constant with respect to $\varepsilon$.

this occurs shortly after the plateau time, $t_p$, which gives an order one Loschmidt echo, $\mathcal{N}(t_p) \approx \exp\left(-\varepsilon \int_0^{t_p} dt' \, e^{\lambda t'}\right) \approx \exp(-1/b^2)$, independent of the error rate [inset, Fig. 10.2(b)].

In contrast, in 1D systems this occurs when $(\varepsilon v_B t^2/v_B t) \sim 1$, at which point the Loschmidt echo has decayed to an *exponentially* small value, $\mathcal{N}(t) \sim \exp(-v_B/\varepsilon)$ [inset, Fig. 10.2(a)]. For small error rates, this implies that large deviations in operator growth are in practice unobservable for locally-interacting systems, since the signal is exponentially small in $1/\varepsilon$. Physically, this is a direct consequence of the asymptotic separation, $\delta\mathcal{S} \ll \overline{\mathcal{S}}$.

## 10.4 Effects of conservation laws

We now show that the above behaviors are strikingly modified when an operator has overlap with a conserved quantity, $\hat{Q} = \sum_i \hat{q}_i$ (e.g. the total spin, or the Hamiltonian). Such systems feature an interplay between hydrodynamics and scrambling, which is embodied by a 'bimodal' profile for unitary time-evolved operators [246, 385]:

$$\hat{M}(t) = \sum_i q(i,t) \, \hat{q}_i + \sum_{\hat{R} \neq \hat{q}_i} c_R(t) \hat{R}. \tag{10.7}$$

The operator contains both small-size components, $\hat{q}_i$, representing the dynamics of the conserved quantity, as well as large-size Pauli strings, $\hat{R}$, representing scrambled information.

The small-size components arise because an operator's overlap with $\hat{Q}$, $\langle\hat{M}(t)\hat{Q}\rangle = \int dx \, q(x,t)$, is conserved in time. As an example, in chaotic 1D systems, one expects the local overlap, $q(i,t) = \langle\hat{M}(t)\hat{q}_i\rangle$, to spread diffusively, which causes the total normalization of the small-size components to decay in time, $\int dx \, |q(x,t)|^2 \sim 1/\sqrt{t}$. This in turn, implies that the total normalization of the large-size Pauli strings is increasing in time; physically, this corresponds to the dynamics of $q(i,t)$ 'emitting' chaotic components, which spread ballistically from thereon. In combination, this leads to a size distribution (Fig. 12.1), $P(\mathcal{S}) \approx \frac{1}{\sqrt{Dt}}\delta(\mathcal{S}-1) + \frac{v_B}{\sqrt{D}}(\frac{3}{2}v_B t - \mathcal{S})^{-3/2}$, where we have assumed that $\mathcal{S}_{q_i} = 1$ [384].

We expect open-system dynamics to damp the large-size components of $\hat{M}$ by a factor $\sim e^{-\varepsilon\mathcal{S}^2/v_B}$, where $\mathcal{S}^2/v_B$ characterizes the space-time volume of a chaotic component[3]. This effectively truncates the size distribution above $\mathcal{S}_{\mathrm{tr}} \sim \sqrt{v_B/\varepsilon}$ (Fig. 12.1). At late times, $v_B t \gtrsim \mathcal{S}_{\mathrm{tr}}$, this suggests that the average operator size will actually *shrink* in time, since small-size components decay more slowly $\sim t^{-1/2}$, than large-size components, $P(\mathcal{S}_{\mathrm{tr}}) \sim t^{-3/2}$. This sharply contrasts with the behavior of operators that do not overlap conserved quantities, where one expects monotonic growth [Fig. 10.2(a)].

To explore this, we simulate the dynamics of a one-dimensional spin chain and measure the OTOC as a proxy for operator growth. For an operator that does not overlap with

---

[3]In systems with conserved quantities, the large-size Pauli strings also feature a heterogeneous structure, in which Pauli strings are more likely to commute with the conserved quantity near the wavefront of operator spreading, and only approach fully random behavior in the interior [246, 385]. In principle, this implies that the wavefront is less susceptible to errors that preserve the conserved quantity.
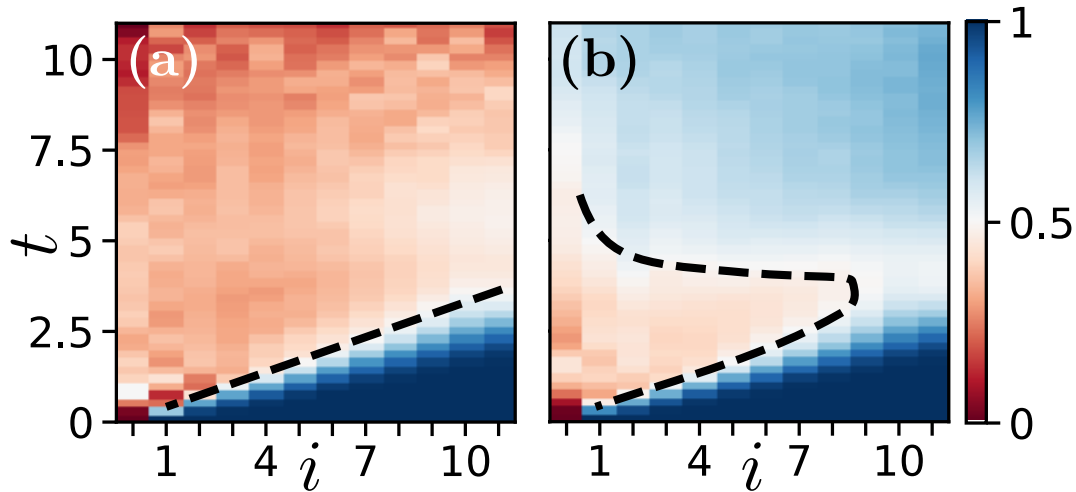
Figure 10.3: The OTOC as a function of time and space for an $N = 12$ one-dimensional spin chain. **(a)** Operators that do not overlap the Hamiltonian exhibit an OTOC which follows a ballistic light cone. **(b)** For an operator that overlaps with the Hamiltonian, the OTOC at a given site $i$ initially decays, before increasing at later times. The specific OTOC we calculate takes the form, $\frac{1}{4} \sum_P \langle e^{-iH_1 t} \hat{M} e^{iH_1 t} \hat{P}_i e^{-iH_2 t} \hat{M} e^{iH_2 t} \hat{P}_i \rangle / \mathcal{N}(t)$, where the forwards and backwards time-evolution are governed by two distinct 1D Hamiltonians, $H_1 = H_2 + \eta \, \delta H$ (see Appendix E for details).

the Hamiltonian, we find that the OTOC decays monotonically following a linear light-cone [Fig. 10.3(a)]. For an operator exhibiting overlap, we find that the decay of the OTOC indeed *reverses* as a function of time, indicative of a decrease in the average operator size [Fig. 10.3(b)]. Interestingly, this insight immediately resolves an apparent disagreement between previous studies of open-system operator growth. In particular, certain studies found that OTOCs were only minimally affected by errors [326, 473], while others found a dramatic reversal of scrambling [450, 529]. We attribute this difference to the presence or absence of conservation laws.

## 10.5   Discussion and outlook

The results of this Chapter lead to a number of implications. First, we provide a new perspective on protocols which divide error-prone OTOC measurements by an independent characterization of the error [326, 418, 450, 473]. In our language, the latter is precisely the normalization, $\mathcal{N}(t) = \langle \hat{M}(t) \, \hat{M}(t) \rangle$. To this end, these protocols will only replicate unitary dynamics when the total error is small ($1 - \mathcal{N} \approx \varepsilon \int_0^t dt \, \overline{\mathcal{S}} \ll 1$) or when size distributions are tightly peaked, $\delta \mathcal{S} \ll \overline{\mathcal{S}}$. Notably, this also applies to the teleportation experiments presented in Chs. 7 and 8. In these protocols, the above division is performed "automatically" within

quantum mechanics, upon post-selecting on the EPR measurement result.

Second, our results suggest a novel protocol for measuring operator size distributions (Fig. 10.4), which circumvents the need to either perform exponentially many measurements [381] or utilize two entangled copies of the system [429]. Specifically, in order to measure the *generating function* of the size distribution, $G_{\mathcal{S}}(\mu) = \sum_{\mathcal{S}} P(\mathcal{S}) e^{-\mu \mathcal{S}}$, we propose the following protocol (Fig. 10.4): (i) prepare an initial state, $\rho = (\mathbb{1} + \hat{M}) \otimes \mathbb{1}^{\otimes N-1}/2^N$, (ii) time-evolve forward, e.g. via a unitary operation, $U$, (iii) apply a set of single-qubit Pauli operators, $\{P_1, \ldots, P_N\}$, (iv) time-evolve backward via $U^{\dagger}$, and (v) measure $\hat{M}$. If the intervening Pauli operators are fixed, this reduces to previous schemes for measuring OTOCs [163]. However, if one randomly samples each Pauli matrix in each experimental shot, with probability $p = (1 - e^{-\mu})/4$ to be $\{X, Y, Z\}$ and probability $1 - 3p$ to be the identity, this in effect implements a decoherence channel, $e^{-\mu \mathcal{S}}$, that explicitly depends on the size superoperator. The fidelity to recover the initial state then gives the generating function via $\mathcal{F} = \frac{1}{2}[1 + \mathcal{N} G_{\mathcal{S}}(\mu)]$, where $\mathcal{N}$ can be measured by setting $\mu = 0$.

Finally, we conjecture that our framework also applies to an alternate scenario (often explored in experiments [138, 139, 416, 418]), where one evolves forward via a Hamiltonian, $H$, and backward via a perturbed Hamiltonian[4], $-H + \eta \, \delta H$ [229]. Naively, this scenario features perturbations that are highly correlated in time and space, and thus outside the Lindbladian framework. However, in a chaotic many-body system, one expects such correlations to quickly decay outside of some thermalization time- ($\tau_{\text{th}}$) and length-scale ($\xi_{\text{th}}$). This assumption leads to a Fermi's golden rule [118, 229] estimate of an effective decoherence rate, $\partial_t \log(\mathcal{N}) \sim \eta^2 \tau_{\text{th}} \xi_{\text{th}} \overline{\mathcal{S}}$, which scales linearly with the average operator size (see Appendix E). Somewhat intriguingly, recent NMR experiments [139] precisely observe this linear scaling with operator size for $\eta \gtrsim 0.1$; this transitions to a square root scaling at smaller $\eta$ and developing a microscopic understanding of this regime remains an open question.

Finally, looking forward, our results also have implications for the classical simulability of open quantum systems—if operator sizes are bounded from above by a constant, $\mathcal{S}_{\varepsilon}$, then time-evolution is in principle efficiently simulable, since the dimension of the accessible operator Hilbert space is polynomial in the system size, $\sim N^{\mathcal{S}_{\varepsilon}}$. A similar idea was recently proposed in diffusive 1D spin chains [384]; our results suggest that it may hold more broadly.

---

[4]Note that our original use of the term Loschmidt echo refers to the case where one evolves forwards and backwards with the same Hamiltonian in the presence of noise and errors. The term Loschmidt echo is also used to describe the scenario here, where one evolves forwards and backwards with two different Hamiltonians.
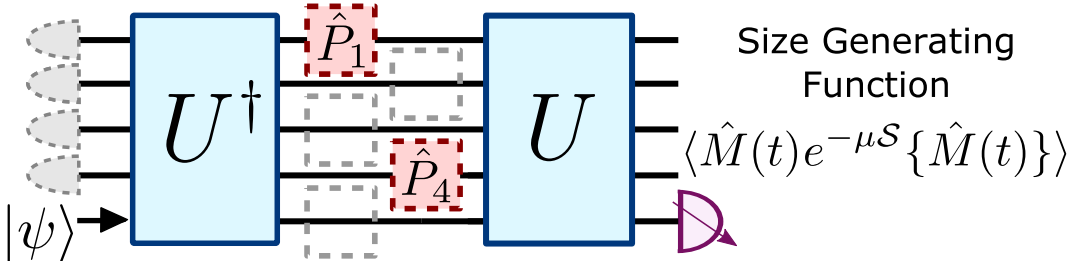
Figure 10.4:    Protocol to measure the generating function, $G_{\mathcal{S}}(\mu)$, of the operator size distribution.  Gray qubits are initially random in the computational basis, $|\psi\rangle$ is an $\hat{M}$ eigenstate, and each site is acted upon randomly by either the identity (gray) or a non-identity Pauli operator (red) in each experimental shot.

# Chapter 11

# Scrambling and complexity in phase space

theoremcounter conjecturecounter definitioncounter

The vast majority of studies of quantum information scrambling—including all other Chapters in Parts III and IV—have focused on discrete variable (DV) systems, where the many-body Hilbert space is composed of a tensor product of local qubits. One of the defining features of such DV scrambling is the notion of operator growth, where the time evolution of an initially simple, local operator $\mathcal{V}$, yields a more complex, late-time operator, $\mathcal{V}(t) = U^{\dagger}(t)\mathcal{V}U(t)$ [398], whose decomposition is dominated by non-local operator strings. A particularly powerful quantitative diagnostic of operator growth is provided by the so-called out-of-time-order correlation (OTOC) function $\langle \mathcal{V}^{\dagger}(t)\mathcal{W}^{\dagger}(0)\mathcal{V}(t)\mathcal{W}(0)\rangle$, which measures the spreading of $\mathcal{V}(t)$ via another local probe operator $\mathcal{W}$ [49, 246, 277, 340, 385, 476, 505].

While discrete variable information dynamics are relatively well understood, both from an information-theoretic perspective and in terms of physical interpretation, their continuous variable (CV) cousin remains poorly explored. This owes to a number of intrinsic subtleties associated with CV systems. For example, their infinite dimensional local Hilbert space leads to technical and conceptual challenges requiring the need for short- and long-distance cutoffs as well as ambiguities related to properly defining the 'volume' of an operator. Moreover, for CV systems, there naturally exists *two* distinct notions of information scrambling; namely, scrambling *within* the phase space of a single degree of freedom and scrambling *across* the phase space of many coupled degrees of freedom.

To this end, the broad goal of this Chapter is to lay out a theoretical foundation for investigating quantum information scrambling in CV systems [65, 89, 94, 114, 182, 200, 386, 398, 405]. This is motivated in part, by an abundance of strongly-interacting, controllable physical systems, whose microscopic degrees of freedom are continuous variable; these include quantum optical systems, cavity/circuit QED, photonic networks and more abstractly, quantum field theories in general.

## 11.1   Summary of results

We begin with a brief summary of the results and structure of this Chapter.

*Analogy between DV and CV systems*—Throughout the Chapter, we find it illuminating to frame our results in analogy with well-known ideas from DV systems. Here, we will begin by introducing this dictionary before summarizing the organization of the remainder of the paper.

In an $N$-qudit DV system (dimension $d = q^N$ with $q$-state qudits), it is common to diagnose scrambling by observing how a single-qudit Pauli operator, $P$, evolves under unitary time evolution, $U(t)$. Since the Pauli operators form a complete basis, the time-evolved $P(t)$ can be re-expanded as,

$$P(t) = \sum_{Q \in \text{Pauli}} f[Q; P(t)] Q \qquad (11.1)$$

where $|f[Q; P(t)]|^2$ can be interpreted as the probability distribution of $P(t)$ over $Q$ and is normalized such that, $\sum_Q |f[Q; P(t)]|^2 = 1$. As aforementioned, scrambling in DV systems corresponds to the fact that strongly-interacting time evolution generically leads the operator expansion in Eqn. (1) to have significant overlaps with 'high-weight' Pauli operators. Here, the weight of a Pauli operator $Q$ quantifies its non-locality and equals the number of qudits on which $Q$ acts non-trivially.

A central motif of our work is the analogy between Pauli operators in DV systems and *displacement operators* in CV systems. In a single-mode CV system (e.g. a simple harmonic oscillator), the displacement operators (which shift a coherent state in phase space) form a complete basis:

$$D(\xi_1, \xi_2) \equiv \exp\left[i(\xi_2 q - \xi_1 p)\right], \qquad (11.2)$$

where $\xi_1$ and $\xi_2$ are (respectively) the shifts of $q$ and $p$, the canonical position and momentum quadrature operators. More generally, an $N$-mode displacement operator can be written as the tensor product of local displacements.

We propose to characterize information scrambling in CV systems by considering the time-evolution of displacement operators, $D(\boldsymbol{\xi}_1; t) \equiv U(t)^\dagger D(\boldsymbol{\xi}_1) U(t)$. To understand why, let us begin by decomposing,

$$D(\boldsymbol{\xi}_1; t) = \frac{1}{\pi^N} \int d^{2N}\boldsymbol{\xi}_2 \, \chi[\boldsymbol{\xi}_2; D(\boldsymbol{\xi}_1; t)] D(-\boldsymbol{\xi}_2), \qquad (11.3)$$

where $\chi(\boldsymbol{\xi}; A) \equiv \text{tr}[AD(\boldsymbol{\xi})]$ is the Wigner characteristic function. Initially, the characteristic function $\chi[\boldsymbol{\xi}_2; D(\boldsymbol{\xi}_1; 0)] = \pi^N \delta(\boldsymbol{\xi}_2 + \boldsymbol{\xi}_1)$ is highly localized; thus, scrambling in CV systems can be identified using the spread of $\chi[\boldsymbol{\xi}_2; D(\boldsymbol{\xi}_1; t)]$ in phase space for generic choices of $\boldsymbol{\xi}_1$, much like the spreading of $P(t)$ into high-weight Pauli operators in the DV case.

With this analogy in hand, let us now introduce the organization and summary of our results.

**Quasi vs genuine scrambling:** In Section 11.2, we introduce an important distinction between two types of scrambling, which we term: quasi scrambling and genuine scrambling. In DV systems, both types are captured by the growth of $P(t)$ into high-weight operators (i.e. large-scale entanglement is generated). The key distinction lies in the nature of the operator *distribution* as quantified by $|f[Q; P(t)]|^2$. In particular, we refer to a system as a quasi scrambler if the operator distribution remains localized on a *single* high-weight operator, and as a genuine scrambler if the distribution is spread out over a *large* number of high-weight operators. This difference in operator distribution behavior can be unambiguously identified via OTOCs and roughly corresponds to whether the scrambling unitary is generated by Clifford (quasi) or non-Clifford (genuine) operators.

Having defined two classes of scrambling in DV systems, we will find that the same distinction applies for CV systems. In particular, we propose that CV quasi scrambling corresponds to situations where the Wigner characteristic function stays localized, while genuine CV scrambling results in the delocalization of the characteristic function. Interestingly, in the case of CV systems, this distinction corresponds to whether the scrambling unitary is generated by Gaussian (quasi) or non-Gaussian (genuine) operators.

**Operator spreading and OTOCs:** In Section 11.3, we develop theoretical tools to study operator spreading in CV systems using CV OTOCs. We begin by establishing an intuitive measure of genuine scrambling, in terms of the volume of a time-evolved operator's distribution in phase space. Then, we present a Fourier transform-like formula which relates this distribution to OTOCs of the time-evolved operator, from which we see that individual OTOCs can detect the distinction between genuine and quasi scramblers. This allows us to propose a physical observable—the OTOC magnitude—which measures the non-Gaussianity of the dynamics. Furthermore, we show that averages of OTOCs, weighted over ensembles of displacement operators, can probe an operator's spread in phase space with tunable short and long-distance cutoffs. This encompasses both operator spreading in real space, as studied in DV systems, as well as within a single mode. Finally, we apply our findings to study operator growth in two examples of genuine scramblers: time-evolution via cubic phase gates and the Hellon-Heiles potential.

**Random Gaussian circuits:** In Section 11.4, we turn our attention to quasi scrambling. Despite the fact that quasi scramblers do not fully spread operators (i.e. they simply map one displacement operator to another), useful insights into genuine scrambling can be obtained by studying the stochastic evolution of quasi scramblers. Indeed, recent progress [68, 198, 199, 203, 211, 213] on the interplay between holography and quantum information theory has revealed that quantum circuits composed of short-range random Clifford unitaries can provide useful intuition for understanding entanglement and operator propagation in DV many-body systems [49, 246, 340, 385, 476, 505].

This motivates us to explore local random Gaussian circuits (Sec. 11.4) as an analytically and numerically tractable toy model of CV scrambling. We observe a number of intriguing features. In particular, we find that such random circuits exhibit exponential growth of displacements *within* each local mode, while operators spread *ballistically* to distant modes. In the former case, the observed exponential growth also leads to other surprising conse-

quences; for example, entanglement in the system grows quadratically in time, in contrast to the previously found linear growth [340]. In the latter case, we observe that the butterfly velocity $v_B$ of the ballistic spread depends on the local growth exponent in a similar fashion to large-$N$ models. To this end, we propose a simple hydrodynamical model which relates the butterfly velocity, the growth exponent and the diffusion constant.

**Unitary designs for CV systems:** Building upon our exploration of random Gaussian unitaries (see Section 11.4), in Section 11.4 we investigate the statistical properties of ensembles of such unitaries. In particular, we attempt to construct CV analogs of Haar randomness and unitary designs[1]. To begin, we provide a plausible definition for unitary designs in CV systems. Using this definition, we find that Gaussian distributed displacement operators asymptotically form a CV 1-design. However, we find that Gaussian unitaries do not form an exact 2-design—consistent with prior results on state designs [62]—but nevertheless capture many qualitative 2-design features in the limit of large squeezing. Finally, we provide a unique generalization of the so-called 'frame potential' to CV systems, enabling the quantitative verification of designs.

**Experimental realizations of CV scrambling:** In Section 11.5, we propose and analyze a concrete experimental realization of CV scrambling in a cavity-QED architecture, where non-Gaussian unitaries are generated via the SNAP (selective number-dependent arbitrary phase) gate [208, 268]. Next, we present concrete protocols for the measurement of both individual and average OTOCs in CV systems. In addition, we also present a CV analog of a teleportation-based protocol for verifying scrambling [274, 517, 518]. Interestingly, while measurement uncertainties may damage the teleported quantum state, our CV teleportation protocol can be made fault-tolerant by using the Gottesman-Kitaev-Preskill (GKP) code (when the squeezing parameter is sufficiently large).

Finally, in Section 11.6, we conclude by offering our perspective on a number of exciting open questions and directions.

## 11.2 Genuine and quasi scrambling

We begin by presenting a broad overview of scrambling in DV and CV systems. Scrambling refers to delocalization of quantum information over the entire system under unitary dynamics $U(t) \equiv \exp\left(-i \int H(t)dt\right)$, where $H(t)$ is the Hamiltonian of the system and can in principle be time dependent. In terms of operators, an initially simple operator $\mathcal{O}$ evolves to a more complex operator $\mathcal{O}(t) = U^\dagger(t)\,\mathcal{O}U(t)$, e.g. local operators become highly non-local [398]. In an $N$-qudit DV system ($d = q^N$ with $q$-state qudits), it is common to diagnose scrambling by observing how a single-qudit Pauli operator $P$ evolves in time. Since Pauli operators form a complete basis $- \frac{1}{d}\operatorname{Tr}\left(PQ^\dagger\right) = \delta_{P,Q}$ for $P, Q \in$ Pauli and

---

[1]In DV systems, our current understanding of quantum chaos and random matrix theory has been deeply influenced by ideas and tools from Haar randomness.

$\frac{1}{d} \sum_{P \in \mathrm{Pauli}} P \mathcal{O} P^\dagger = \mathrm{Tr}(\mathcal{O}) \boldsymbol{I}$ – the time-evolved $P(t)$ can be expanded as

$$P(t) = \sum_{Q \in \mathrm{Pauli}} f[Q; P(t)] Q, \qquad \sum_Q |f[Q; P(t)]|^2 = 1, \tag{11.4}$$

where the second constraint results from the unitarity of $U(t)$. As such, $|f[Q; P(t)]|^2$ can be interpreted as a probability distribution over $Q$ for each $P(t)$.

Scrambling in DV systems corresponds to the growth of $f[Q; P(t)]$ such that the quantity $|f[Q; P(t)]|^2$ for 'high-weight' Pauli operators becomes significant. Here the weight of a Pauli operator $Q$ quantifies the non-locality of $Q$, and equals the number of qudits on which $Q$ acts non-trivially. Within this correspondence, we identify two different classes of scrambling dynamics. Time-evolution $U$ is called *quasi scrambling* if the operator distribution $|f[Q; P(t)]|^2$ remains concentrated on only a few Paulis $Q$ even as the weight of these Paulis grows significantly. In contrast, time-evolution $U$ is called *genuine scrambling* if $|f[Q; P(t)]|^2$ has support on many high-weight Pauli operators for a generic low-weight Pauli operator $P(0)$.

Interestingly, these two classes of scramblers are related to Clifford and non-Clifford unitary operators. Clifford operators are unitary operators which transform Pauli operators into Pauli operators; $U P U^\dagger \in$ Pauli for all $P \in$ Pauli. Examples include the Hadamard gate and the Control-Not gate. Under a random Clifford operator, the time-evolved operator $P(t)$ becomes a high-weight $\sim O(N)$ Pauli operator. However, by definition, the operator distribution $f[Q; P(t)]$ remains concentrated on a single Pauli $Q = P(t)$, hence it is only quasi scrambled. On the other hand, if $U$ is a non-Clifford operator, e.g. a Haar random unitary, the operator distribution becomes almost uniform,

$$|f[Q; P(t)]|^2 \sim \frac{1}{d^2 - 1} \qquad Q \neq \boldsymbol{I}, \tag{11.5}$$

achieving genuine scrambling.

Many previous studies in quantum information literature recognize random Clifford unitaries as scramblers since they delocalize operators and create nearly maximal entanglement when applied to arbitrary product states (see [73, 278] for instance). However, Clifford operators represent only a very restricted subset of all possible time-evolutions, and one expects that many aspects of thermalization in many-body quantum systems require more complex unitaries to capture. Hence, we think it is important to distinguish two classes of scramblers. A certain plausible definition of scrambling which distinguishes two different classes was proposed in Ref. [517] based on late-time asymptotic behavior of OTOCs.

We now turn our attention to CV systems. A central motif of our work will be an analogy between Pauli operators in DV systems and *displacement operators* in CV systems. In a single-mode CV system (e.g. the simple harmonic oscillator), the displacement operator shifts a coherent state by position $\xi_1$ and momentum $\xi_2$ in phase space. It takes the form

$$D(\xi_1, \xi_2) \equiv \exp\left[i(\xi_2 q - \xi_1 p)\right], \tag{11.6}$$
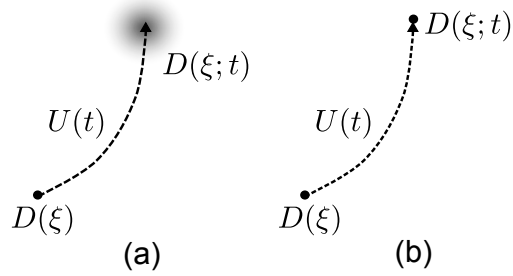
Figure 11.1: Operator spreading in two types of CV scrambling. (a) In genuine scrambling, an initial displacement operator time-evolves into a sum of many displacements, spread throughout phase space. (b) In quasi scrambling, displacements may move around phase space, but remain localized.

where $q$ and $p$ are the canonical position and momentum quadrature operators. Similar to Paulis, the displacement operators of an $N$-mode CV system are formed as tensor products of local displacement operators, written as

$$D\left(\boldsymbol{\xi}\right) \equiv \bigotimes_{k=1}^{N} D^{k}\left(\xi_{2k-1}, \xi_{2k}\right),\tag{11.7}$$

where $D^{k}\left(\xi_{2k-1}, \xi_{2k}\right)$ is the mode $k$ single-mode displacement operator, and $\boldsymbol{\xi} \in R^{2N}$ is a $2N$-component vector of displacements. Like Paulis, displacement operators form a complete operator basis, obeying $\mathrm{tr}\left(D\left(\boldsymbol{\xi}\right) D\left(\boldsymbol{\xi}'\right)\right) = \pi^{N}\delta\left(\boldsymbol{\xi}+\boldsymbol{\xi}'\right)$ and $1/\pi^{N} \int d^{2N}\boldsymbol{\xi}\ D\left(\boldsymbol{\xi}\right) A D^{\dagger}\left(\boldsymbol{\xi}\right) = \mathrm{tr}\left(A\right)\boldsymbol{I}$.

Inspired by the similarities between displacement operators and Paulis, we will characterize scrambling in CV systems by considering the time-evolution of displacement operators, $D\left(\boldsymbol{\xi}_{1};t\right) \equiv U\left(t\right)^{\dagger} D\left(\boldsymbol{\xi}_{1}\right) U\left(t\right)$. The completeness of displacement operators allows the decomposition

$$D\left(\boldsymbol{\xi}_{1};t\right) = \frac{1}{\pi^{N}} \int d^{2N}\boldsymbol{\xi}_{2}\ \chi\left[\boldsymbol{\xi}_{2}; D\left(\boldsymbol{\xi}_{1};t\right)\right] D\left(-\boldsymbol{\xi}_{2}\right),\tag{11.8}$$

where $\chi\left(\boldsymbol{\xi}; A\right) \equiv \mathrm{tr}\left[A D\left(\boldsymbol{\xi}\right)\right]$ is known as the Wigner characteristic function. Initially, the characteristic function $\chi\left[\boldsymbol{\xi}_{2}; D\left(\boldsymbol{\xi}_{1};0\right)\right] = \pi^{N}\delta\left(\boldsymbol{\xi}_{2}+\boldsymbol{\xi}_{1}\right)$ is highly localized in phase space; scrambling in CV systems can therefore be characterized by the growth of $\chi\left[\boldsymbol{\xi}_{2}; D\left(\boldsymbol{\xi}_{1};t\right)\right]$. As in the DV systems, we identify two distinct classes of scrambling dynamics (visualized in Fig. 11.1). Time-evolution $U\left(t\right)$ is *quasi scrambling* if $\chi\left[\boldsymbol{\xi}_{2}; D\left(\boldsymbol{\xi}_{1};t\right)\right]$ remains highly localized in phase space, but spreads to multiple modes. Time-evolution $U\left(t\right)$ is *genuine scrambling* if $\chi\left[\boldsymbol{\xi}_{2}; D\left(\boldsymbol{\xi}_{1};t\right)\right]$ spreads significantly over phase space for generic choices of $\boldsymbol{\xi}_{1}$.

This separation of CV scramblers can be related to a common classification of CV unitaries into Gaussian and non-Gaussian operators. A complete introduction to Gaussian

unitaries can be found in Ref. [485]; we provide a brief overview here. We begin by condensing notation, defining the vector of quadrature operators $\boldsymbol{x} = (q_1, p_1, \cdots, q_N, p_N)$. This allows us to concisely write an $N$-mode displacement operator as $D(\boldsymbol{\xi}) = \exp\left(i\boldsymbol{x}^T\boldsymbol{\Omega}\boldsymbol{\xi}\right)$, defining the block diagonal matrix $\boldsymbol{\Omega} = \bigoplus_{k=1}^{N} \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. The product of displacement operators is now given by the simple addition,

$$D(\boldsymbol{\xi}_1) D(\boldsymbol{\xi}_2) = e^{-i\boldsymbol{\xi}_1^T \boldsymbol{\Omega}\boldsymbol{\xi}_2} D(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2). \tag{11.9}$$

Gaussian unitaries are generated by Hamiltonians that are second order in the quadrature operators. One can represent a Gaussian unitary via its action on quadrature operators, which takes the form

$$U_{\boldsymbol{S},\boldsymbol{d}}^{\dagger} \boldsymbol{x} U_{\boldsymbol{S},\boldsymbol{d}} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{d}, \tag{11.10}$$

where the unitary is labelled by a $2N$-component displacement $\boldsymbol{d}$ and a $2N \times 2N$ symplectic matrix $\boldsymbol{S}$, $\boldsymbol{S}\boldsymbol{\Omega}\boldsymbol{S}^T = \boldsymbol{\Omega}$. From this, one can show that Gaussian unitaries transform displacement operators into other displacement operators according to

$$U_{\boldsymbol{S},\boldsymbol{d}}^{\dagger} D(\boldsymbol{\xi}) U_{\boldsymbol{S},\boldsymbol{d}} = \exp\left(i\boldsymbol{d}^T\boldsymbol{\Omega}\boldsymbol{\xi}\right) D\left(\boldsymbol{S}^{-1}\boldsymbol{\xi}\right), \tag{11.11}$$

analogous to the action of Clifford operators on Paulis in DV systems. As such, the Wigner characteristic function of $D(\boldsymbol{\xi}_1; t)$ remains highly localized under time-evolution by a Gaussian unitary, and so we identify Gaussian unitaries as quasi scramblers. Genuine scramblers correspond to non-Gaussian unitaries, generated by Hamiltonians that are third or higher order in the quadrature operators. This can arise from interactions, hard boundary conditions (e.g. quantum billiard systems), or non-linear gates such as the single-mode cubic phase gate [172] or the Kerr effect [136] in quantum optical experiment.

## 11.3  Operator spreading in genuine scrambling

In this section, we develop a basic formalism to characterize a time-evolved operator's distribution in *phase space* as a probe of genuine CV scrambling. We begin by demonstrating that one aspect of this distribution—its volume—can be measured in two ways: via averages of *time-ordered* correlation (TOC) functions of (generically non-local) displacement operators, and via the so-called frame potential, previously used to study the complexity and pseudorandomness of DV unitary ensembles [399]. While the two measures are identical at infinite temperature, they differ when regularizing by a density matrix, which will be necessary in CV systems. To probe finer-grained aspects of operator spreading we turn to OTOCs, establishing a precise relation between a time-evolved operator's phase space distribution and its OTOC with displacement operators. One consequence of this relation is a constraint on OTOC decay: specifically, the magnitude of the OTOC can *only* decay during genuine scrambling (non-Gaussian dynamics), and is strictly 1 in quasi scrambling (Gaussian dynamics). Further, this relation prompts us to consider more general averages

of OTOCs over ensembles of displacement operators. We show that average OTOCs measure a coarse-grained density of an operator's phase space distribution; their change in time thus characterizes a 'flow' of the distribution. To conclude, we apply these tools to study two specific non-Gaussian Hamiltonians, an exactly-solvable cubic potential and the chaotic Henon-Heiles potential, and show that both models lead to operator spreading and OTOC decay.

## Phase space volume and the frame potential

We begin by introducing our notion of an operator's 'phase space volume'. We show that in DV systems one can define such a volume for time-evolved Pauli operators, (roughly) by counting the number of Pauli strings with non-zero weight when expanding the operator in the Pauli basis. This can be probed by TOCs at infinite-temperature or, equivalently, the $k = 1$ frame potential. In generalizing this to CV systems, one encounters various divergences due to the lack of an infinite-temperature limit, as the CV system's Hilbert space is not bounded. To regulate this, we expand our definition of phase space volume to be with respect to a normalized density matrix $\rho$. In sufficiently low-temperature regimes, this 'coarse-grained' volume behaves qualitatively differently from the previous volume. We demonstrate this in multiple examples for the specific case of a thermal density matrix with respect to the number operator Hamiltonian.

### Defining phase space volume for CV systems

We first introduce an operator volume in DV systems. Consider a system of $N$ qudits, with total Hilbert space dimension $d = q^N$. As Pauli operators form a complete basis, the time-evolution of a Pauli operator $P$ may be decomposed as $P(t) = \sum_{Q \in \text{Pauli}} f[Q; P(t)] Q$. The coefficients $f[Q; P(t)]$ coincide with the TOC of $P$ and $Q$ with respect to the infinite-temperature density matrix $\rho_\infty = \frac{1}{d}\mathbf{1}$,

$$f[Q; P(t)] = \mathcal{C}_1(P(t), Q(0)) \equiv \frac{1}{d}\text{tr}\left(P(t)Q^\dagger(0)\right). \tag{11.12}$$

One can show that $\sum_Q |f[Q; P(t)]|^2 = 1$. This allows one to consider the *ensemble* of Pauli operators $\mathcal{E}[P(t)] = \{Q | Q \sim |f[Q; P(t)]|^2\}$, defined by the normalized probability distribution $|f[Q; P(t)]|^2$. One can now define a volume of the ensemble $\mathcal{E}[P(t)]$ corresponding to the number of $Q$ with significant $|f[Q; P(t)]|^2$. This can be made rigorous using entropies of the distribution $|f[Q; P(t)]|^2$. Specifically, the exponential of the Rényi-2 entropy

$$\text{vol}(\mathcal{E}[P(t)]) \equiv 2^{S^{(2)}_{\mathcal{E}[P(t)]}} = \frac{1}{\sum_{Q \in \text{Pauli}} |f[Q; P(t)]|^4} \tag{11.13}$$

provides a good measure of such a volume. The Rényi-2 entropy of the ensemble is related to the frame potential [399], which we find corresponds to an inverse volume:

$$\mathcal{F}_{\mathcal{E}[P(t)]} \equiv \mathbb{E}_{U,V \sim \mathcal{E}[P(t)]}\left\{\left|\text{tr}\left(U^\dagger V\right)\right|^2\right\} = d^2\, 2^{-S^{(2)}_{\mathcal{E}[P(t)]}}, \tag{11.14}$$

where $\mathbb{E}_{U,V\sim\mathcal{E}[P(t)]}$ denotes the expectation value when $U, V$ are sampled independently from the ensemble $\mathcal{E}[P(t)]$.

In generalizing this relation to CV systems, we encounter one of our first obstacles in working with infinite-dimensional systems: the traces used in DV expressions, regulated by the infinite-temperature density matrix, are often infinite in the CV context, where infinite-temperature is not well-defined. For instance, when trying to characterize the volume of a time-evolved displacement operator $D\left(\boldsymbol{\xi}_1; t\right)$, the analog of $f\left[Q; P(t)\right]$ would naturally be the characteristic function $\chi\left[\boldsymbol{\xi}; D\left(\boldsymbol{\xi}_1; t\right)\right]$. However, the volume of the characteristic function is not well-defined, as $\left|\chi\left[\boldsymbol{\xi}; D\left(\boldsymbol{\xi}_1; t\right)\right]\right|^2$ has infinite norm [for example, at time $t = 0$, $\chi \sim \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_1)$].

To remedy such divergences, it is natural to consider TOCs at a finite temperature,

$$\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho = \operatorname{tr}\left[\rho D\left(\boldsymbol{\xi}_1; t\right) D\left(\boldsymbol{\xi}_2\right)\right]. \tag{11.15}$$

Expressed in terms of the operators' characteristic functions, the TOC can be viewed as $\chi\left[\boldsymbol{\xi}; D\left(\boldsymbol{\xi}_1; t\right)\right]$ 'smeared' over a width in phase space determined by $\chi\left[\boldsymbol{\xi}; \rho\right]$,

$$\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho =$$
$$\frac{1}{\pi^N} \int d^{2N}\boldsymbol{\xi} \; \chi\left[\boldsymbol{\xi}; D\left(\boldsymbol{\xi}_1; t\right)\right] \; e^{i\boldsymbol{\xi}^T \boldsymbol{\Omega} \boldsymbol{\xi}_2} \chi\left(\boldsymbol{\xi} - \boldsymbol{\xi}_2; \rho\right). \tag{11.16}$$

For the sake of illustration, throughout the Chapter we will frequently take $\rho$ to be the thermal density matrix for the number operator Hamiltonian, $\tilde{\rho}_{n_{\text{th}}} \sim \bigotimes_{k=1}^N e^{-\beta_{n_{\text{th}}} a_k^\dagger a_k}$, which we refer to as the 'thermal density matrix' for convenience. Here the effective temperature $\beta_{n_{\text{th}}}^{-1} = \ln\left(1 + \frac{1}{n_{\text{th}}}\right)^{-1}$ is proportional to the mean photon number per mode $n_{\text{th}}$ in the CV limit $n_{\text{th}} \gg 1$. The characteristic function of the thermal density matrix has a width $\sim 1/n_{\text{th}}$ in phase space, $\chi\left(\boldsymbol{\zeta}; \tilde{\rho}_{n_{\text{th}}}\right) = \exp\left(-\left(n_{\text{th}} + 1/2\right)|\boldsymbol{\zeta}|^2\right)$. In the high-temperature limit ($n_{\text{th}} \gg 1$), this width becomes small, and $\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho$ becomes proportional to the unregulated $\chi\left[\boldsymbol{\xi}_2; D\left(\boldsymbol{\xi}_1; t\right)\right]$.

Unlike their infinite temperature counterparts, finite temperature TOCs possess a well-defined norm $\mathcal{N}_\rho \equiv \int d\boldsymbol{\xi}_2 |\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho|^2 = \pi^N \operatorname{tr}\left(\rho^2\right)$. This allows us to proceed similarly to the DV case and consider the ensemble of displacement operators $\mathcal{E} = \{D\left(\boldsymbol{\xi}_2\right) | \boldsymbol{\xi}_2 \sim |\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho|^2 / \mathcal{N}_\rho\}$. The volume of such an ensemble can be defined similarly to Eq. (11.13),

$$\operatorname{vol}\left(\mathcal{E}; \rho\right) = \left(\frac{1}{4\pi}\right)^N \frac{\mathcal{N}_\rho^2}{\int d\boldsymbol{\xi} \; |\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho|^4}. \tag{11.17}$$

In Appendix F, we use this definition to provide quantum analogs of Liouville's theorem and the Kolmogorov-Sinai entropy for Gaussian time-evolution.

### Phase space volume and the CV frame potential

While the above satisfactorily defines the phase space volume of CV operators, we find it interesting to also generalize Eq. (11.14), which provides a powerful interpretation of the
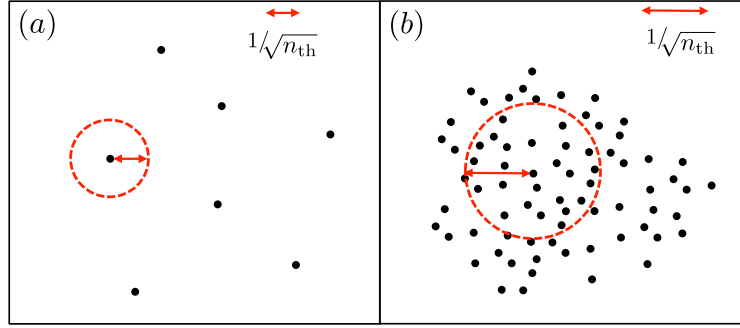
Figure 11.2:    The finite temperature frame potential measures the volume of unitary ensembles (black dots) coarse-grained over a distance $1/\sqrt{n_{\text{th}}}$ in phase space (red), where $n_{\text{th}}$ is the number of photons in the thermal density matrix (proportional to the temperature). This leads to different behavior for (a) sparsely-distributed ensembles (high temperature) vs. (b) densely-distributed ensembles (low temperature).

frame potential as a phase space volume. To do so we again require regulation by a density matrix $\rho$, which we insert into the frame potential in two distinct ways, providing phase space interpretations for each. Further discussion of CV frame potentials is contained in Section 11.4, where they are used to verify unitary designs in CV systems.

To begin, we consider the following finite-temperature frame potential [399],

$$\mathcal{F}_{\mathcal{E}}\left(\rho\right) = \mathbb{E}_{U,V \sim \mathcal{E}}\left\{\left|\operatorname{tr}\left(\rho\, U^{\dagger}V\right)\right|^{2}\right\}. \tag{11.18}$$

For the thermal density matrix, this frame potential measures the inverse volume of an ensemble 'smeared' over distances $1/\sqrt{n_{\text{th}}}$, in units of $1/n_{\text{th}}^{N}$. To see this, consider an individual term $|\operatorname{tr}\left(\tilde{\rho}_{n_{\text{th}}} U^{\dagger}V\right)|^{2}$ for displacements $U = D(\boldsymbol{\xi}_{1})$, $V = D(\boldsymbol{\xi}_{2})$. One computes

$$
\begin{aligned}
\left|\operatorname{tr}\left(\tilde{\rho}_{n_{\text{th}}} D(\boldsymbol{\xi}_{1})^{\dagger}D(\boldsymbol{\xi}_{2})\right)\right|^{2} &= e^{-(2n_{\text{th}}+1)|\boldsymbol{\xi}_{1}-\boldsymbol{\xi}_{2}|^{2}} \\
&\simeq \begin{cases} 1 & \text{if } |\boldsymbol{\xi}_{1} - \boldsymbol{\xi}_{2}| \lesssim \frac{1}{\sqrt{n_{\text{th}}}}, \\ 0 & |\boldsymbol{\xi}_{1} - \boldsymbol{\xi}_{2}| \gtrsim \frac{1}{\sqrt{n_{\text{th}}}}. \end{cases}
\end{aligned} \tag{11.19}
$$

Hence, the thermal frame potential treats two sampled displacements $U$, $V$ as identical if they are within distance $\lesssim 1/\sqrt{n_{\text{th}}}$ of each other. In the infinite-temperature limit this distance goes to zero, and one finds a correspondence to our previous definition of volume Eq. (11.17),

$$\lim_{n_{\text{th}}\to\infty} \operatorname{vol}\left(\mathcal{E}; \tilde{\rho}_{n_{\text{th}}}\right) = \lim_{n_{\text{th}}\to\infty}\left(\frac{1}{8n_{\text{th}}}\right)^{N}\frac{1}{\mathcal{F}_{\mathcal{E}}^{(1)}\left(\tilde{\rho}_{n_{\text{th}}}\right)}. \tag{11.20}$$

To further illustrate the connection between the finite-temperature frame potential and volume, we examine two example distributions, or *ensembles*, of unitaries, as depicted in

Fig. 11.2. First, consider a discrete ensemble of $M$ displacement operators distributed sparsely in phase space compared to $1/\sqrt{n_{\text{th}}}$, shown in Fig. 11.2(a). Here, the only nonzero contributions to the frame potential come from $U = V$. This occurs a fraction $1/M$ of the time, giving $\mathcal{F}_{\mathcal{E}} \approx 1/M$—intuitively, the smeared ensemble has a volume $(1/\sqrt{n_{\text{th}}})^{2N}$ per operator, which gives a volume $M$ in the units $(1/n_{\text{th}})^N$. Second, consider an ensemble of displacement operators densely distributed with density $\omega$ over some volume $\mathcal{V}$. Here, each displacement $U$ receives significant contributions from all operators within a $2N$-dimensional ball of radius $1/\sqrt{n_{\text{th}}}$ about $U$. There are $\sim \omega \left(1/\sqrt{n_{\text{th}}}\right)^{2N}$ such operators for each $U$, giving an inverse volume

$$\mathcal{F}_{\mathcal{E}} \left(\tilde{\rho}_{n_{\text{th}}}\right) \sim \frac{\omega \left(1/\sqrt{n_{\text{th}}}\right)^{2N}}{\mathcal{V}\omega} \sim \left(\frac{1}{n_{\text{th}}}\right)^N \frac{1}{\mathcal{V}}, \tag{11.21}$$

which is the inverse original volume, in the prescribed units. These two limits, sparsely and densely distributed, may equivalently be thought of as high and low temperature limits, as the length scale $1/\sqrt{n_{\text{th}}}$ by which the limits were defined is set by the inverse temperature.

The volume interpretation of the frame potential also extends to continuous ensembles of displacement operators. We present this for the particular case of a Gaussian-distributed ensemble,

$$\mathbb{D}_{\boldsymbol{\xi}_0, \boldsymbol{V}} = \left\{ D\left(\boldsymbol{\xi}\right) | \boldsymbol{\xi} \sim P_D^G\left(\boldsymbol{\xi}; \boldsymbol{\xi}_0, \boldsymbol{V}\right) \right\}, \tag{11.22}$$

where $P_D^G\left(\cdot; \boldsymbol{\xi}_0, \boldsymbol{V}\right)$ is a Gaussian distribution with mean $\boldsymbol{\xi}_0 \in R^{2N}$ and covariance matrix $\boldsymbol{V} \in R^{2N} \times R^{2N}$. The frame potential can be computed, giving

$$\mathcal{F}_{\mathbb{D}_{\boldsymbol{\xi}_0, \boldsymbol{V}}} \left(\tilde{\rho}_{n_{\text{th}}}\right) = \prod_{\ell=1}^{2N} \frac{1}{\sqrt{1 + 4\lambda_\ell \left(2n_{\text{th}} + 1\right)}}, \tag{11.23}$$

where the eigenvalues $\lambda_\ell$ of $\boldsymbol{V}$ give the squared width of the Gaussian along the $\ell^{\text{th}}$ eigenvector. An intuitive notion of volume would be the product of these widths, $\sqrt{\det \boldsymbol{V}} = \prod_\ell \sqrt{\lambda_\ell}$. At high temperatures ($n_{\text{th}}\lambda_\ell \gg 1 \, \forall \ell$, see Fig. 11.3(a)), we indeed find

$$\mathcal{F}_{\mathbb{D}_{\boldsymbol{\xi}_0, \boldsymbol{V}}} \left(\tilde{\rho}_{n_{\text{th}}}\right) \simeq \left(\frac{1}{8n_{\text{th}}}\right)^N \frac{1}{\sqrt{\det \boldsymbol{V}}}. \tag{11.24}$$

At lower temperatures this correspondence does not hold. Specifically, if $n_{\text{th}} < 1/\lambda_\ell$ for some $\ell$ (shown Fig. 11.3(b)), their contributions to the volume saturate to $O(1/n_{\text{th}}^N)$ constants. As before, this is related to the 'smearing' due to the finite phase space resolution of the low temperature thermal density matrix.

An alternative route to characterizing the volume of a CV operator lies in a 'twice-regularized' finite-temperature frame potential, which we introduce in detail in Sec. 11.4 and Appendix F. In the previous measure of volume, we inserted $\rho$ two separate times, once in the TOC and once in the finite-temperature frame potential. Here, we instead consider the
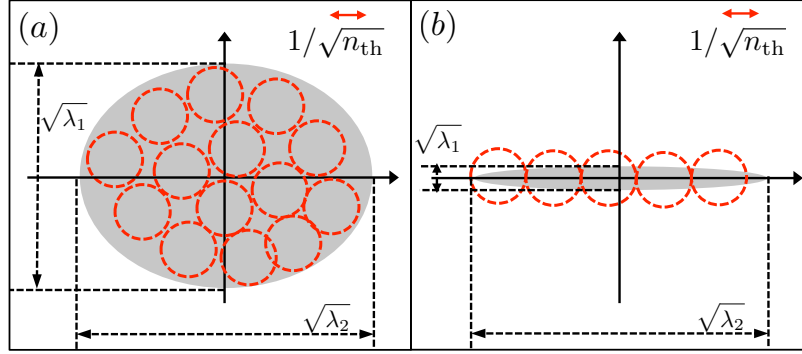
Figure 11.3:   Schematic of the volume measured by the frame potential for two continuous ensembles of displacement operators (grey). (a) For an ensemble with large widths $\lambda_1, \lambda_2 \gg 1/n_{\text{th}}$, it reproduces the traditional volume $\sqrt{\lambda_1 \lambda_2}$. (b) If the ensemble is narrow in some quadrature, say $\lambda_1 \ll 1/n_{\text{th}}$, it instead measures the 'coarse-grained' volume $\sqrt{\lambda_2/n_{\text{th}}}$. Here $n_{\text{th}}$ is the number of photons in the thermal density matrix.

ensemble of the *unnormalized* characteristic function, $\mathcal{E} = \{D(\boldsymbol{\xi}_2) \,|\, \boldsymbol{\xi}_2 \sim |\chi[\boldsymbol{\xi}; D(\boldsymbol{\xi}_1; t)]|^2\}$, and measure its volume using the twice-regularized frame potential

$$\mathcal{J}_{\mathcal{E}}^{(1)}(\rho) \sim \mathbb{E}_{U,V \sim \mathcal{E}}\left\{|\text{tr}\left(\sqrt{\rho}\, U^\dagger \sqrt{\rho}\, V\right)|^2\right\}. \tag{11.25}$$

Intuitively, the frame potential both 'smears' displacements over a width $\chi[\boldsymbol{\xi}, \rho]$, and weights operators by their preservation of the low-energy subspace, measured by $\text{tr}\left(\sqrt{\rho}\, U^\dagger \sqrt{\rho}\, U\right)$. For the thermal density matrix $\tilde{\rho}_{n_{\text{th}}}$, one computes $\sqrt{\tilde{\rho}_{n_{\text{th}}}} = \left(\frac{n'_{\text{th}}+1}{\sqrt{n_{\text{th}}+1}}\right)^N \tilde{\rho}_{n'_{\text{th}}}$ with $(n'_{\text{th}}+1)^2/n'^2_{\text{th}} \equiv (n_{\text{th}}+1)/n_{\text{th}}$, which gives

$$\text{tr}\left(\sqrt{\tilde{\rho}_{n_{\text{th}}}} D^\dagger(\boldsymbol{\xi}_1) \sqrt{\tilde{\rho}_{n_{\text{th}}}} D(\boldsymbol{\xi}_2)\right) = \left(\frac{(n'_{\text{th}}+1)^2}{(n_{\text{th}}+1)\,(2n'_{\text{th}}+1)}\right)^N$$

$$\times \exp\left(-\frac{|\boldsymbol{\xi}_1|^2 + |\boldsymbol{\xi}_2|^2}{2(2n'_{\text{th}}+1)}\right) \exp\left(-\frac{|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|^2 n'_{\text{th}}(n'_{\text{th}}+1)}{(2n'_{\text{th}}+1)}\right). \tag{11.26}$$

As expected, this contribution is significant for nearby $|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2| \lesssim 1/\sqrt{n_{\text{th}}}$ displacements that approximately preserve the subspace of $\lesssim n_{\text{th}}$ photons, $|\boldsymbol{\xi}_1|, |\boldsymbol{\xi}_2| \lesssim \sqrt{n_{\text{th}}}$.

## Operator spreading and OTOCs

As we have seen, an operators' phase space distribution can be characterized by TOCs. However, such a characterization is not always convenient, nor well-matched to physical observables. For instance, once the phase space volume of an operator becomes large, its

characterization will require the measurement of a number of TOCs comparable to the dimension of the relevant many-body Hilbert space, many of which will involve highly non-local operators.

This motivates us to turn our attention to OTOCs. In thermalizing many-body systems, the decay of OTOCs detects the spread of local operators in real space [49, 68, 198, 199, 203, 211, 213, 246, 249, 252, 340, 385, 476, 505]. In few-body CV systems, OTOCs have also found use due to their correspondence with diagnostics of classical chaos [405].

In contrast to previous work on OTOCs in CV systems, we focus on OTOCs of displacement operators. We begin by establishing a precise Fourier transform-like relation between these OTOCs and operator distributions in phase space. We find that Gaussian and non-Gaussian unitaries, previously distinguished by their ability to spread operators in phase space, also have starkly different behavior on OTOCs. Gaussian time-evolution *cannot* cause the OTOC magnitude to decay, and can only change the OTOC's phase—the decay of OTOCs thus serves as an indicator of non-Gaussian time-evolution. Extending the relation between OTOCs and operator spreading, we show that the OTOC, when averaged over an ensemble of displacement operators, measures a time-evolved operator's support in a 'ball', of designated width and center, in phase space. This can probe operator spreading both in real space, by averaging over local displacement operators, as well as within the Hilbert space of a single mode, by varying the average displacement of the ensemble. Later, in Section 11.5, we will discuss measurement schemes for both OTOCs and average OTOCs, showing that both are efficiently measurable using only Gaussian operations.

**OTOCs and operator spreading**

We begin by deriving an explicit relation between an operator's distribution in phase space and OTOCs. Again, it is helpful to first consider DV systems. OTOCs at infinite temperature are defined as

$$\mathcal{C}_2\left(P(t), R(0)\right) \equiv \frac{1}{d} \operatorname{Tr}\left[P^\dagger(t) R^\dagger(0) P(t) R(0)\right]. \tag{11.27}$$

Expanding $P(t) = \sum_Q f\left[Q; P(t)\right] Q$, we have

$$\mathcal{C}_2\left(P(t), R(0)\right) = \sum_{Q \in \text{Pauli}} \left|f\left[Q; P(t)\right]\right|^2 \mathcal{C}_2\left(Q, R\right). \tag{11.28}$$

Since Pauli operators commute up to a phase, $QR = e^{i\phi_{Q,R}} QR$, the OTOC of two Paulis $\mathcal{C}_2\left(Q, R\right) \equiv \frac{1}{d} \operatorname{Tr}\left(Q^\dagger R^\dagger Q R\right)$ is an overall phase. This allows the following the inverse transformation:

$$\frac{1}{d^2} \sum_{R \in \text{Pauli}} \mathcal{C}_2\left(P, R\right) \mathcal{C}_2\left(R, Q\right)^\star = \delta_{P,Q}, \tag{11.29}$$

which can be proven by using the orthogonality of the phase space point operators [175]. Applying this to Eqs. (11.28) we have

$$|f\,[Q;P(t)]\,|^2 = \frac{1}{d^2} \sum_{R(0)\in\text{Pauli}} \mathcal{C}_2\,(P(t),R(0))\,\mathcal{C}_2\,(R,Q)^\star. \tag{11.30}$$

Thus, the probability distribution $|f\,[Q;P(t)]\,|^2$ can be obtained from OTOCs by a transformation similar to the discrete Fourier transform. A relation akin to this one was previously derived in Ref.[399].

In analogy to the DV case, we begin by introducing the CV OTOC with respect to displacement operators,

$$\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho = \text{tr}\left[\rho D^\dagger\,(\boldsymbol{\xi}_1;t)\,D^\dagger\,(\boldsymbol{\xi}_2)\,D\,(\boldsymbol{\xi}_1;t)\,D\,(\boldsymbol{\xi}_2)\right]. \tag{11.31}$$

By using Eq. (11.8) this OTOC can be written in terms of the characteristic function $\chi\,[\boldsymbol{\xi};D\,(\boldsymbol{\xi}_1;t)]$ as

$$\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho = \frac{1}{\pi^{2N}}\int d^{2N}\boldsymbol{\xi}\,d^{2N}\boldsymbol{\xi}'\,e^{-2i(\boldsymbol{\xi}_2-\boldsymbol{\xi}/2)^T\boldsymbol{\Omega}\boldsymbol{\xi}'}$$
$$\times\chi^\star\,[\boldsymbol{\xi};D\,(\boldsymbol{\xi}_1;t)]\,\chi\,(\boldsymbol{\xi}-\boldsymbol{\xi}';\rho)\,\chi\,[\boldsymbol{\xi}';D\,(\boldsymbol{\xi}_1;t)]\,, \tag{11.32}$$

similar to Eqs. (11.28), but with the characteristic function $\chi\,(\boldsymbol{\zeta};\rho)$ of $\rho$ now acting as a transformation kernel. For the thermal density matrix in the infinite temperature limit, one has $\chi\,(\boldsymbol{\xi}-\boldsymbol{\xi}';\tilde{\rho}_{n_{\text{th}}})\propto\delta\,(\boldsymbol{\xi}-\boldsymbol{\xi}')$ and achieves Fourier-like relations between TOCs and the OTOC:

$$\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho \propto \int d^{2N}\boldsymbol{\xi}\,|\mathcal{C}_1(\boldsymbol{\xi}_1,\boldsymbol{\xi};t)_\rho|^2 e^{-2i\boldsymbol{\xi}_2^T\boldsymbol{\Omega}\boldsymbol{\xi}}, \tag{11.33}$$

$$|\mathcal{C}_1(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho|^2 \propto \int d^{2N}\boldsymbol{\xi}\,\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi};t)_\rho e^{2i\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{\xi}_2}, \tag{11.34}$$

analogous to Eqs. (11.28, 11.30). At finite temperature, the second of these becomes

$$\int d^{2N}\boldsymbol{\xi}_2\,e^{2i\boldsymbol{\xi}_2\boldsymbol{\Omega}\boldsymbol{\xi}'}\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho = \int d^{2N}\boldsymbol{\xi}\,e^{i\boldsymbol{\xi}_2^T\boldsymbol{\Omega}\boldsymbol{\xi}'}$$
$$\times\chi^\star\,[\boldsymbol{\xi};D\,(\boldsymbol{\xi}_1;t)]\,\chi\,(\boldsymbol{\xi}-\boldsymbol{\xi}';\rho)\,\chi\,[\boldsymbol{\xi}';D\,(\boldsymbol{\xi}_1;t)]\,. \tag{11.35}$$

For the thermal density matrix, this implies that the Fourier transform of OTOCs addresses probability distributions coarse-grained over a width $\sim\frac{1}{\sqrt{n_{\text{th}}}}$.

We further illustrate this relation between OTOCs and operator spreading with a brief example. Consider an operator that spreads to have width $V$ in phase space, $\chi\,[\boldsymbol{\xi};D\,(\boldsymbol{\xi}_1;t)]\sim\exp\left(|\boldsymbol{\xi}-\boldsymbol{\xi}_1|^2/\,(2V)\right)$. This leads to an OTOC $\mathcal{C}_2(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;t)_\rho\sim\exp\left(-V|\boldsymbol{\xi}_2|^2\right)$, where a larger phase space width implies greater decay of the OTOC.

**Average OTOCs**

We have seen that the decay of OTOCs detects the increase in operator volume characteristic of genuine scrambling. In this section, we demonstrate that OTOCs can also probe finer-grained aspects of operator spreading. Specifically, we show that *averages* of OTOCs over ensembles of displacement operators measure a coarse-grained distribution of a time-evolved operator in phase space. We begin by reviewing the use of average OTOCs to characterize real space operator spreading in DV systems. Following this, we turn to the simplest example of a CV average OTOC, which detects the amount that an operator has spread outside a ball of some radius about the identity. We then generalize this, and show that the center of the ball, as well as its widths in every direction of phase space, can be tuned by varying the ensemble over which the OTOC is averaged. This heuristic geometrical picture suggests a hydrodynamical interpretation of operator spreading, where changes of OTOCs in time characterize a flow of the operator's distribution through the surfaces of these balls. In Section 11.4, we develop such a description for random Gaussian circuits.

In DV systems, it is well-known that averaging the OTOC over ensembles of Pauli operators can detect operator spreading in real space. For example, consider the probability that an operator $P(t)$ has non-trivial support on the $j$-th qudit,

$$\mathcal{W}_j \equiv \sum_{Q \in \text{Pauli} \, : \, Q|_j \neq I} |f\left[Q; P(t)\right]|^2, \tag{11.36}$$

where $Q|_j$ represents the Pauli operator content at $j$-th qudit. This can be rewritten in terms of OTOCs as [213]

$$1 - \mathcal{W}_j = \frac{1}{d^2} \sum_{R \in \text{Pauli} \, : \, R = I \otimes \cdots I \otimes R_j \otimes I \cdots \otimes I} \mathcal{C}_2\left(P(t), R\right) \tag{11.37}$$

where $R$ are single qubit Pauli operators acting on $j$th qubit. This shows that the generic smallness of OTOCs, namely those averaged over single-qudit Pauli operator, implies higher weights in the operator spreading of $P(t)$.

In CV systems, averaging the OTOC over ensembles of displacement operators can probe operator spreading not only in real space, but also phase space. We define the average OTOC over an ensemble $\mathcal{E}$ of displacement operators as:

$$\overline{\mathcal{C}_2}\left(D\left(\boldsymbol{\xi}_1; t\right), \mathcal{E}\right)_\rho \equiv \mathbb{E}_{V \sim \mathcal{E}}\left[\mathcal{C}_2\left(D\left(\boldsymbol{\xi}_1; t\right), V\right)_\rho\right]. \tag{11.38}$$

There are two important regularizations in this OTOC to keep in mind: the finite extent of ensemble $\mathcal{E}$, and the density matrix $\rho$. To gain insight on their effects, we begin with a fully symmetric Gaussian ensemble of displacements,

$$\mathbb{D}_n = \left\{ D\left(\boldsymbol{\xi}\right) | \boldsymbol{\xi} \sim P_D^G\left(\boldsymbol{\xi}; n\right) \equiv \exp\left(-|\boldsymbol{\xi}|^2/n\right) / \left(\pi n\right)^N \right\}. \tag{11.39}$$
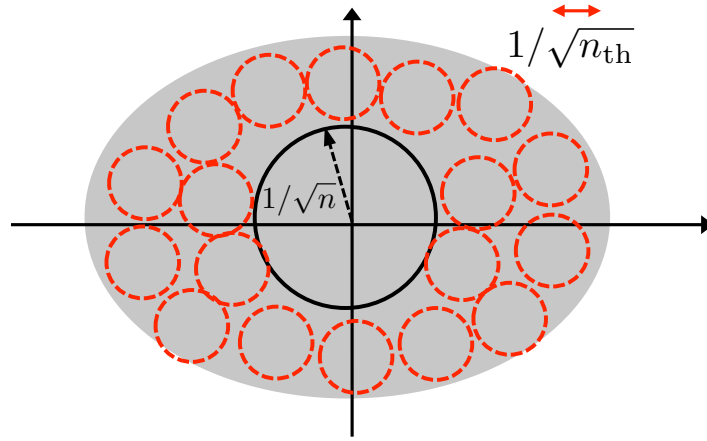
Figure 11.4:   The average OTOC measures the extent that an operator (grey) has spread outside a ball of distance $1/\sqrt{n}$ (black) in phase space, after coarse-graining over scales $1/\sqrt{n_{\rm th}}$ (red). Here $n$ is the width of the ensemble that the OTOC is averaged over, and $n_{\rm th}$ is the number of photons in the thermal density matrix.

To understand the effect of the ensemble size $n$, consider the average OTOC for an unspread displacement operator $D\left(\boldsymbol{\xi}_1;t\right) = D\left(\boldsymbol{\xi}\right)$. One finds

$$\overline{\mathcal{C}_2}\left(D\left(\boldsymbol{\xi}\right),\mathbb{D}_n\right)_\rho = \exp\left(-n|\boldsymbol{\xi}|^2\right), \tag{11.40}$$

which is small unless $|\boldsymbol{\xi}| \leq 1/\sqrt{n}$. This suggests that, for an arbitrary $D\left(\boldsymbol{\xi}_1;t\right)$ with characteristic function $\chi\left[\boldsymbol{\xi};D\left(\boldsymbol{\xi}_1;t\right)\right]$, significant contributions to the OTOC arise only from regions $|\boldsymbol{\xi}| \leq 1/\sqrt{n}$. This is confirmed by an explicit calculation of the average OTOC,

$$\overline{\mathcal{C}_2}(D\left(\boldsymbol{\xi}_1;t\right),\mathcal{E})_\rho = \frac{1}{\pi^{2N}}\int d^{2N}\boldsymbol{\xi}d^{2N}\boldsymbol{\xi}'\exp\left(-n|\boldsymbol{\xi}|^2\right)e^{i\boldsymbol{\xi}'^T\boldsymbol{\Omega}\boldsymbol{\xi}}$$
$$\times \chi^\star\left[\boldsymbol{\xi}';D\left(\boldsymbol{\xi}_1;t\right)\right]\chi\left[\boldsymbol{\xi}'-\boldsymbol{\xi};\rho\right]\chi\left[\boldsymbol{\xi};D\left(\boldsymbol{\xi}_1;t\right)\right], \tag{11.41}$$

where contributions are damped by the same exponential factor $\exp\left(-n|\boldsymbol{\xi}|^2\right)$. On the other hand, the finite phase space resolution of $\rho$ (i.e. the width of $\chi\left[\boldsymbol{\xi}'-\boldsymbol{\xi};\rho\right]$) introduces mixing between $\chi^\star\left[\boldsymbol{\xi}';D\left(\boldsymbol{\xi}_1;t\right)\right]$ and $\chi\left[\boldsymbol{\xi};D\left(\boldsymbol{\xi}_1;t\right)\right]$, which serves to coarse-grain the operator's phase space distribution on a scale set by $\rho$. For a thermal state $\tilde{\rho}_{n_{\rm th}}$, this occurs for $|\boldsymbol{\xi}'-\boldsymbol{\xi}| \leq 1/\sqrt{n_{\rm th}}$. Combining the two regularizations (Fig. 11.4), we see that the decay of the average OTOC probes the extent that the operator distribution, coarse-grained by $\rho$, is spread outside a ball of radius $\sim 1/\sqrt{n}$ about the identity ($\boldsymbol{\xi} = 0$). Intuitively, one likely wants to set the coarse-graining to be on a scale smaller than the ball radius, $1/\sqrt{n_{\rm th}} < 1/\sqrt{n}$.

From the analysis above, we see a clear distinction between volume characterization by the OTOC and the frame potential. Namely, the frame potential measures the overall volume

of the coarse-grained operator spreading, whereas the average OTOC measures the extent to which the operator has spread outside the ball of radius $1/\sqrt{n}$. Broadly, this implies that the average OTOC decay earns contributions only from operators with large displacements (compared to $1/\sqrt{n}$). This feature is intuitively favorable for characterizing scrambling and operator growth in CV systems because larger displacement operators are more complex, requiring larger energy to implement.

One can also consider the average OTOC with respect to more generic Gaussian distributions $\mathbb{D}_{\boldsymbol{\xi}_0, \boldsymbol{V}}$ defined in Eq. (11.22). These detect the portion of a coarse-grained operator's distribution within a distorted ball of width(s) determined by $\boldsymbol{V}^{-1}$ about a center $\boldsymbol{\xi}_0$. As such, in principle, the average OTOC can measure a coarse-grained local probability density anywhere in the phase space with tunable short ($\sim 1/\sqrt{n_{\mathrm{th}}}$) and long ($\sim 1/\sqrt{n}$) distance cut-offs. Such measures include probing operator spreading in real space, by choosing the widths of $\boldsymbol{V}$ to be large for modes inside some spatial region and small otherwise. Finally, we note that our focus on Gaussian ensembles is not only analytically convenient; in Sec. 11.5 we show that such average OTOCs are naturally measured in quantum optical experiments.

## Measuring non-Gaussianity with OTOCs

We have seen that Gaussian and non-Gaussian dynamics have different abilities to spread operators in phase space, termed genuine and quasi scrambling. Separately, we showed that operators' phase space distributions are related to OTOCs via Eq. (11.32) (with Eqs. (11.33-11.34) as limiting cases). Here we complete this triangle of relations, showing that decay of the OTOC *magnitude* measures non-Gaussianity and operator spreading.

Once again, we begin by constructing a measure of non-Cliffordness for DV time-evolution. Using the commutation $PQ = e^{i\phi_{P,Q}}QP$ of Pauli operators, one can show that OTOCs with respect to Paulis satisfy ($P \neq I$)

$$\begin{aligned} \left| \mathcal{C}_2(P(t), R) \right| &= 1, \quad \forall\, R \in \mathrm{Pauli} \Leftrightarrow U \text{ is Clifford}, \\ &< 1, \quad \exists\, R \in \mathrm{Pauli} \Leftrightarrow U \text{ is non-Clifford}. \end{aligned} \tag{11.42}$$

Hence, the decay of the amplitude of OTOCs is sensitive to non-Cliffordness of $U$. This is another way to see why Clifford unitaries should not be called genuine scramblers. This prompts us to consider the sum of OTOC amplitudes as a faithful measure of non-Cliffordness:

$$\begin{aligned} \mathcal{M}^{DV} &\equiv \frac{1}{d^4} \sum_{P,R \in \mathrm{Pauli}} \left| \mathcal{C}_2(P(t), R) \right|^2 \\ &= \frac{1}{d^2} \sum_{P \in \mathrm{Pauli}} \left| f\left[Q; P(t)\right] \right|^4 \\ &= \frac{1}{d^2} \sum_{P \in \mathrm{Pauli}} \mathrm{vol}(\mathcal{E}[P(t)])^{-1}, \end{aligned} \tag{11.43}$$

where decay of $\mathcal{M}^{CV}$ from 1 indicates non-Clifford behavior, and where we use Eqs. (11.13) to relate $\mathcal{M}^{CV}$ to the phase space volume of time-evolved Pauli operators. This provides an explicit relation between OTOCs, operator spreadings and non-Cliffordness.

Analogous relations can be derived for CV systems. Using the commutation $D\left(\boldsymbol{\xi}_1\right) D\left(\boldsymbol{\xi}_2\right) = e^{-2i\boldsymbol{\xi}_1{}^T \boldsymbol{\Omega}\boldsymbol{\xi}_2} D\left(\boldsymbol{\xi}_2\right) D\left(\boldsymbol{\xi}_1\right)$, the OTOC under Gaussian time-evolution can be computed exactly, giving the overall phase $\mathcal{C}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho = e^{-2i\left(\boldsymbol{S}^{-1}\boldsymbol{\xi}_1\right)^T \boldsymbol{\Omega}\boldsymbol{\xi}_2}$ with magnitude 1. One can show that ($\boldsymbol{\xi}_1 \neq \mathbf{0}$ and $\rho$ full rank)

$$\left|\mathcal{C}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho\right| = 1, \quad \forall \boldsymbol{\xi}_2 \Leftrightarrow U \text{ is Gaussian.} \tag{11.44}$$
$$< 1, \quad \exists \boldsymbol{\xi}_2 \Leftrightarrow U \text{ is non-Gaussian.} \tag{11.45}$$

This suggests the following measure of non-Gaussianity

$$\mathcal{M}^{CV} \equiv \mathbb{E}_{\boldsymbol{\xi}_1 \sim \mathcal{E}_1, \boldsymbol{\xi}_2 \sim \mathcal{E}_2} \left|\mathcal{C}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_\rho\right|^2 \tag{11.46}$$

with respect to some ensembles of displacement operators $\mathcal{E}_1, \mathcal{E}_2$, where decay of $\mathcal{M}^{CV}$ from 1 indicates non-Gaussian behavior. In Section 11.5, we present a measurement protocol of this quantity. Interestingly, our proposed quantity $\mathcal{M}^{DV}$ is closely related to a recent work on the stabilizer test by Gross, Nezami and Walter [177]; it is an interesting future problem to relate $\mathcal{M}^{CV}$ to an analogous 'Gaussianity test'.

## Examples of genuine scramblers

Having established a foundation for characterizing scrambling and operator spreading in CV systems, we now apply these tools to two specific examples of non-Gaussian Hamiltonians: an exactly-solvable cubic phase gate in Section 11.3 and the quantum chaotic Henon-Heiles potential [148, 538, 539] in Section 11.3. We show that both lead to operator spreading in phase space and the decay of OTOCs. Each Hamiltonian is of the form

$$H_{NG} = \sum_{k=1}^{N} p_k^2/2m + V\left(q_1, \cdots, q_N\right), \tag{11.47}$$

where the non-Gaussianity arises from a non-linear potential $V\left(q_1, \cdots, q_N\right)$. Such Hamiltonians encompass a variety of many-body phases, as well as the few-mode billiard systems often studied in quantum chaos. In addition to the examples given here, in Section 11.5 we describe genuine scrambling operations suitable to be realized in present-day quantum optics experiment.

### Cubic phase gate

We begin with an analytically tractable toy model: a single-mode Hamiltonian with a cubic potential $V = q^3/3!$, in the $m \to \infty$ limit [172]. Here we can solve exactly for the

time-evolved displacement operator $D(\alpha; t)$. We use the Hadamard lemma

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \cdots \equiv \tilde{B} \tag{11.48}$$

and its straightforward extension $e^A f(B) e^{-A} = f(\tilde{B})$, for functions $f(B)$ with a Taylor expansion. Taking $e^A \equiv U(t) = \exp(-i\gamma t q^3/3!)$ and $f(B) = D(\alpha; t = 0)$, we find

$$D(\alpha; t) = \exp\left[i\left(\mathrm{Im}(\alpha) q - \mathrm{Re}(\alpha) p + \gamma t \mathrm{Re}(\alpha) q^2\right)\right]. \tag{11.49}$$

To compute the OTOC, we apply the Hadamard lemma again, obtaining

$$D^\dagger(\alpha; t) D^\dagger(\beta) D(\alpha; t)$$
$$= e^{i\left(-\mathrm{Im}(\beta)q + \mathrm{Re}(\beta)p + 2\mathrm{Im}(\alpha^\star \beta) + 2\gamma t \mathrm{Re}(\alpha)\mathrm{Re}(\beta)q\right)}, \tag{11.50}$$

which gives the OTOC

$$\mathcal{C}_2(\alpha, \beta; t)_\rho = \exp(i\theta) \chi\left(2i\gamma t \mathrm{Re}(\alpha) \mathrm{Re}(\beta); \rho\right). \tag{11.51}$$

Here the phase factor is given by

$$\theta = 2\mathrm{Im}(\alpha^\star \beta) + 2\gamma t \mathrm{Re}(\alpha) \mathrm{Re}(\beta) \mathrm{Re}(\alpha + \beta), \tag{11.52}$$

and $\chi(\mathrm{Re}(\xi) + i\mathrm{Im}(\xi); \rho)$ is the characteristic function of $\rho$. For a thermal state $\tilde{\rho}_{n_{\mathrm{th}}}$, we have

$$\mathcal{C}_2(\alpha, \beta; t)_{\tilde{\rho}_{n_{\mathrm{th}}}} = e^{i\theta} e^{-2(2n_{\mathrm{th}}+1)(\mathrm{Re}(\alpha)\mathrm{Re}(\beta)t)^2 \gamma^2} \tag{11.53}$$

indicating a Gaussian decay of the OTOC in time. From this, the operator's phase space distribution (diagnosed by TOCs) can be calculated via Eq. (11.34). For large $n_{\mathrm{th}}$, we have

$$|\mathcal{C}_1(\alpha, \beta'; t)_{\tilde{\rho}_{n_{\mathrm{th}}}}|^2 \approx \delta\left(\mathrm{Re}(\alpha) - \mathrm{Re}(\beta')\right)$$
$$\times \exp\left[-\frac{\left(\mathrm{Im}(\beta') - \mathrm{Im}(\alpha) + \gamma t \mathrm{Re}(\alpha)^2\right)^2}{2(2n_{\mathrm{th}} + 1)\gamma^2 \mathrm{Re}(\alpha)^2 t^2}\right]. \tag{11.54}$$

We see that the imaginary part of $D(\alpha; t)$ spreads as a Gaussian in phase space, while the real part does not spread (a consequence of the simplicity of our Hamiltonian, which commutes with $q$). The width of the operator distribution increases linearly in $t$, according to

$$\left|\mathrm{Im}(\beta')\right| \simeq \sqrt{2(2n_{\mathrm{th}} + 1)}\left|\mathrm{Re}(\alpha)\right| t\gamma. \tag{11.55}$$
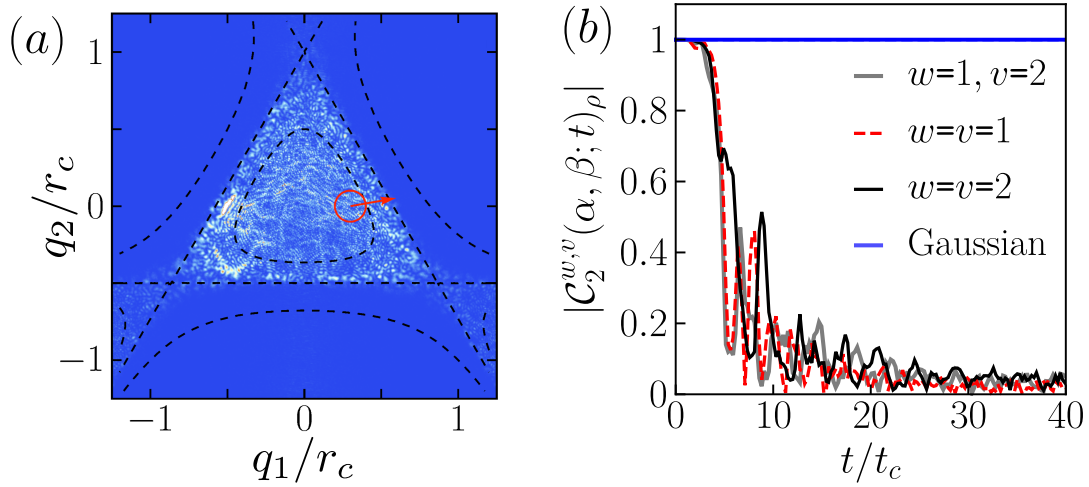
Figure 11.5: Time-evolution under the Henon-Heiles potential. (a) The initial $(t = 0)$ wave function is localized in the red circle with momentum shown by the red arrow. Coloring displays the probability density of the final state $(t = 40t_c)$, with blue indicating zero probability. Black dashed lines indicate contours $V = 2V_C, V_C, V_C/2$ of the potential, and time is measured in units of $t_c = r_C/\sqrt{2V_C/m}$. (b) The OTOC amplitude $|\mathcal{C}_2^{w,v}(\alpha, \beta; t)_{|\gamma_1\rangle|\gamma_2\rangle}|$ for a two-mode coherent state $\gamma_1 = 0.15r_C + i\sqrt{7V_C/40}\cos(10°), \gamma_2 = i\sqrt{7V_C/40}\sin(10°)$ and displacement operators $\alpha = (1-i)/4, \beta = (1+i)/4$ on modes $v, w \in \{1, 2\}$. The OTOC decay for a different coherent state $\gamma_1 = 0.125r_C + i0.375\sqrt{V_C}\cos(20°), \gamma_2 = 0.125r_C + i0.375\sqrt{V_C}\sin(20°)$ and displacements $\alpha = (1-i)/2, \beta = (0.4 + 0.3i)$ exhibits similar behavior (not shown). In contrast, the OTOC amplitude under Gaussian dynamics does not decay (blue).

## Henon-Heiles potentials

The previous example of a genuine scrambler was integrable, to enable analytic treatment. To probe operator spreading in a non-integrable system, we consider the two-mode Henon-Heiles potential

$$V(q_1, q_2) = U\left(q_1^2 + q_2^2\right)/2 + \lambda\left(q_1^2 q_2 - q_2^3/3\right), \tag{11.56}$$

which has a long history of study in both quantum and classical chaos [148, 538, 539]. We choose this potential partially with an eye on experiment: Hamiltonians that are low-order polynomials in $q, p$ are likely easier to realize in quantum optics experiment than those with hard-cut offs, such as billiard systems. This potential has a local minimum at the origin, and three saddle points at distance $r_C \equiv U/\lambda$ and energies $V_C = U^3/6\lambda^2$ [see Fig. 11.5(a)]. The classical orbits are chaotic for energies above $V_C/2$ [210]. In the quantum system, Ref. [538, 539] showed that initially local quantum states spread out in phase space, eventually approaching an equilibrium distribution.

We numerically study the Henon-Heiles potential with parameters $m = 1/2, U = 1, \lambda =$

0.025. In agreement with previous works [538, 539], we observe that an initial wavefunction localized in phase space spreads out over the entire classically allowed spatial region $V < V_C$ [Fig. 11.5(a)].  To study operator spreading, we numerically compute the OTOC for pure coherent states, $\rho = |\boldsymbol{\gamma}\rangle \langle\boldsymbol{\gamma}|$.  In practice, we expect such OTOCs to replicate much of the behavior of OTOCs with respect to the thermal density matrix, as, intuitively, all information about initial states other than their energy will be 'forgotten' at late times due to the chaotic dynamics.  Moreover, since the thermal state is a mixture of coherent states, $\tilde{\rho}_{n_{\mathrm{th}}} \sim \int d\boldsymbol{\gamma}^{2N} e^{-\beta_{n_{\mathrm{th}}}|\boldsymbol{\gamma}|^2} |\boldsymbol{\gamma}\rangle \langle\boldsymbol{\gamma}|$, thermal OTOCs can be obtained exactly from an average over coherent state OTOCs.  As shown in Fig. 11.5(b), the OTOC for a coherent state with energy $\sim V_C$ quickly decays on a timescale $\sim t_c = r_C/\sqrt{2V_C/m}$, roughly the time required to move from the origin to a saddle point for the classical Hamiltonian.  To verify the generic nature of this decay, we computed the OTOC for a different initial state of similar energy, with respect to different displacement operators, and observe similar behavior.  Finally, as a counterpoint, we also compute the same OTOCs for a Gaussian potential, $V(q_1, q_2) = U(q_1^2 + q_2^2)/2$.  As expected the OTOC amplitude remains unity, indicating no genuine scrambling.

## 11.4   Operator distributions in quasi scrambling

In this Section, we attempt to learn more about genuine scrambling dynamics by studying *quasi* scrambling systems.  Although we were able to exactly numerically simulate genuine scrambling Hamiltonians in Section 11.3, we were limited to few-mode Hamiltonians due to the exponential complexity of a many-mode Hilbert space.  In contrast, quasi scrambling (Gaussian) time-evolution can be efficiently simulated—with $N$ modes, one only needs to keep track of the $2N \times 2N$ symplectic matrix $\boldsymbol{S}$.

It is initially surprising that quasi scrambling can teach us anything about general scrambling systems, and we begin Section 11.4 by briefly elaborating on the sense in which quasi scrambling unitaries are, and are not, capable of scrambling.  Following this, we explore random circuits of Gaussian unitaries on both single (Section 11.4) and many (Section 11.4) mode systems.  Unlike DV random circuits [49, 68, 198, 246, 340, 385, 476, 505], we find that the accessible single-mode CV Hilbert space *grows exponentially* in time, related to a tunable parameter, the *squeezing*, of the Gaussian operation. This in turn leads a squeezing-dependent ballistic spreading of operators in many-mode circuits, as well as an unusual quadratic growth of entanglement entropy.  To theoretically capture these results, we introduce a hydrodynamical model of operator spreading that is accurate in the low-squeezing regime.

In DV systems, the expectation that quasi scrambling random circuits can mimic aspects of actual physical systems is justified by the notion of unitary designs.  In Section 11.4, we adapt the definition of unitary designs to CV systems.  We provide explicit results for CV 1-designs formed by displacement operators, and analogs of 2-designs formed by Gaussian unitaries.  CV designs necessarily involve 'cutting-off' an ensemble of unitaries at some finite extent, which we show can lead to ensembles that mimic design behavior on certain subspaces
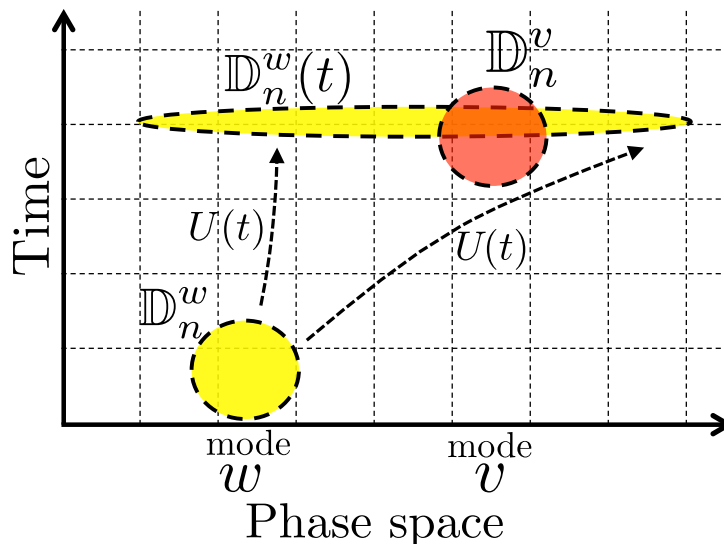
Figure 11.6: Rough illustration of quasi scrambling and the quantum Liouville's theorem. While the global volume of an ensemble $\mathbb{D}_n^w(t)$ of time-evolved displacement operators (yellow) remains fixed under time-evolution by a Gaussian unitary $U(t)$, the projected volume on the mode $v$ (measured via an average over local displacements; red), and the coarse-grained volume (roughly, the number of boxes the ensemble occupies) may increase.

of a Hilbert space, but not the entire space.

## Random Gaussian circuits

Before addressing operator spreading in random Gaussian circuits, we find it useful to discuss the extent that such systems can scramble. As we have seen, quasi scrambling unitaries cannot increase the volume of displacement operators in phase space. In Appendix F, we prove a quantum Louiville theorem that expands this non-increase to volumes of *ensembles* of displacement operators. Interestingly however, Gaussian dynamics can 'squeeze' such ensembles so that their phase space volume *appears* to increase after coarse-graining by a density matrix $\rho$, as depicted in Figs. 11.3(b), 11.6.

Hints of this scrambling power were in fact already present when we considered average OTOCs. While quasi scrambling unitaries cannot cause any individual OTOC to decay, they can randomize the OTOC phase; when averaging many OTOCs, this leads to decay just as does genuine scrambling[2]. In what follows, we will be interested in averaging over the time-evolution *itself* (i.e. the random circuits), but similar themes hold.

---

[2]Even this limited amount of scrambling is somewhat surprising, given that Gaussian time-evolution corresponds to free bosons (possessing a quadratic Hamiltonian). In contrast, free *fermions* cannot scramble operators in any sense—time-evolved local operators will always consist of only few-body terms.

**Single-mode: growth of the accessible local Hilbert space**

We begin by studying random Gaussian circuits on a single mode. In DV systems, a single-qudit circuit would seem trivial—the product of Haar random unitaries is also Haar random, and so sequential applications of them would have no interesting dynamics. In contrast, the set of CV Gaussian unitaries is unbounded, and to choose a random unitary we must cut-off this ensemble using a finite 'squeezing' parameter. The squeezing is *not* invariant under composition of Gaussians; we will study its increase in time, and this increase's effect on the size of the accessible Hilbert space.

Borrowing terminology from quantum optics, we decompose a general $N$-mode Gaussian operation into a product of 'passive linear optics' operations and 'squeezing' operations. This is known as the Euler decomposition, and takes the form

$$U_{\boldsymbol{S}} = U_{\boldsymbol{K}} U_{\boldsymbol{S}(\{r_k\})} U_{\boldsymbol{L}}. \tag{11.57}$$

The passive linear optics operations preserve photon number, and are described by symplectic orthogonal matrices $\boldsymbol{K}, \boldsymbol{L} \in SpO(N)$. Single-mode squeezing operations, which increase/decrease photon number by mixing creation and annihilation operators of each mode $k$, are characterized by their strengths $r_k$ and represented by the diagonal matrix $\boldsymbol{S}(\{r_k\}) = \bigoplus_{k=1}^{N} \mathrm{Diag}\left(e^{r_k}, e^{-r_k}\right)$. This squeezing operation multiplies a states' width in the $q_k$-quadratures by $e^{r_k}$, and in the $p_k$-quadratures by $e^{-r_k}$. (See Appendix F for a more in-depth introduction to Gaussian operations.)

In our single-mode Gaussian circuit, we consider a sequence $U_1, U_2, \ldots, U_T$ of random Gaussian unitaries. We take the passive linear optics operations $\boldsymbol{K}, \boldsymbol{L}$ of each unitary to be distributed uniformly according to the Haar measure, and the squeezing strengths $r_t$ (at time step $t$) to be small, drawn from some probability distribution $P(r_t)$. We study the growth of the total squeezing $r_T^{\mathrm{tot}}$ of the compounded unitary $U_T U_{T-1} \ldots U_1$. Averaging over angles, we find that the mean and variance of $r_T^{\mathrm{tot}}$ increase as

$$\langle r_{T+1}^{\mathrm{tot}} \rangle = \langle r_T^{\mathrm{tot}} \rangle + \frac{1}{2} \langle r_{T+1}^2 \rangle, \tag{11.58}$$

$$\mathrm{var}\left(r_{T+1}^{\mathrm{tot}}\right) = \mathrm{var}\left(r_T^{\mathrm{tot}}\right) + \frac{1}{2} \langle r_{T+1}^2 \rangle. \tag{11.59}$$

After many time steps, this approaches a normal distribution $r_T^{\mathrm{tot}} \sim \mathcal{N}\left(\mu, \sigma^2\right)$ with equal mean and variance, $\mu = \sigma^2 = \sum_{t=1}^{T} \langle r_t^2/2 \rangle = T \langle r_t^2/2 \rangle$. We verify this numerically in Fig. 11.7(a).

We can also track the time-evolution of a displacement operator $D(\alpha(t))$ under the random Gaussian circuit. At large squeezing, the amplitude of a typical evolved displacement is dominated by the displacement's component on the axis amplified by $e^{r_T^{\mathrm{tot}}}$. From our previous results, the factor $e^{r_T^{\mathrm{tot}}}$ obeys a log-normal distribution $\mathcal{LN}\left(T \langle r_t^2/2 \rangle, T \langle r_t^2/2 \rangle\right)$, and so the amplitude obeys

$$|\alpha(t)| \sim \mathcal{LN}\left(\mu = T \langle r_t^2/2 \rangle |\alpha(0)|, \sigma^2 = T \langle r_t^2/2 \rangle |\alpha(0)|^2\right). \tag{11.60}$$
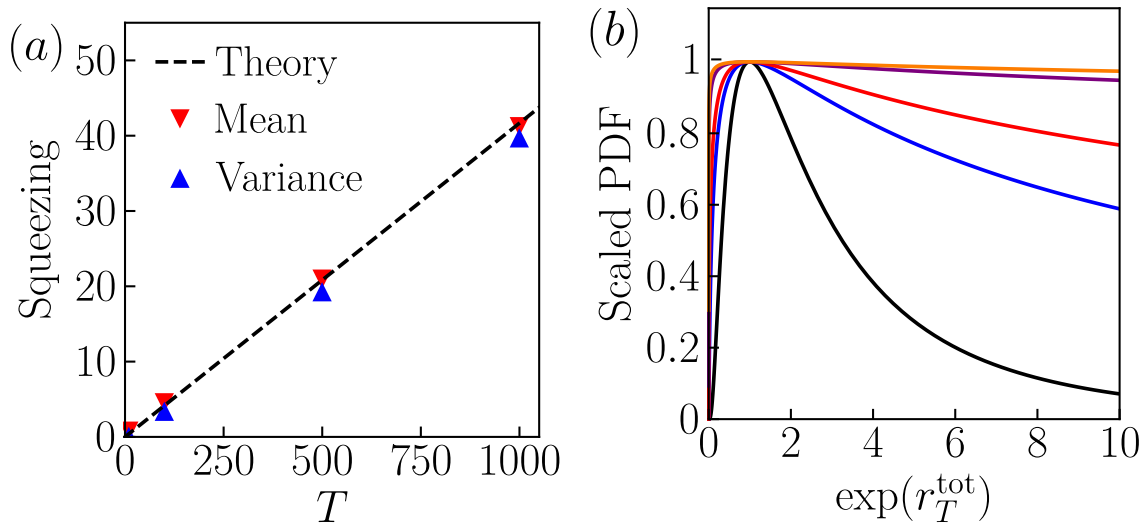
Figure 11.7: (a) Increase of the mean and variance of the total squeezing $r_T^{\text{tot}}$ for a single-mode Gaussian circuit, with individual squeezings drawn from a uniform distribution in $[0, 1/2]$ (other distributions exhibit similar behavior). Each data point is calculated from 2000 samples. At large $T$, the mean and variance are equal and close to the theoretical prediction $\frac{1}{2} \sum_{t=1}^{T} \langle r_t^2 \rangle = (1/24)T$ (black dashed line). (b) Scaled probability density function of the log-normal distribution for $e^{r_T^{\text{tot}}}$ (mean/variance $1, 5, 10, 50, 100$, from bottom to top).

This resembles a constant distribution in the range $|\alpha(0)|e^{-T\langle r_t^2/2\rangle} < |\alpha(t)| < |\alpha(0)|e^{T\langle r_t^2/2\rangle}$ [shown in Fig. 11.7(b)]. The phase space volume $\sim (e^{T\langle r_t^2/2\rangle})^2$ available to the time-evolved displacement operator increases exponentially in time.

Squeezing's effect on states parallels its effect on displacement operators. Note that, for large squeezing, the number of photons in a state typically increases as $n \to e^{2r_T^{\text{tot}}} n$ (as seen by writing $n = a^\dagger a \sim q^2 + p^2 \to e^{2r}q^2 + e^{-2r}p^2 \sim e^{2r}n$). Therefore, over different circuit realizations, the time-evolved state will have some chance to be in any of the $\sim (e^{T\langle r_t^2/2\rangle})^2 n$ states of photon number $\lesssim (e^{T\langle r_t^2/2\rangle})^2 n$; hence our claim that the size of the Hilbert space 'accessible' to the system grows exponentially.

**Many-modes: ballistic propagation**

We now turn to the effect of squeezing in many-mode random Gaussian circuits. Again owing to the Euler decomposition of Gaussian unitaries in Eqs. (11.57), we construct a general many-mode circuit by interleaving layers of single-mode squeezing and layers of multi-mode passive linear optics (i.e. beamsplitters and phase phase-shifters), as shown in Fig. 11.8. To capture the behavior of locally-interacting physical systems, we take the passive linear optics operations to be nearest-neighbor (decomposable into nearest-neighbor
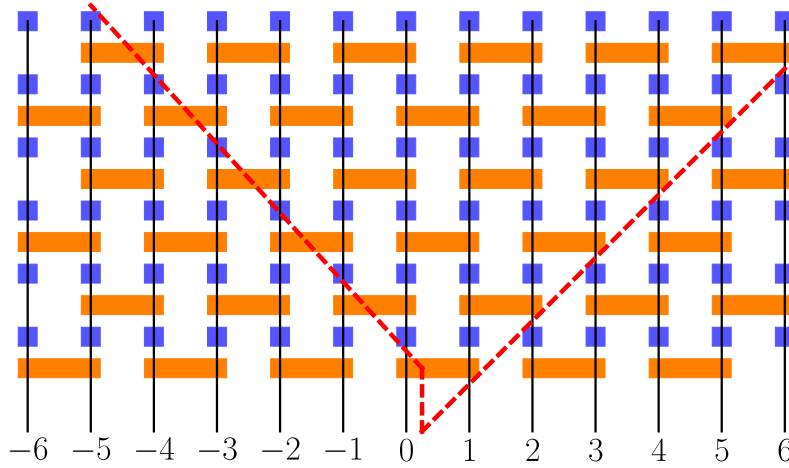
Figure 11.8: Schematic of the local random Gaussian circuit for 13 modes. The orange rectangles are random two-mode passive linear optics (beamsplitters and phase shifters) and the blue rectangles are random single-mode squeezers. The red dashed line shows the lightcone of mode 0.

beamsplitters and single-mode phase shifting operations). Each of these is described by a two-mode Gaussian unitary $U_{\boldsymbol{L}_{2x,t}}$, where $\boldsymbol{L}_{2x,t}$ is a random symplectic orthogonal matrix. Each time step also includes single-mode squeezing of amplitudes $r_{x,t}$ drawn uniformly from the interval $[0, R]$.

We characterize operator spreading in these circuits using the average OTOC. Specifically, we consider the time-evolution of an initial displacement $D^0(\alpha)$ (localized on mode 0) under an ensemble $\mathbb{C}_R$ of random Gaussian circuits. This gives rise to an ensemble of displacements $\mathbb{D}^0(\alpha; t) = \{U^\dagger(t) D^0(\alpha) U(t) | U(t) \sim \mathbb{C}_R\}$. We measure the support of this ensemble on mode $x$ with the OTOC

$$\overline{\mathcal{C}_2}\left(\mathbb{D}^0(\alpha; t), \mathbb{D}_n^x\right)_\rho = \mathbb{E}_{\boldsymbol{\xi} \sim \mathbb{D}^0(\alpha; t)}\left[\exp\left(-n\left(\boldsymbol{\xi}_{2x}^2 + \boldsymbol{\xi}_{2x+1}^2\right)\right)\right], \tag{11.61}$$

averaged over both $\mathbb{D}^0(\alpha; t)$ and a local displacement ensemble $\mathbb{D}_n^x$ of width $n$. Per Section 11.3, decay of the average OTOC indicates that significant portions of time-evolved displacements are distributed outside the ball of radius $\frac{1}{\sqrt{n}}$ about the identity on mode $x$.

For low squeezing $R$, operator spreading in the random Gaussian circuit can be captured by a simple hydrodynamical equation. The central object of this equation is the amplitude squared of the operator on the mode $x$, $f(x, t) = \boldsymbol{\xi}_{2x}^2 + \boldsymbol{\xi}_{2x+1}^2$. To motivate the hydrodynamical description, note that, in the absence of squeezing, the total amplitude $F(t) = \int dx\, f(x, t)$ is conserved due to the orthogonality of the matrix $\boldsymbol{S}$ [see Eq.(11.57)]. In this regime, random beamsplitters lead to diffusion of $f$. Introducing single-mode squeezing breaks conservation of $F$ and leads, on average, to its exponential growth. Together, these
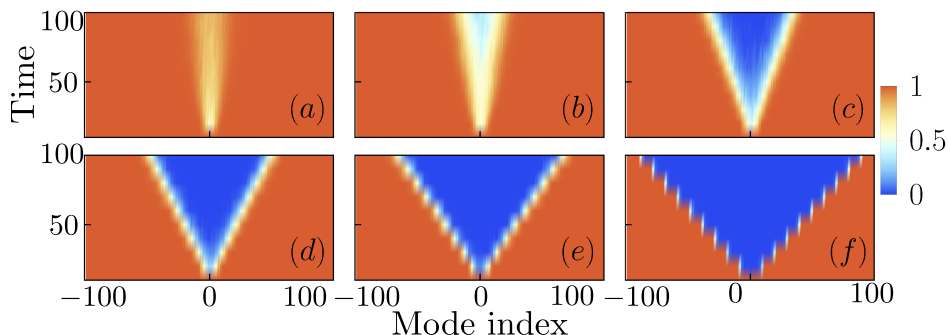
Figure 11.9: Average OTOC in the random Gaussian circuit, for a thermal density matrix $n_{\text{th}} = 5$, as a function of both space (x-axis) and time (in units of 10, y-axis). The butterfly velocity increases to its upper bound of 1 as the squeezing is increased [$R = 0, 0.2, 0.4, 0.6, 0.9, 2$, from (a)-(f)]. Each average is obtained from 100 samples.

suggest the hydrodynamical equation

$$\partial_t \overline{f}(x,t) = D\partial_x^2 \overline{f}(x,t) + c_R \overline{f}(x,t), \tag{11.62}$$

where we denote the average of $f$ over the circuit ensemble as $\overline{f}(x,t)$, and introduce the diffusion constant $D$ and the growth exponent $c_R$. Solving this equation, we find that an initially local operator $\overline{f}(x,0) = \delta(x)$ spreads according to

$$\overline{f}(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt} + c_R t\right). \tag{11.63}$$

From Eq. (11.61), the growth of $\overline{f}$ leads to decay of the OTOC, which becomes sizable when $\overline{f} \sim 1$. This decay spreads ballistically with a wavefront $x_f(t) = v_B t$, where we define the butterfly velocity

$$v_B = \sqrt{4Dc_R}. \tag{11.64}$$

Intriguingly, this relation between the butterfly velocity and the diffusion constant closely resembles that found for a coupled SYK chain [182, 505] and weakly interacting diffusive metal [367], if one identifies the single-mode growth exponent $c_R$ with the Lyapunov exponent.

We verify our hydrodynamic model numerically on a system of $2L + 1$ modes, indexed by integers from $[-L, L]$. The total operator amplitude $n$ indeed grows exponentially, with a growth exponent proportional to the squeezing $c_R \sim R$, see Fig. 11.10(a). The variance in position $\langle x^2 \rangle$ increases linearly in time, consistent with diffusive behavior at $D = 1/2$. We numerically extract the wavefront $x_f(t)$ by finding the farthest mode with average OTOC $< 0.5$. The wavefront spreads ballistically with a squeezing-dependent butterfly velocity, as shown in Figs. 11.9, 11.10(c). For small $c_R$, this velocity agrees with the hydrodynamical relation Eq. (11.64).
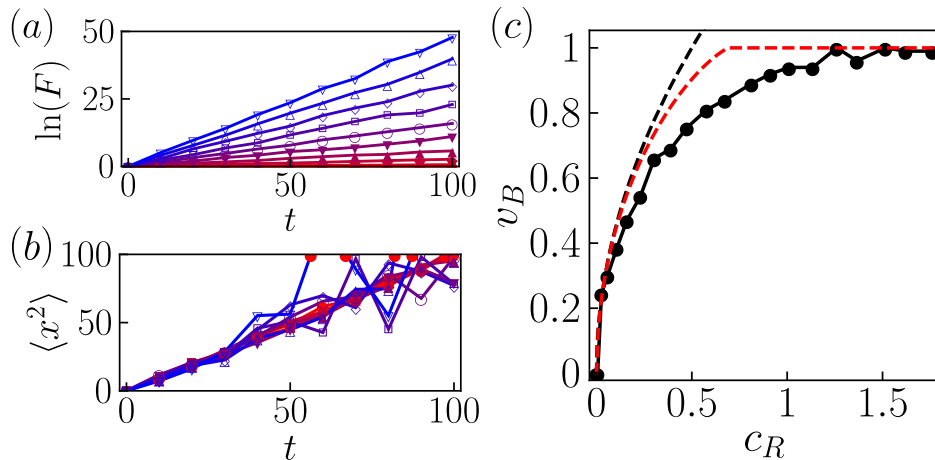
Figure 11.10: Hydrodynamical description of the many-mode random Gaussian circuit. (a) The total operator amplitude $F$ increases exponentially in time, with growth exponent proportional to the squeezing ($R = 0, 0.1, \cdots, 0.9$, from red to blue). (b) The variance in position $\langle x^2 \rangle$ grows linearly in time, indicating diffusive behavior. (c) These combine to give a ballistic spread of the OTOC decay, with a squeezing-dependent butterfly velocity $v_B$. The black dashed curve is the hydrodynamical prediction $v_B = \sqrt{4Dc_R}$, the red dashed curve is the binomial analysis, and black dots are numerics. Each average is obtained from 100 samples.

At larger squeezing $c_R \sim 1$, the butterfly velocity saturates to a maximum value of 1, and our hydrodynamical description does not apply. This maximum velocity is set by the nearest-neighbor coupling of the circuit, and we can capture this saturation by taking this discreteness into account. Note that, averaged over circuit realizations, a beam-splitter between modes $x$ and $x + 1$ acts to average the values of $f$ on each mode: $\overline{f}(x, t+1) = \overline{f}(x+1, t+1) = [\overline{f}(x, t) + \overline{f}(x+1, t)]/2$. Under this process, an initially local $f$ will spread as a binomial distribution. Combining with squeezing, we predict

$$\overline{f}(x, t) = \text{Bi}(t, x)\, e^{c_R t} = \binom{t}{\frac{x}{2} + \frac{t}{2}} \frac{1}{2^t}\, e^{c_R t}. \tag{11.65}$$

From this, we can solve for the butterfly velocity using only the approximation $t \gg 1$ [but not the further approximation $x_f(t) \ll t$, which would reproduce the Gaussian of Eq. (11.63)]. As shown in Fig. 11.10(c), this indeed more accurately captures the squeezing dependence of the butterfly velocity[3].

---

[3]We observe systematic overestimation of the velocity by theory near the point of saturation $e^{c_R} = 2$. We speculate that this arises from finite-size effects and discrepancy between the average OTOC, $\overline{\exp\left(-nf(x, t)\right)}$, and the quantity we theoretically model, $\exp\left(-n\overline{f}(x, t)\right)$. These effects can only decrease the numerical estimate of the velocity due to the strict upper bound $v_B \leq 1$.
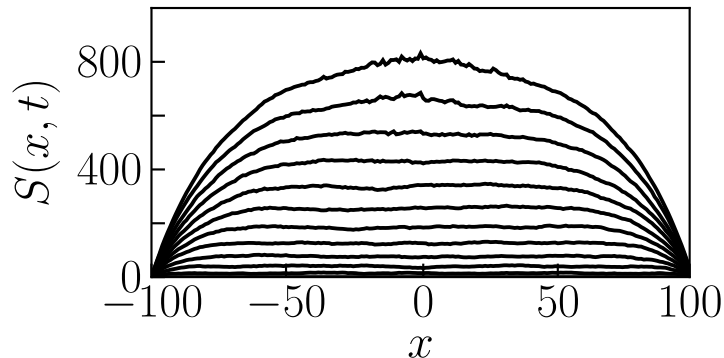
Figure 11.11: The entanglement entropy $S(x, t)$ for a random Gaussian circuit with $L = 201$ modes and squeezing $R = 0.2$ ($t = 0, 80, 160, \cdots, 800, 880$, from bottom curve to top). The entropy initially increases quadratically in time, then begins to saturate to $S(x, t) \sim t(L - |x|)$.

In addition to operator spreading, entanglement growth is a key diagnostic of scrambling in many-body systems. To study it in our model, we bipartition the system at position $x$ and calculate the entanglement entropy $S(x, t)$ between the left and right subsystems as a function of time. As shown in Fig. 11.11, for a system initially in the vacuum product state, we find that the average entanglement growth across the center cut $h(t) = \overline{S(0, t)}$ is *quadratic* in time $\sim t^2$, in contrast to the linear growth characteristic of DV systems [341]. We can understand this in terms of the exponential growth of the accessible Hilbert space found in Section 11.4. In DV systems, the hard cut-off of the local Hilbert space means that qudits near the cut quickly become maximally entangled across the cut; the linear growth $\sim t$ arises from a ballistic 'spread' of entanglement with faraway modes. In the CV case, one still receives this ballistic factor of $t$, *in addition* to a factor $t \sim \log(e^{c_R t})$ from the growth of already-entangled modes. In a finite-size system, at some time $T \sim L$ all modes will contribute to entanglement across the cut, and we expect this growth to saturate to a linear behavior $S(x, t) \sim t(L - |x|)$. This is seen in Fig. 11.11, although we are limited in system size and evolution time due to the increased ill-conditioning of the state's covariance matrix (used for efficient numerical simulation of Gaussian evolution) under squeezing.

In addition to the average entanglement, we study its fluctuations across circuit realizations, as measured by the standard deviation $w(t) = \sqrt{\overline{(S(0, t) - h(t))^2}}$. In DV systems, such fluctuations are predicted to lie in the Kardar-Parisi-Zhang (KPZ) universality class [341], scaling with time as $\sim t^{1/3}$. In contrast, in Fig. 11.12 we observe fluctuations scaling linearly with time $\sim t$. We suspect that this arises from a dominance of fluctuations in squeezing over KPZ fluctuations, but postpone a full theoretical model to future work.
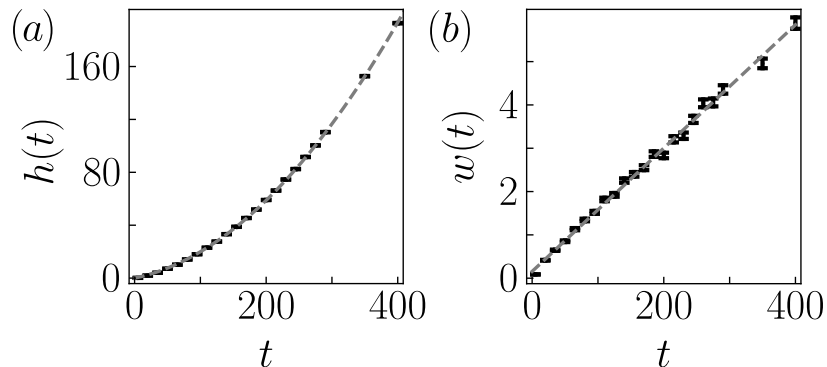
Figure 11.12: (a) Average entanglement entropy $h(t)$ across the center cut (black dots), and a quadratic polynomial fit $0.1t + 0.001t^2$ (grey line). (b) Fluctuation $w(t)$ of the entanglement entropy (black dots), and a linear polynomial fit $0.014t$ (grey line). Each point is obtained from 1000 independent samples for $R = 0.2$ and $L = 400$. Error bars indicate standard-deviations.

## CV unitary designs

In DV systems, the utility of quasi scrambling random circuit models is justified by an assumption that aspects of physical systems can be modeled by local, Haar random unitaries [20, 113, 115, 203, 361, 447]. When applicable, this assumption has incredible power—it can be rigorously shown that averages over Haar random unitaries can be replicated by much simpler unitary ensembles, known as unitary $k$-designs. By simulating the system using the simpler $k$-designs (or computing the Haar averages explicitly), one encounters a rare opportunity to study quantum chaotic behavior in an analytically and numerically tractable setting [49, 68, 198, 246, 340, 385, 476, 505].

In DV systems, a unitary $k$-design is defined as distribution of unitaries that replicates the $k^{\text{th}}$ moments of the Haar ensemble [24, 122, 402],

$$\mathbb{E}_{\mathcal{E}}\left[f(U, U^\dagger)\right] = \mathbb{E}_{\text{Haar}}\left[f(U, U^\dagger)\right] \qquad (11.66)$$

for any polynomial $f(U, U^\dagger)$ of order $\leq k$ in $U, U^\dagger$. Designs represent a hierarchy of increasingly random behavior, which has been tied to the complexity of unitary ensembles [399]. They can also inform our understanding of scrambling in systems with a notion of locality: local randomization of the system, as diagnosed by TOCs, is captured by 1-designs, while entanglement generation and operator delocalization are captured by 2-designs [399]. In DV systems, it is well-known that the set of all Pauli operators form a 1-design, and Clifford operators a 2-design.

The extension of unitary designs to CV systems is initially unclear, as there is no Haar ensemble in an infinite-dimensional system. Despite this, we notice that *averages* over Haar unitaries can remain well-defined in the CV limit. To see this, note that the Haar expectation

of $f(U, U^\dagger)$ can be computed power by power, which leads to an equivalent definition of a unitary $k$-design:

$$\mathbb{E}_{\mathcal{E}}\left[U^{\otimes k} \otimes (U^\dagger)^{\otimes k}\right] = \mathbb{E}_{\text{Haar}}\left[U^{\otimes k} \otimes (U^\dagger)^{\otimes k}\right],\qquad(11.67)$$

where the operator $U^{\otimes k} \otimes (U^\dagger)^{\otimes k}$ acts on $2k$-copies of the original system. This reformulation is convenient because the RHS can be computed explicitly; it is found to be a sum of permutation operators between the $2k$ system copies, with coefficients that depend on the dimension $d$ [399]. We use this to *define* a CV $k$-design as an ensemble which satisfies Eqs. (11.67) in the limit $d \to \infty$, keeping only lowest order terms in $1/d$.

In Sections 11.4, we use this definition to show that displacement operators form a CV 1-design. Intriguingly, previous work has indicated that Gaussian *states* cannot form a CV state 2-design [62]; in Section 11.4, we show an analogous result for Gaussian unitaries, arising from the need to regulate, or cut-off, CV unitary ensembles. For 1-designs, regularization naturally leads to ensembles that replicate 'energy-constrained' random behavior. For 2-designs, regularization relates to the squeezing of the Gaussian unitaries, and we find that scrambling by Gaussian unitaries *necessarily* increases the 'energy' (defined roughly, by the choice of regularization) of the system. Intriguingly, large-squeezing Gaussian unitaries do exhibit some similar behavior to 2-designs in DV systems.

To verify CV designs, in Section 11.4 we adapt a DV quantity which measures closeness to Haar randomness, the *frame potential*, to CV unitary ensembles. We find that the DV finite-temperature frame potential [399] has no nontrivial lower bound for CV systems, owing to the potentially infinite size of CV ensembles. To remedy this, we define a 'twice-regulated' frame potential, which only receives contributions from unitaries that approximately preserve a 'low-energy' subspace.

Although higher designs remain largely unknown in both DV [271, 484, 534] and CV systems, we note that our differentiation between quasi and genuine scrambling matches that between 2-designs and higher designs. Specifically, the volume of operators, measured by either the fourth power of the TOC or the OTOC squared, consists of four copies each of $U$ and $U^\dagger$. Its 'Haar average', corresponding to a large phase space volume of unitary time-evolved operators, is therefore replicated by $(k \geq 4)$-designs.

Before proceeding, we contrast our work with previous results on unitary designs in CV systems. Ref. [419] proposes to construct a CV 1-design by using the number and phase operators as generators; however, as they point out, their construction relies on a finite-dimension Hilbert space cut-off. Another proposal, Ref. [176], defines Haar randomness and $k$-designs in CV systems using the isomorphism between unitary matrices $U(N)$ and orthogonal symplectic matrices $SpO(2N)$, the latter describing operations in passive linear optics. This approach only encompasses Gaussian operations without squeezing, and is not suited to capture chaotic behavior in interacting CV systems.

**CV 1-design**

We begin with a brief review of 1-designs in DV systems. For $k = 1$, Eqs. (11.67) becomes

$$\mathbb{E}_{\mathcal{E}}\left[U \otimes U^{\dagger}\right] = \frac{1}{d} S_{\leftrightarrow} \tag{11.68}$$

where the swap operator $S_{\leftrightarrow}$ interchanges the copies of the system acted on by $U$ and $U^{\dagger}$. By rearranging indices, one can also define 1-designs by any of the equivalent conditions:

$$\mathbb{E}_{\mathcal{E}}\left[U \, A \, U^{\dagger}\right] = \frac{1}{d} \operatorname{tr}(A) \, \boldsymbol{I}, \tag{11.69}$$

$$\mathbb{E}_{\mathcal{E}}\left[\operatorname{tr}(U^{\dagger} A) \, U\right] = \frac{1}{d} A, \tag{11.70}$$

$$\mathbb{E}_{\mathcal{E}}\left[U \otimes U^{*}\right] = |\mathrm{EPR}\rangle\langle\mathrm{EPR}| , \tag{11.71}$$

where $A$ is an arbitrary operator on the system, and $|\mathrm{EPR}\rangle \equiv \frac{1}{\sqrt{d}}\sum_i |i\rangle |i\rangle^{*}$, with $\{|i\rangle\}$ a complete basis of states, is an EPR pair between two copies of the system. Moreover, as seen from Eqs. (11.70), all four definitions are equivalent to requiring that $\mathcal{E}$ forms a complete operator basis. Pauli operators, as a complete operator basis, thus form a DV 1-design.

The equivalence of 1-designs and complete operator bases immediately suggests a uniform ensemble of displacement operators as a candidate CV 1-design. Indeed, displacement operators satisfy similar relations

$$\left(\frac{1}{\pi}\right)^{N} \int d^{2N}\boldsymbol{\xi} D\left(\boldsymbol{\xi}\right) \otimes D^{\dagger}\left(\boldsymbol{\xi}\right) = S_{\leftrightarrow} \tag{11.72}$$

$$\left(\frac{1}{\pi}\right)^{N} \int d^{2N}\boldsymbol{\xi} \, D\left(\boldsymbol{\xi}\right) A \, D^{\dagger}\left(\boldsymbol{\xi}\right) = \operatorname{tr}(A) \, \boldsymbol{I}. \tag{11.73}$$

However, this ensemble is not normalized, due to the infinite volume of phase space. This compensates for the factors of $\frac{1}{d}$ in Eqs. (11.68-11.70), which go to zero in the CV limit. Regularizing this ensemble, in addition to being convenient, is also physically motivated. For example, the dynamics of a Hamiltonian system will be constrained by energy, and we shouldn't expect an unbounded ensemble of displacements to mimic typical scrambling behavior. To regularize this, we consider the Gaussian ensemble $\mathbb{D}_n$ of displacements, defined in Eq. (11.39). Gaussians are natural due to their stability under addition, which suggests that sequential applications of random displacements will asymptotically approach a Gaussian distribution. The ensemble $\mathbb{D}_n$ satisfies

$$\lim_{n\to\infty} n^{N}\int d^{2N}\boldsymbol{\xi} \, P_{D}^{G}\left(\boldsymbol{\xi};n\right) D\left(\boldsymbol{\xi}\right) A \, D^{\dagger}\left(\boldsymbol{\xi}\right) = \operatorname{tr}(A) \, \boldsymbol{I} \tag{11.74}$$

$$\lim_{n\to\infty} n^{N}\int d^{2N}\boldsymbol{\xi} \, P_{D}^{G}\left(\boldsymbol{\xi};n\right) D\left(\boldsymbol{\xi}\right) \otimes D^{\dagger}\left(\boldsymbol{\xi}\right) = S_{\leftrightarrow} \tag{11.75}$$

$$\lim_{n\to\infty} \int d^{2N}\boldsymbol{\xi} \, P_{D}^{G}\left(\boldsymbol{\xi};n\right) D\left(\boldsymbol{\xi}\right) \otimes D^{*}\left(\boldsymbol{\xi}\right) = |\mathrm{EPR}\rangle\langle\mathrm{EPR}| , \tag{11.76}$$

where the operator $A$ has a well-defined characteristic function, the conjugate $D^*(\boldsymbol{\xi})$ flips the signs of the displacement's momentum quadratures[4], and the CV EPR pair is defined as $|EPR\rangle = \lim_{\beta\to\infty} \sum_n e^{-\beta n} |n\rangle |n\rangle / \mathcal{N}_\beta$ ($\mathcal{N}_\beta$ chosen for normalization) [36], and the CV EPR pair is defined as $|\text{EPR}\rangle = \lim_{\beta\to\infty} \sum_n e^{-\beta n} |n\rangle |n\rangle / \mathcal{N}_\beta$ ($\mathcal{N}_\beta$ chosen for normalization) [36].

Intriguingly, at finite $n$, analogs of Eqs.(11.74-11.76) still hold for *states with mean photon number* $n_{\text{th}} < n$. For example, $n^N \int d^{2N}\boldsymbol{\xi}\, P_D^G(\boldsymbol{\xi};n) D(\boldsymbol{\xi}) \otimes D^\dagger(\boldsymbol{\xi})$ will act as the swap operator on the subspace of $\lesssim n$ photons, but not for higher photon number (we verify this in Appendix F). This can be understood intuitively. As we have seen, a thermal density matrix $\tilde{\rho}_{n_{\text{th}}}$ can resolve distances $1/\sqrt{n_{\text{th}}}$ in phase space. To swap such states, the ensemble needs to have nontrivial commutation $\sim e^{i\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{\xi}_0}$ [see, e.g. Eq. (11.9)] with displacements of this minimum distance $|\boldsymbol{\xi}_0| \sim 1/\sqrt{n_{\text{th}}}$. This occurs when $|\boldsymbol{\xi}| \gtrsim 1/|\boldsymbol{\xi}_0|$, or $\sqrt{n} \gtrsim \sqrt{n_{\text{th}}}$. We speculate that these regulated designs may arise naturally when approximating scrambling quantum dynamics. For example, time-evolution under a static Hamiltonian will generally be energetically restricted to some subspace of the total Hilbert space. Approximating scrambling behavior in such systems would require operators which respect this subspace.

The regulated CV 1-design arises naturally in physical contexts. We briefly discuss two such examples. In CV state tomography [123, 309], one aims to estimate the density matrix $\rho$ of an unknown quantum state through its characteristic function $\chi(\boldsymbol{\xi};\rho) \equiv \text{tr}[\rho D(\boldsymbol{\xi})]$. In reality, one can only perform a finite number of measurements, and can therefore only estimate $\chi(\boldsymbol{\xi};\rho)$ in a certain region of phase space. Conventionally, one chooses to sample $\boldsymbol{\xi}$ according to a Gaussian distribution, obtaining the reconstruction $\rho' \sim \int d^{2N}\boldsymbol{\xi}\, P_D^G(\boldsymbol{\xi};n)\, \text{tr}[\rho D(\boldsymbol{\xi})] D(-\boldsymbol{\xi})$, which suffers from Gaussian additive noise of strength $1/n$ [470]. A second application of CV 1-designs concerns designs for quantum *states*. In DV systems, a state 1-design is obtained by applying a unitary 1-design to a computational basis state. For CV systems, an analogous procedure gives a Gaussian distributed ensemble of coherent states. This ensemble has important applications in quantum information processing. For instance, it can be used as a basis of encoding states to achieve the classical capacity of one-mode bosonic Gaussian channels [470], or in CV quantum key distribution protocols [178].

**CV 2-design**

In DV systems, the set of Clifford unitaries forms a 2-design [399]. They obey a defining equation analogous to Eq. (11.68),

$$\mathbb{E}_{\mathcal{E}}\left[(U \otimes U) A (U^\dagger \otimes U^\dagger)\right] = \frac{1}{d^2-1}\left[\boldsymbol{I}\,\text{tr}(A) + \right.$$

---

[4]Partial transpose corresponds to the time reversal, $p \to -p, q \to q$, which leads to $D(\xi_1,\xi_2) \to D(-\xi_1,\xi_2)$, $\chi(\xi_1,\xi_2;U^T) = \chi(-\xi_1,\xi_2;U)$, and $W(q,p;U^T) = W(q,-p;U)$ [438]. On the other hand, complex conjugate corresponds to $p \to p, q \to -q$, which leads to $D(\xi_1,\xi_2) \to D(\xi_1,-\xi_2)$, $\chi(\xi_1,\xi_2;U^\star) = \chi(\xi_1,-\xi_2;U)$, and $W(q,p;U^\star) = W(-q,p;U)$. Note that $W(q,p;A) = W(-q,-p;A)$ for a Hermitian operator $A$.

$$\left. S_{\leftrightarrow}\mathrm{tr}\left(S_{\leftrightarrow}A\right) - \frac{1}{d}\boldsymbol{I}\mathrm{tr}\left(S_{\leftrightarrow}A\right) - \frac{1}{d}\boldsymbol{I}\mathrm{tr}\left(S_{\leftrightarrow}A\right)\right]. \tag{11.77}$$

It is insightful to observe the action of this quantum channel on Pauli matrices $A = P_1 \otimes P_2$:

$$\mathbb{E}_{\mathcal{E}}\left[\left(U \otimes U\right) P_1 \otimes P_2 \left(U^{\dagger} \otimes U^{\dagger}\right)\right]$$
$$= \begin{cases} \boldsymbol{I} \otimes \boldsymbol{I} & \text{if } P_1 = P_2 = \boldsymbol{I}; \\ \frac{1}{d^2-1}\sum_{P \neq \boldsymbol{I}} P \otimes P^{\dagger} & \text{if } P_1 = P_2 \neq \boldsymbol{I}; \\ 0 & \text{if } P_1 \neq P_2. \end{cases} \tag{11.78}$$

Intuitively, a random Clifford unitary transforms a non-identity Pauli matrix to any other non-identity Pauli with equal probability (along with the constraint $P_1 = P_2$).

In the CV case, we take the $d \to \infty$ limit of Eq. (11.77) as the definition of a 2-design[5],

$$\mathbb{E}_{\mathcal{E}}\left[\left(U \otimes U\right) A \left(U^{\dagger} \otimes U^{\dagger}\right)\right]$$
$$\xrightarrow[d \to \infty]{} \frac{1}{d^2}\left[\boldsymbol{I}\mathrm{tr}\left(A\right) + S_{\leftrightarrow}\mathrm{tr}\left(S_{\leftrightarrow}A\right)\right]. \tag{11.80}$$

We begin by exploring a particular ensemble $\mathcal{E}_r$ of Gaussian unitaries that comes close to satisfying this definition. To define this ensemble, it is helpful to decompose a given Gaussian into the product of a quadratic operation and a displacement, $U = D(\boldsymbol{d}/2)U_{\boldsymbol{S}}$, and Euler decompose the former as $\boldsymbol{S} = \boldsymbol{K}\boldsymbol{S}(\{r_i\})\boldsymbol{L}$, with single-mode squeezings of strength $r_i$ (see Section 11.4). When acted on a displacement operator $A = D(\xi_1) \otimes D(\xi_2)$, Eq. (11.80) leads to a sum over a distribution of transformed displacements $D(\boldsymbol{S}\xi_1) \otimes D(\boldsymbol{S}\xi_2)$, similar to Eqs. (11.77-11.78). A Gaussian distribution over $\boldsymbol{d}$ gives this distribution a factor of $\delta(\xi_1 - \xi_2)$, analogous to the DV requirement $P_1 = P_2$. The distribution is rotationally symmetric if $\mathbf{K}, \mathbf{L}$ are Haar distributed. Finally, we take the squeezings $r_i = r$ to be large. In the many-mode limit, one can show (via the Central Limit theorem) that this gives displacements approximately Gaussian distributed, with a width $e^r|\xi_1|$ proportional to the initial displacement $\xi_{\mathbf{1}}$. Concretely, we find

$$\mathbb{E}_{\mathcal{E}_r}\left[\left(U^{\otimes 2}\right) D(\xi_1) \otimes D(\xi_2) \left(U^{\dagger \otimes 2}\right)\right]$$
$$\approx \delta(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2) \int d^{2N}\boldsymbol{\xi}\, P_D^G\left(\boldsymbol{\xi}; e^r|\boldsymbol{\xi}_1|\right) \left[D\left(\boldsymbol{\xi}\right) \otimes D^{\dagger}\left(\boldsymbol{\xi}\right)\right], \tag{11.81}$$

a CV analog of Eq. (11.78).

The relation Eq. (11.81) is close to the 2-design definition Eq. (11.80). When $\xi_1 = 0$, both have a RHS proportional to the identity. When $\xi_1 \neq 0$, the RHS of Eq. (11.81)

---

[5]We note that a similar CV definition of higher $k$-designs is possible. Using expressions in Ref. [399], one can derive

$$\mathbb{E}_{\mathcal{E}}\left[\left(U\right)^{\otimes k} A \left(U^{\dagger}\right)^{\otimes k}\right] \xrightarrow[d \to \infty]{} \frac{1}{d^k}\sum_{\pi} W_{\pi}\mathrm{tr}\left(W_{\pi}^{-1}A\right), \tag{11.79}$$

where $W_{\pi}$ performs the permutation $\pi$ on the $k$ system copies.

is proportional to a SWAP operator on subspaces with less than $\sim e^r |\boldsymbol{\xi}_1|$ photons (see our discussion in Section 11.4). However, the constant of proportionality for the latter is off by a displacement-dependent factor of $1/(e^r |\xi_1|)^N$. This arises from the fact that, wherever a displacement $D(\xi)$ is transformed to, the displacement $D(c\,\xi)$ is transformed to a displacement $c$ times as large (because Gaussian unitaries act linearly on $\xi$). The width of the transformed displacements' distribution must then be proportional to $|\xi_1|$, leading to the prefactor $1/|\xi_1|^N$. This prefactor is similar to that found in Ref. [62], when considering an ensemble of single-mode Gaussian states.

It is unclear how physically fundamental this discrepancy is. Arguing for its physicality is that it seems as if it would arise in any soft regularization of CV systems. For instance, when regularized by the thermal density matrix $\tilde{\rho}_{n_{\text{th}}}$, information about the state is contained in displacements as small ('low-energy') as $\xi \sim 1/\sqrt{n_{\text{th}}}$ and as large ('high-energy') as $\xi \sim \sqrt{n_{\text{th}}}$. If the lowest-energy operator is evolved to a high-energy operator by some elements of the 2-design, then the product of many low-energy operators (which form a larger displacement, i.e. a higher energy operator) must be taken to an operator of even higher energy, outside of the subspace defined by $\tilde{\rho}_{n_{\text{th}}}$. This seems to suggest that it is necessary for a 2-design to increase the energy/dimension of the effective CV Hilbert space, leading to the 'energy-dependent' prefactor $1/|\xi_1|^N$. On the other hand, this restriction may be unique to the linear action of Gaussian unitaries on displacements, and it is an open question whether an ensemble of non-Gaussian unitaries can satisfy Eq. (11.80).

Nevertheless, 2-designs are interesting in many-body physics because they can model physical processes, and the ensemble $\mathcal{E}_r$ *does* possess these more qualitative properties. For instance, a key characteristic of a 2-design is the ability to generate entanglement. In DV systems, a typical Clifford unitary applied to a product state will generate near maximal entanglement between any two subsystems. In CV systems, we numerically investigate the entanglement generated by the ensemble $\mathcal{E}_r$ by applying a randomly sampled Gaussian on an initially unentangled two-mode system in a pure state[6]. We find that the entanglement generated grows linearly with the squeezing $r$, as shown in Fig. 11.13. This is expected, as entanglement is proportional to the logarithm of the number of states accessible to each subsystem, which increases $\sim e^{2r}$ under squeezing.

We also find that time-evolution in the many-mode random Gaussian circuit in Section 11.4 converges to $\mathcal{E}_r$ (with asymptotically increasing squeezing) at long times, demonstrating that $\mathcal{E}_r$ can capture operator growth and OTOC decay as observed in those circuits. The displacement amplitudes can be seen to be Gaussian distributed locally due to the beam-splitters, with typical width $e^{c_R t}/\sqrt{2L+1}$ as $t \to \infty$. The single-mode circuit in Section 11.4 does *not* asymptotically converge to $\mathcal{E}_r$, although this is less surprising since our arguments for $\mathcal{E}_r$ were valid only in the many-mode limit. Instead, we observe that a displacement $D(\alpha(t))$ is time-evolved such that its *amplitude* obeys a uniform distribution, which gives a

---

[6]We note that sampling from $\mathcal{E}_r$ is efficient [106, 389, 408]. The most difficult component is sampling the $N-$mode random unitary from passive linear optics $\mathbb{B}$. This can be accomplished by arranging $N(N-1)/2$ number of beam splitters in a depth $N$ optical circuit [106].
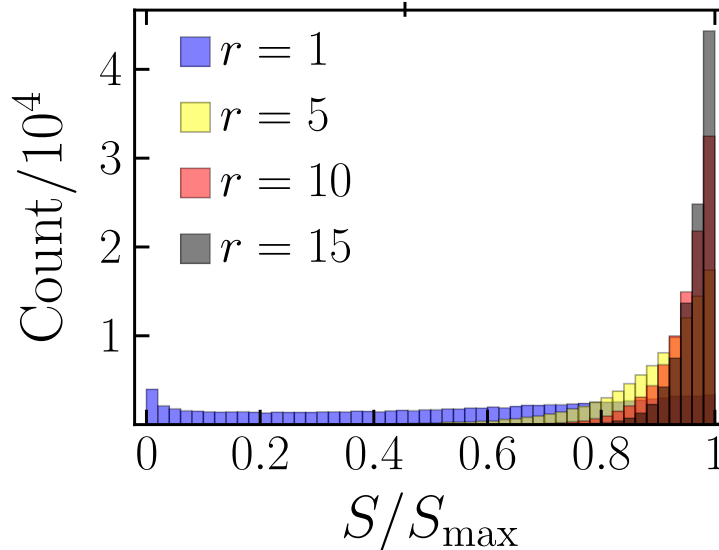
Figure 11.13: Distribution of entanglement entropy $S$ after applying random unitaries drawn from $\mathcal{E}_r$ on the two-mode vacuum state, for various squeezing strengths $r$, scaled by its (squeezing-dependent) maximum $S_{\max}$. Each distribution contains $10^4$ sampled unitaries.

distribution $P(\alpha) \sim 1/|\alpha|$ for the displacement.

**Finite temperature frame potential**

In DV systems, the closeness of a unitary ensemble to a $k$-design is measured by the $k^{\text{th}}$ frame potential [399],

$$\mathcal{F}_\mathcal{E} \equiv \mathbb{E}_{U,V \sim \mathcal{E}} \left\{ \left| \text{tr} \left( U^\dagger V \right) \right|^{2k} \right\}. \tag{11.82}$$

The frame potential is 1 for a trivial ensemble, and decays to a minimum $k!/d^{2k}$ when $\mathcal{E}$ is a $k$-design. A 'finite-temperature' generalization of the frame potential also exists [399], which takes the form

$$\mathcal{F}_\mathcal{E}^{(k)} (\rho) = \mathbb{E}_{U,V \in \mathcal{E}} \left\{ \left| \text{tr} \left( (\rho)^{\frac{1}{k}} U^\dagger V \right) \right|^{2k} \right\}, \tag{11.83}$$

and decays from 1 to a $\rho$-dependent constant $\sim 1/d^{2k}$.

We begin our discussion of CV frame potentials by demonstrating the obstacles encountered when applying DV frame potentials to CV unitary ensembles. A naive application of Eq. (11.82) to CV unitaries, say displacements $U = D(\boldsymbol{\xi}_U), V = D(\boldsymbol{\xi}_V)$, is ill-defined due to traces diverging like $\sim [\delta(\boldsymbol{\xi}_U - \boldsymbol{\xi}_V)]^{2k}$. When discussing operator volumes in Section 11.3, we regulated these divergences with a density matrix $\rho$, via the finite-temperature frame potential Eq. (11.83). Unfortunately, since the frame potential is only lower-bounded by 0 in the CV case, it is not clear how to use its decay to make sharp statements about an

ensemble's validity as a $k$-design. For instance, for our candidate 1-design $\mathbb{D}_n$ we find

$$\mathcal{F}_{\mathbb{D}_n}^{(k)}\left(\tilde{\rho}_{n_{\mathrm{th}}}\right) \sim \left(1/k^2 n n_{\mathrm{th}}\right)^N, \tag{11.84}$$

which indeed approaches 0 as $n \to \infty$. However, ensembles which are *not* 1-designs lead to similar decay. Consider two such ensembles: the sphere of displacements with fixed amplitude $\tilde{\mathbb{D}}_n = \left\{\otimes_{k=1}^N D^k\left(\sqrt{n}e^{i\theta_k}\right)|\theta_k \sim [0, 2\pi]\right\}$, and the ensemble of phase shifts $\mathbb{R} = \left\{\otimes_{j=1}^N U_{\boldsymbol{R}(\theta_j)}|\theta_j \sim [0, 2\pi]\right\}$. We find

$$\mathcal{F}_{\tilde{\mathbb{D}}_n}^{(k)}\left(\tilde{\rho}_{n_{\mathrm{th}}}\right) \sim \left(1/k\sqrt{n n_{\mathrm{th}}}\right)^N \tag{11.85}$$

$$\mathcal{F}_{\mathbb{R}}^{(k)}\left(\tilde{\rho}_{n_{\mathrm{th}}}\right) \sim \left(1/k n_{\mathrm{th}}\right)^N. \tag{11.86}$$

In particular, the frame potential for $\tilde{\mathbb{D}}_n$ decays to 0 as $n \to \infty$, just as it does for the 1-design $\mathbb{D}_n$. For $k = 1$, this can be understood through Section 11.3. The finite-temperature frame potential measures the ensemble's inverse volume in phase space—it will decay to 0 as long as this volume increases to infinity, regardless of whether the ensemble approaches a 1-design. This suggests that we may be asking the wrong question: to be a 1-design on the entire Hilbert space the ensemble must have infinite phase space volume, and accurately capturing this volume will necessarily be difficult.

Motivated by our regularization of 1- and 2-designs, we instead seek to characterize whether a unitary ensemble can form a design *on a particular, 'low-energy' subspace*, defined by some density matrix $\rho$. To do so, we introduce a new 'twice-regulated' frame potential that weights unitaries based on their preservation of this subspace. This takes the form (see Appendix F),

$$\mathcal{J}_{\mathcal{E}}^{(k)}\left(\rho\right) = \frac{\int_{\mathcal{E}} dU \int_{\mathcal{E}} dV \left| \mathrm{Tr}\left\{U^\dagger \rho^{\frac{1}{2k}} V \rho^{\frac{1}{2k}}\right\}\right|^{2k}}{\left[k! \int_{\mathcal{E}} dU \left| \mathrm{Tr}\left\{U^\dagger \rho^{\frac{1}{k}} U \rho^{\frac{1}{k}}\right\}\right|^k\right]^2}$$

$$\geq \frac{1}{\mathcal{H}^{(k)}(\{\rho_i\})}. \tag{11.87}$$

It has no upper bound, but decays to a strict nonzero lower bound determined by the entropies of $\rho$, $\mathcal{H}^{(k)}(\{\rho_i\}) \approx k! \, \mathrm{Tr}\left\{\rho^{\frac{1}{k}}\right\}^{2k}$. As we saw in Section 11.3, at $k = 1$, this frame potential measures the inverse coarse-grained volume of displacements obeying $|\boldsymbol{\xi}| < \sqrt{n_{\mathrm{th}}}$. This volume is upper bounded by $\sim n_{\mathrm{th}}^N$, leading to the nonzero lower bound of the frame potential. In Appendix F, we show that the frame potential measures an operator distance between the LHS and RHS of Eq. (11.67), with respect to the density matrix $\rho$.

We can use the twice-regulated frame potential to verify our regularized 1-design. We find:

$$\mathcal{J}_{\mathbb{D}_n}^{(1)}\left(\tilde{\rho}_{n_{\mathrm{th}}}\right) \approx \left(\frac{(1 + 2\left(n_{\mathrm{th}}/n\right))^2}{1 + 4\left(n_{\mathrm{th}}/n\right)}\right)^N \geq 1, \tag{11.88}$$

which indeed decays to its lower bound for $n \gg n_{\mathrm{th}}$. We postpone evaluation of the $k = 2$ frame potential for the large-squeezing Gaussian ensemble $\mathcal{E}_r$ to future work.

# 11.5 Experimental verifying scrambling

Here, we consider the experimental detection of scrambling in CV systems. We begin in Section 11.5 with the implementation of the scrambling dynamics themselves, providing concrete, precisely-controllable schemes to realize genuine scrambling dynamics using the so-called SNAP gate in a cavity-QED architecture. We numerically simulate these dynamics when possible, and mention open questions which may be addressed by experiment. Turning towards detection, in Section 11.5 we introduce concise measurement schemes for TOCs, individual OTOCs, and average OTOCs. These schemes rely only on Gaussian operations, as well as the ability to experimentally implement the (possibly non-Gaussian) scrambling operations $U(t)$ and $U^\dagger(t)$. Unfortunately, all of these schemes are prone to confusing scrambling with decoherence and experimental error, a problem well-known in DV systems. We address this in Section 11.5 by introducing a robust teleportation-based measurement scheme, adapted from that in Ref. [518].

## Experimental realization of scramblers

In this Section, we present concrete proposals for the experimental realization of CV genuine scrambling dynamics. While we can numerically simulate these models in single- and few-mode systems, many questions arise that are beyond the scope of exact numerics, and are ripe for experimental input. For instance, the implementation of genuine scrambling random circuits could probe the accuracy of our conjecture in Section 11.4, that quasi scrambling circuits can replicate aspects of OTOC decay and operate spreading in genuine scrambling. Additionally, our proposal will also prove apt for probing scrambling in *number-conserving* CV systems, which may behave qualitatively different from the non-conserving models of Section 11.4.

Measuring scrambling behavior (as discussed in the following Sections 11.5 and 11.5) will necessarily entail the precise implementation of not only the scrambling unitary $U$, but also either its inverse $U^\dagger$ or its conjugate $U^*$, requiring a high amount of experimental control. Additionally, realizing genuine scrambling requires strong non-Gaussian operations. Candidate experimental platforms for detecting CV scrambling should feature both of these properties, and might include non-linear crystals, cavities [208, 232, 475], and optical Floquet systems [283].

For concreteness, we focus on cavity-QED architectures. All Gaussian operations (displacements, beamsplitters, and squeezing operations) can be implemented in these systems [232]. Furthermore, non-Gaussian effects are typically much stronger than in other platforms. We focus on a particular non-Gaussian gate that has already been implemented, the so-called Selective Number-dependent Arbitrary Phase (SNAP) gate [208, 268]. Diagonal in the photon number basis, it takes the form

$$S_N\left(\{\theta_n\}\right) = \sum_{n=0}^{\infty} e^{i\theta_n} |n\rangle \langle n|, \tag{11.89}$$
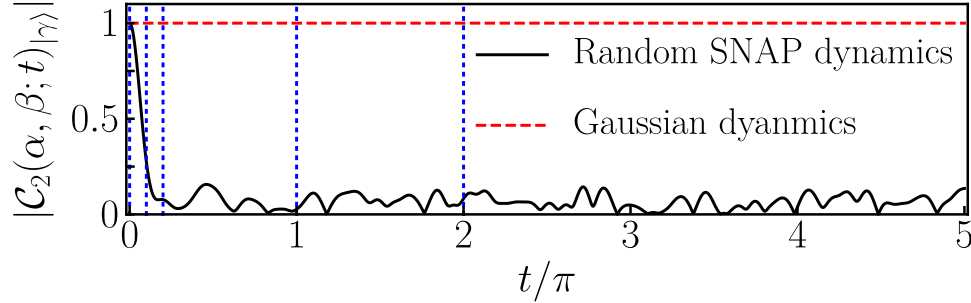
Figure 11.14: OTOC amplitude $|\mathcal{C}_2(\alpha, \beta; t)_{|\gamma\rangle}|$ for a time-dependent single-mode random SNAP gate $S_N(t)$ (black), for $\alpha = 2 + 2i, \beta = 2 - 2i, \gamma = 8$. Blue lines indicate the times of the TOC snapshots in Fig. 11.15. For contrast, the OTOC amplitude does not decay under Gaussian dynamics (red).

where the phase $\theta_n \in [0, 2\pi)$ of the $n$-photon number state can in principle be controlled arbitrarily. The SNAP gate can be experimentally realized using a cavity coupled to a transmon qubit, with the Hamiltonian

$$H = \omega_c a^\dagger a + \omega_q |e\rangle\langle e|e + \chi a^\dagger a |e\rangle\langle e|e + \Omega(t) e^{iw_q t} |e\rangle \langle g| , \qquad (11.90)$$

where $a$ is the annihilation operator of the cavity and $|g\rangle$ ($|e\rangle$) is the ground (excited) state of the qubit. When the qubit frequency shift $\chi$ is larger than both the qubit and cavity transition line-widths $\omega_q, \omega_c$, and the drive $\Omega(t) = \Omega_n(t) e^{-in\chi t}$ is weak, one can apply a phase selectively to the $n$-photon state $|n\rangle\langle n|n$, i.e. $|g, n\rangle \rightarrow e^{i\theta_n} |g, n\rangle$, while keeping all other states invariant. A drive composed of multiple frequencies $n$ allows the independent implementation of multiple such phases, realizing the SNAP gate. In experiments reported by Ref. [208], phases for up to $n = 11$ are precisely controlled. With a cavity lifetime $\sim 50\,\mu s$, qubit relaxation times $\sim 20 - 30\,\mu s$, and $\chi \sim$ MHz, this allowed the implementation of up to 15 SNAP gates with fidelity $> 0.96$ per gate.

The SNAP gate can simulate a variety of effective cavity Hamiltonians. For example, the Kerr nonlinearity $K_R(t) \equiv \exp(-itH^2)$ with $H^2 = (p^2 + q^2)^2$ is realized via a single SNAP gate with time-dependent phases $\theta_n = -t(4n + 2)^2$. In fact, when combined with Gaussian operations, the SNAP gate is universal for the realization of all Hamiltonians polynomial in quadrature operators [302]. While in principle this allows one to realize the Henon-Heiles potential and the cubic phase gate of Section 11.3, such Hamiltonians generically require long sequences of fast SNAP gates, which may prove less feasible for experiment.

With near-term experiments in mind, we introduce several genuine scramblers composed of only a moderate number of SNAP gates. We begin by studying operator spreading in the phase space of a single-mode, under a single time-dependent SNAP gate. As an example, we consider the random SNAP gate $S_N(t) = \sum_{n=0}^{\infty} e^{iw_n t} |n\rangle \langle n|$, with 'energies' $w_n$ distributed
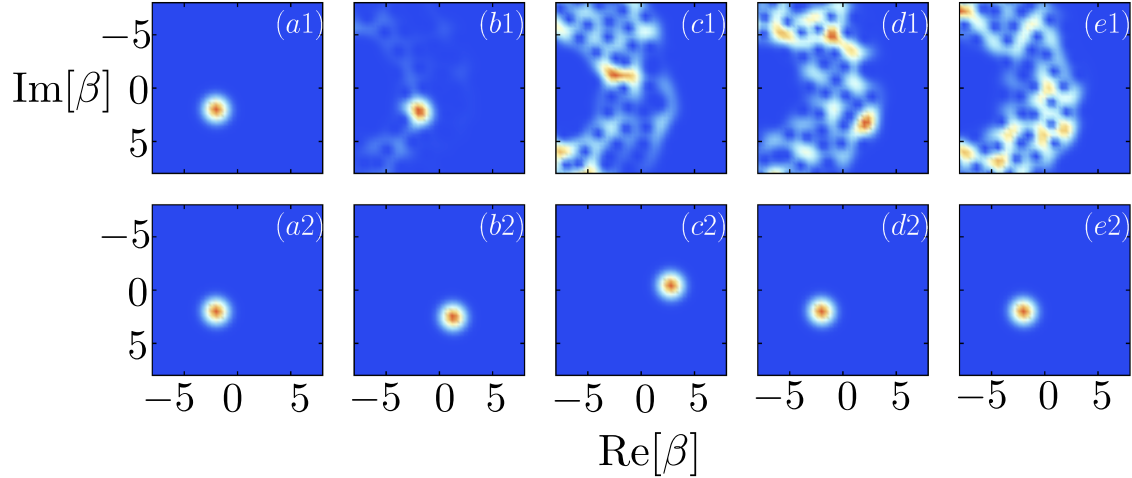
Figure 11.15: (a1)-(e1) Snapshots of the phase space distributions of a displacement operator $\alpha = 2 + 2i$ under the time-dependent random SNAP gate $S_N(t)$, as measured by the TOC $|\mathcal{C}_1(\alpha, \beta; t)_{|\gamma\rangle}|^2$, with $\gamma = 8$ (blue indicates zero TOC). Time increases from left to right $[t/\pi = 0, 0.1, 0.2, 1, 2, 5,$ from (a1)-(e1); see Fig. 11.14]. (a2)-(e2) For contrast, the distribution remains localized under Gaussian dynamics ($H^1 = p^2 + q^2$). All plots share the same x- and y-axes, the real and imaginary parts of $\beta$, respectively.

uniformly in $[0, 2\pi)^7$. Typical of genuine scrambling behavior, we find numerically that OTOCs decay to small values in $O(1/2\pi)$ time and remain small afterwards (see Fig. 11.14). To visualize the operator spreading responsible for this decay, one can calculate and plot the square of the TOC $|\mathcal{C}_1(\alpha, \beta; t)_{|\gamma\rangle}|^2$ [see Eq. (11.15)], shown in Fig. 11.15(a). As anticipated, the TOC spreads to occupy a larger phase space volume on a time scale similar to that of the OTOC decay. For contrast, under time-evolution by a *Gaussian* Hamiltonian $H^1 = p^2 + q^2$, the OTOC amplitude is fixed at unity, and the TOC remains localized in phase space [Fig. 11.15(b)].

Moving forward, we consider the use of SNAP gates to study genuine scrambling in multi-mode systems. Such behavior is tremendously difficult to numerically simulate due to the exponential size of the multi-mode Hilbert space, and would benefit greatly from experimental input. As a first example, the inclusion of SNAP gates in random circuit models like those in Section 11.4 would break the Gaussianity of the dynamics, and allow one to study operator spreading and entanglement formation of generic locally interacting CV systems. In addition, SNAP gates conserve photon number, and are thus particularly well-suited for probing CV scrambling in the presence of conservation laws. To this end, one might consider a random circuit of only passive linear optics and SNAP gates (no squeezing), organized

---

[7]We choose a random SNAP gate, instead of the better-known Kerr gate, due to a periodicity revival in the Kerr gate: $K_R(\pi/16) = \exp(-i\pi/4)\, \boldsymbol{I}$.
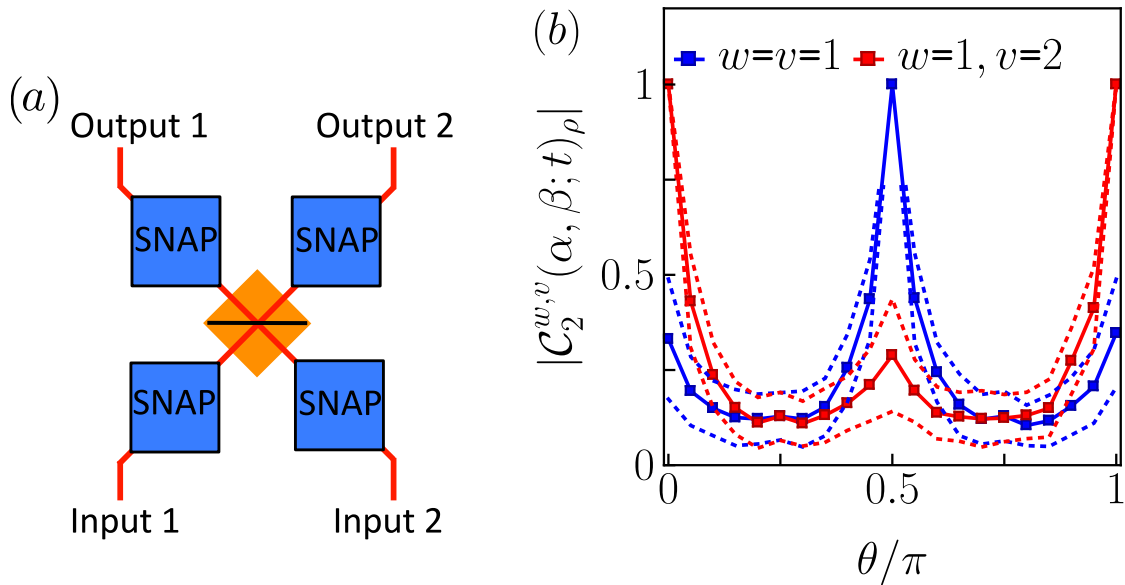
Figure 11.16: (a) Simple two-mode model of a SNAP gate random circuit. The center block represents a random beamsplitter. (b) The average OTOC amplitude ($\alpha = (1 + i)/2, \beta = (1 - i)/2, \gamma_1 = \gamma_2 = 2$) under after the circuit, for both single-mode scrambling ($w = v = 1$) and multi-mode scrambling ($w = 1, v = 2$), as a function of the beamsplitter transmissivity $\cos^2 \theta$. Dashed curves indicate standard deviations. Each data point is averaged over 100 samples.

similarly to Fig. 11.8, with SNAP gates replacing single-mode squeezing operations. It would be interesting to observe these circuits' entanglement growth in time: since number conservation seems to forbid the squeezing-induced Hilbert space growth of Section 11.4, one might guess that conservation laws cause the system to saturate to DV-like behavior $S \sim t$, in contrast to our previous result $S \sim t^2$.

For instructive purposes, we numerically simulate the simplest of such multi-mode, number-conserving circuits: a single-layer circuit on a two-mode system, as shown in Fig. 11.16(a). The circuit consists of a beamsplitter with transmissivity $\cos^2 \theta$ sandwiched between four random SNAP gates, each with phases iid uniformly in $[0, 2\pi)$. To characterize operator spreading, we compute the amplitude of the OTOC with respect to single-mode displacement operators, averaged over the random SNAP gates. As expected the OTOC strongly depends on $\theta$, which controls the mixing of the two modes [see Fig. 11.16(b)]. At $\theta = 0$, there is no mixing between modes, and so $|\mathcal{C}_2^{1,2}(\alpha, \beta; t)| = 1$. At $\theta = \pi/2$, mode 1 and mode 2 are swapped, giving $|\mathcal{C}_2^{1,1}(\alpha, \beta; t)| = 1$. The mixing is maximized at $\theta = \pi/4$. Somewhat surprisingly, this value maximizes not only two-mode scrambling (i.e. it minimizes $|\mathcal{C}_2^{1,2}(\alpha, \beta; t)|$), but also single-mode scrambling (i.e. it *also* minimizes $|\mathcal{C}_2^{1,1}(\alpha, \beta; t)|$).
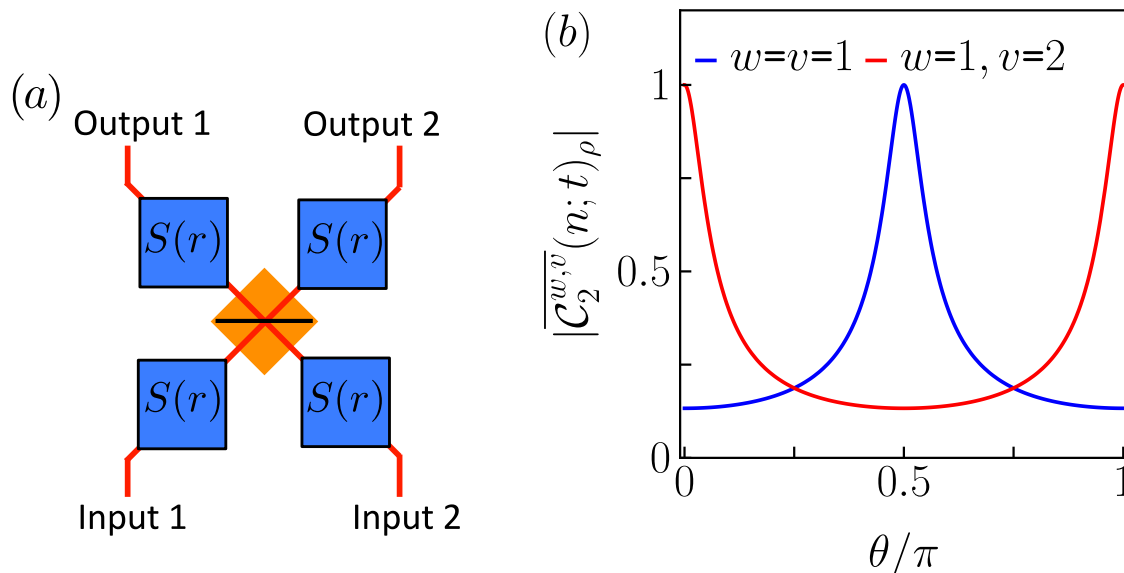
Figure 11.17: (a) A two-mode Gaussian analog to the random SNAP gate circuit of Fig. 11.16, where the SNAP gates are replaced with single-mode squeezing of strength $r = 1$. (b) The average OTOC ($n = 1$) for single-mode ($w = v = 1$) and multi-mode ($w = 1, v = 2$) scrambling, as a function of the beamsplitter transmissivity $\cos^2 \theta$.

The experimental realization of genuine scrambling circuits would also enable a powerful check on our assumption in Section 11.4, that aspects of genuine scrambling could simulated using averages over quasi scrambling systems. For example, we consider a quasi scrambling analog (Fig. 11.17) to our previous SNAP gate circuit (Fig. 11.16), with squeezing operations of strength $r$ instead of SNAP gates. In Appendix F, we compute the circuit's average OTOC $\overline{\mathcal{C}_2^{\omega,v}}(n;t)_\rho \equiv \overline{\mathcal{C}_2}(\mathbb{D}_n^\omega(t), \mathbb{D}_n^v)$ [Fig. 11.17(b)]. We find qualitatively similar behavior to the (non-averaged) OTOC of the random SNAP gate circuit, supporting our assumption. Nonetheless, some deviation is observed at the minimally-mixing values $\theta = 0, \pi/2$. We speculate that this is due to the additional conservation of the individual modes' photon numbers at these values in the SNAP gate circuit, which may inhibit OTOC decay and complicates this particular comparison.

## Measurement of TOCs and OTOCs

The most direct way to measure the amplitude of TOCs and OTOCs is to sequentially apply the operators in the correlation function to a state, and measure the probability to remain in that state. For instance, to measure the TOC $|\langle \boldsymbol{\gamma}|U^\dagger D(\boldsymbol{\xi}_1) U D(\boldsymbol{\xi}_2)|\boldsymbol{\gamma}\rangle|^2$, one would apply $D(\boldsymbol{\xi}_2)$, then $U$, then $D(\boldsymbol{\xi}_1)$, then $U^\dagger$, then measure the probability to be in state $|\boldsymbol{\gamma}\rangle$. This is depicted for TOCs and OTOCs in Figs. 11.18(a,c). For simplicity, we

consider correlations with respect to a coherent state $|\boldsymbol{\gamma}\rangle$. The probability to be in the coherent state can be measured by performing a displacement $D(-\boldsymbol{\gamma})$ and measuring the probability to be in the vacuum state.

One can measure TOCs and OTOCs themselves, and not just their amplitudes, using a control qubit and an interferometric scheme similar to Refs. [451, 512]. As shown in Fig. 11.18(b,d), here one initializes the control qubit in the state $(|0\rangle + |1\rangle)/\sqrt{2}$ and performs different operations on the CV system given different states of the control qubit. The complex-valued TOC/OTOC is found by measuring $X + iY$, where $X, Y$ are Pauli operators on the control qubit.

Finally, we present a concise scheme to measure average OTOCs, as introduced in Section 11.3. The only change from our individual OTOC schemes is that we now use an ancillary mode, prepared in the Gaussian state $|\psi_0\rangle \sim \int_{-\infty}^{\infty} dq e^{-\frac{1}{2}q^2/\Delta^2}|q\rangle$ (expressed in the position basis), to perform the 'ensemble' of displacements on the CV system of interest[8]. This is done using the SUM gate,

$$
\begin{array}{c}
|\psi_0\rangle \;\rule[0.5ex]{2.5cm}{0.4pt}\!\!\bullet\!\!\rule[0.5ex]{2.5cm}{0.4pt} \\[1ex]
\qquad\qquad\qquad\Big| \\[-0.5ex]
|\gamma\rangle \;\rule[0.5ex]{1cm}{0.4pt}\!\boxed{\text{SUM}}\!\rule[0.5ex]{1.5cm}{0.4pt}
\end{array}
\quad,
\tag{11.91}
$$

which acts on quadrature operators as

$$q_1 \rightarrow q_1 \qquad\qquad p_1 \rightarrow p_1 - p_2 \tag{11.92}$$

$$q_2 \rightarrow q_1 + q_2 \qquad p_2 \rightarrow p_2. \tag{11.93}$$

This implements a displacement $D(q)$ with probability $\sim e^{-q^2/\Delta^2}$ (shown by tracing out the ancilla). Average OTOCs require one to sample pairs of displacement operators $D(\boldsymbol{\xi})$ and $D(-\boldsymbol{\xi})$ in a correlated manner, which can be achieved by using the same ancilla for each displacement of the pair[9].

## Robust teleportation-based protocol

In this section, we turn our attention to a theoretical/conceptual question concerning the verification of scrambling. While OTOCs can characterize the phenomena of scrambling when measured perfectly, they are sensitive to experimental noise and decoherence, challenging the experimental measurement of scrambling [518]. For instance, loss and thermal noise, the most common imperfections in optical systems, both cause the OTOC to decay just as scrambling time-evolution would (see Appendix F). It is therefore desirable to characterize scrambling in a way which clearly distinguishes scrambling from such errors.

---

[8]A Gaussian thermal state, instead of a Gaussian pure state, could also be used as the ancilla.

[9]We note that this procedure for taking averages over Gaussian distributions of displacement operators is performed using only Gaussian operations. From Section 11.3, this implies that one can measure the non-Gaussianity of an operation $U \otimes U \otimes U^\dagger \otimes U^\dagger$ via only Gaussian operations, suggesting a possible resource theoretic interpretation of the quantity.

Figure 11.18: Measurement protocols for (a) the TOC amplitude $|\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_{|\gamma\rangle}|^2$, (b) the TOC $\mathcal{C}_1(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_{|\gamma\rangle}$, (c) the OTOC amplitude $|\mathcal{C}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_{|\gamma\rangle}|^2$, and (d) the OTOC $\mathcal{C}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2; t)_{|\gamma\rangle}$.

A robust verification protocol of scrambling has been recently proposed [518] and experimentally realized [274] for DV systems. This protocol draws inspiration from ideas in quantum gravity, and can be viewed as a many-body generalization of quantum teleportation [517]. Here, we briefly describe a similar teleportation-based protocol for CV scrambling. For simplicity, we restrict our attention to measuring scrambling by Gaussian unitaries, where the protocol will succeed with probability unity assuming no experimental error[10].

We begin our analysis with the ideal, error-free case, and demonstrate our protocol's robustness to error after. We seek to characterize quasi scrambling by a Gaussian unitary $U$ acting on two CV modes. Fig. 11.19 displays the set-up for the teleportation-based protocol. Initially, the system is prepared in two EPR pairs on modes $(2, 2')$ and $(1', R)$ whereas an arbitrary quantum state $|\psi\rangle$ is prepared on 1. For a later purpose, it is convenient to use the stabilizer formalism to characterize EPR pairs. For a two mode system with position and momentum operators $(q, p)$, $(q', p')$, the CV EPR pair $|\text{EPR}\rangle$ is defined as the $P = Q = 0$ eigenstate of operators

$$P = p + p' \qquad Q = q - q' \tag{11.94}$$

Note that this corresponds to the infinite squeezing limit of the two-mode squeezed state.

Next, we apply $U$ on $(1, 2)$, and its complex conjugate $U^\star$ on modes $(2', 1')$. While the protocol works with an arbitrary quasi scrambling unitary, we simplify our treatment by

---

[10]In DV systems, the analogous protocol succeeds with probability unity for Clifford unitaries, but also with lesser probability $1/d^2$ for non-Clifford unitaries. This probability goes to zero in the CV limit, hence our avoidance of non-Gaussian unitaries. In reality, imperfections in EPR preparation and measurement will 'cut-off' this CV limit, and one might expect that the protocol can also succeed with finite probability for non-Gaussian unitaries.
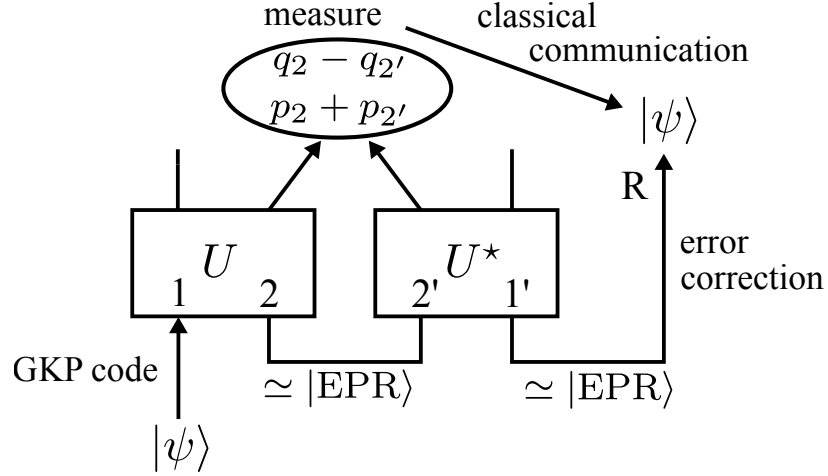
Figure 11.19:   Schematic of the teleportation-based protocol for robustly measuring CV scrambling. An initial state $|\psi\rangle$ and two CV EPR pairs are time-evolved under the unitary of interest $U$ and its conjugate $U^*$. After, a pair of modes $2, 2'$ are measured via the operators $q_2 - q'_2$ and $p_2 + p_{2'}$. The measurement outcome is used to error correct the original state, and the teleportation fidelity serves as a robust measure of information scrambling between modes 1 and 2 by $U$. In the presence of measurement uncertainties and imperfection in EPR preparations, the GKP encoding enables error-correction.

considering the following family of quasi scrambling unitaries:

$$\begin{aligned} p_1 &\to mp_1 + (m+1)p_2 & q_1 &\to mq_1 - (m-1)q_2 \\ p_2 &\to (m-1)p_1 + mp_2 & q_2 &\to -(m+1)q_1 + mq_2 \end{aligned} \tag{11.95}$$

where $m$ is arbitrary real number. When $m \neq 0, -1, 1$, observe that $U$ delocalizes any single-mode displacement operators to a two-mode displacement operator, the criteria for a non-trivial quasi scrambler. It is convenient to write the above transformation in the following, inverted manner:

$$\begin{aligned} mp_1 - (m+1)p_2 &\to p_1 & mq_1 + (m-1)q_2 &\to q_1 \\ -(m-1)p_1 + mp_2 &\to p_2 & (m+1)q_1 + mq_2 &\to q_2. \end{aligned} \tag{11.96}$$

The unitary $U$ contains squeezing since the total amplitude of displacements changes. The amount of squeezing $e^r \sim m$ plays an important role in fault-tolerance of the teleportation protocol, as we will see later.

After applying $U$ and $U^\star$, we measure $(2, 2')$ with the following operators:

$$P_{22'} = p_2 + p_{2'} \qquad Q_{22'} = q_2 - q_{2'}. \tag{11.97}$$

We send the measurement outcomes $Q_{22'}$ and $P_{22'}$ as a classical message to $R$. By applying the inverse transformation from Eq. (11.96), at $t = 0$, we have

$$P_{22'} = \left( -(m-1)p_1 + mp_2 \right) + \left( -(m-1)p_{1'} + mp_{2'} \right) \tag{11.98}$$

$$Q_{22'} = \left( (m+1)q_1 + mq_2 \right) - \left( (m+1)q_{1'} + mq_{2'} \right). \tag{11.99}$$

Since $p_2 + p_2' = 0$ and $q_2 - q_{2'} = 0$, we arrive at

$$P_{22'} = -(m-1)(p_1 + p_1') \qquad Q_{22'} = (m+1)(q_1 - q_{1'}). \tag{11.100}$$

Hence, the protocol teleports the following state on $R$:

$$D\left( \left( \frac{Q_{22'}}{m+1}, -\frac{P_{22'}}{m-1} \right) \right) |\psi\rangle \tag{11.101}$$

which is just a state shifted from $|\psi\rangle$. Since this shift can be corrected using the classical message, this protocol teleports the quantum state $|\psi\rangle$.

While we have focused on a particular family of quasi scramblers, our treatment generalizes to generic quasi scrambling unitaries $U$ that delocalize single-mode displacement operators. Namely, we can show that the teleported state is $|\psi\rangle$ up to displacements that can be undone using knowledge of the measurement result.

So far we have assumed that all implementations of the protocol are perfect. In actuality, one might incur errors due to decoherence, or a mismatch between the experimentally-implemented unitaries $U$, $U^*$. This protocol is robust to both of these effects, as discussed in great detail for DV systems in Ref. [518]. There are also further imperfections unique to the CV limit. For example, it is not possible to prepare perfect CV EPR pairs, and one must approximate them with two-mode squeezed states. In addition, measurements of $Q_{22'}$ and $P_{22'}$ will involve some uncertainty.

Both of these imperfections lead to unknown displacement errors on the teleported state, and, at a calculational level, appear similar to inserting some density matrix $\rho$ in the tensor contractions corresponding to the EPR pair and measurement. This link to finite-temperature scrambling, as well as experimental relevance, motivate us to qualitatively address when teleportation can succeed despite these imperfections. We consider the above protocol for a general $N$-mode Gaussian unitary $U_{\mathbf{S}}$ (taking displacements $\xi \to \mathbf{S}\xi$, where $\mathbf{S}$ is a symplectic matrix), in the presence of a measurement error, i.e. one recorded $(Q_{22'}, P_{22'})$ but actually measured $(Q_{22'}, P_{22'}) + \Delta\xi$. Imperfections due to an imperfect EPR pair—a two-mode squeezed state of $\sim n_{\mathrm{EPR}}$ photons—are treated similarly: they arise from the fact that small displacements $D(\Delta\xi)$ leave the state unchanged for $\Delta\xi < 1/\sqrt{n_{\mathrm{EPR}}}$. To see these errors' effect, one can compute the displacement $D(\Delta\mathbf{z})$ on $|\psi\rangle$ that would have given rise to the measurement error: $\Delta\xi = \mathbf{S}_{1,m}\Delta\mathbf{z}$, where $\mathbf{S}_{1,m}$ is the $2 \times 2$ submatrix of $\mathbf{S}$ between mode 1 and the measured mode $m$. Due to the error, the protocol decodes the 'wrong' state, off by a displacement $\Delta\mathbf{z}$. However, as we have seen, small displacements are only distinguishable from the identity operation on subspaces of greater than $1/|\Delta\mathbf{z}|^2$ photons. Therefore, if $|\psi\rangle$

has $\lesssim n_\psi$ photons, the error has no effect when $\sqrt{n_\psi} < 1/|\Delta \mathbf{z}|$. Since an $N$-mode unitary with single-mode squeezing $e^r$ has elements of typical magnitude $\sim e^r/\sqrt{N}$, the protocol can only teleport states of photon number $\sqrt{n_\psi} < e^r/(\sqrt{N}|\Delta\xi|)^{11}$. This makes sense: a state $|\psi\rangle$ is described by displacement operators separated by distance $\sim 1/\sqrt{n_\psi}$ in phase space. This translates to a distance $\sim (e^r/\sqrt{N}) \times (1/\sqrt{n_\psi})$ on the measured mode after application of $U$. Our teleportation constraint implies that we can only teleport states when this phase space distance is resolvable despite the imperfections, $(e^r/\sqrt{N}) \times (1/\sqrt{n_\psi}) > \Delta\xi$.

We also hope to make the teleportation protocol fault-tolerant to these imperfections by using quantum error-correction. Gottesman, Kitaev and Preskill have proposed a way of encoding a finite-dimensional qudit into a single oscillator. This error-correcting code, called the GKP code, enables us to correct small displacement errors. Suppose we prepare a qudit input state and encode it into mode 1 via the GKP encoding. For simplicity, we consider the error from faulty measurements, and assume that $Q_{22'}$ and $P_{22'}$ suffer from uncertainties $\Delta Q_{22'}$ and $\Delta P_{22'}$. When the quantum state is teleported to $R$, it suffers from displacement noises $D(q_R)$ and $D(p_R)$ of $|q_R| \sim \frac{\Delta Q_{22'}}{m+1}$ and $|p_R| \sim \frac{\Delta P_{22'}}{m-1}$. By using the GKP code, displacement errors can be corrected if $m$ is sufficiently large, and thus the teleportation protocol can be made fault-tolerant. We remark, however, that the aforementioned protocol does not suppress errors from the imperfect EPR preparations on $(2, 2')$ and $(1', R)$.

Observe that the induced error can only be suppressed when the squeezing $m$ is large. There appears to be an intriguing relation between the amount of squeezing and the amount of teleported information. This leaves an interesting future problem concerning the upper bound on the information capacity under energy constraints. Another relevant future problem concerns the relation between this protocol and the average OTOC. For the quasi scramblers in Eq. (11.96), the average OTOCs becomes small as $m$ increases. Hence, in the presence of measurement uncertainties, the smaller average OTOCs enables larger amount of fault-tolerant teleportation.

## 11.6 Discussion and outlook

In this Chapter, we sought to understand scrambling in CV systems from a quantum information theoretic perspective. Scrambling in CV systems had been studied in a number of seminal previous works [65, 94, 114, 200, 386, 405], where OTOCs for quadrature operators and number operators are explored in the context of specific example Hamiltonians. These

---

[11]For an imperfect EPR pair, an additional constraint arises from squeezing's effect on the two-mode squeezed state. Although squeezing increases the state's typical photon number, it vastly reduces its width in certain, typically many-mode, quadratures. This implies that there exist many-mode displacements of magnitude $|\Delta\xi| \sim e^r/\sqrt{n_{\text{EPR}}}$ that leave the many-body state (after the application of $U$, $U^*$) unchanged. Projecting to the measured mode, these are equivalent to measurement errors $\sim e^r/(\sqrt{N}\sqrt{n_{\text{EPR}}})$. The previous arguments then give the constraint $\sqrt{n_\psi} > \sqrt{n_{EPR}}$, which is a more powerful constraint at high squeezing $e^r > \sqrt{N}$. We see that squeezing can only assist teleportation up to a certain limit, after which the teleportation is bounded solely by the strength of the imperfections.

studies revealed that the OTOC with quadratures operators can enable a quantum-classical correspondence within the Ehrenfest time [114, 405]. Compared to these previous studies, our work provides: (i) general interpretations of TOCs and OTOCs in terms of operator spreading, (ii) an investigation of the scrambling dynamics of generic local circuits based upon a CV analog of 2-designs, (iii) a quantum-classical correspondence in phase space, via our choice to use displacement operators in OTOCs, and (iv) an experimental blueprint for probing scrambling in cavity QED systems.

Our characterization of CV information dynamics is more appropriate in certain physical contexts than others. Specifically, a key idea of our work is to introduce a density matrix representing the Hilbert space of interest, which led to a smooth regularization of various quantities. Throughout the paper we choose a thermal density matrix $\tilde{\rho}_{n_{\text{th}}}$—other choices may lead to different coarse-graining procedures at short-distances and cut-offs at long-distances.

For this description to apply, the system of interest must feature a large local Hilbert space with a smooth cut-off. The relevant notion of locality is determined by the operators with which we wish to probe the system, and does not necessarily have to agree with the UV lattice cut-off of the system. Notably, these conditions can hold even for large but finite dimensional systems, whenever the quantum states under study are associated with a wide range of energy scales. More specifically, consider a Hamiltonian measuring some energy $H_0$, with eigenstates $|E_n\rangle, 1 \leq n \leq N$. When $N$ is large and the relevant states $\rho$ have a distribution $\langle E_n|\rho|E_n\rangle$ spread out over a range of energies $E_n$, and decaying smoothly towards those states with larger energies, a CV description applies.

We conclude this Chapter by mentioning a few open directions in the study of CV quantum information dynamics. The first set of open directions relates to our understanding of genuine scrambling. Most importantly, the speed and saturation of operator volume increase characterized by OTOC decay requires further study. For displacement-operator-based OTOCs, the connection between the initial decay and the Lyapunov exponent in classical chaotic systems can be further explored. This may lead to a deeper understanding of different classes of non-Gaussian unitaries. Related to this, it would be interesting to numerically investigate the effects of the small, but finite mass term $m$ in the cubic phase gate model in Sec. 11.3. Eventually, one hopes to derive bounds on operators' volume increase when the system is constrained by certain conservation law.

The second set of directions concerns our understanding of local random Gaussian circuits. In Sec 11.4, we identified a *q*uadratic growth of entanglement and linear increase of fluctuation for such systems, arising from an increase in the accessible local Hilbert space. However, a full theoretical model that explains the deviation from the KPZ scaling still needs to be developed. Such a model could deepen our understanding of scrambling dynamics in general CV systems.

Additionally, extending the study in Sec. 11.5 on models with photon number conservation law to a larger scale may deepen our understanding of conservation laws' consequences for scrambling dynamics. Performing the SNAP gate-based experiment proposed in Section 11.5 would be very instructive on this open problem, as numerical simulation becomes

difficult for greater than a few modes. For state-of-art experimental platforms, the realization of this experiment is plausible in the near future. Adding squeezing to such system would also allow the verification of the theoretical results, and assumptions, in Sec 11.4. Our work lays a solid theoretical foundation for such experimental studies.

The third set of directions regards the construction of CV unitary $k$-designs. We still understand very little about higher designs in the CV case. Evaluating the new frame potential in Eq. (F.31) for $k \geq 2$, and more general ensembles, as well as relating it to notions of complexity, are important future directions. Additionally, in light of our findings on Gaussian 2-designs, it remains an open question whether there exists a more appropriate definition of higher CV designs that is compatible with soft energy regularization. Nevertheless, our CV 1-design and 2-design 'analog' may still be useful for applications such CV state tomography (1-design) and compressed sensing (2-design) [401].

Another interesting direction concerns the definition of a 'size' for time-evolved operators at finite temperature. In DV systems at infinite temperature, the size of operators $O(t)$ corresponds to the average number of qubit supports in the Pauli decomposition of $O(t)$. As discussed in Section 11.3, the size of $O(t)$ can be measured by OTOCs. However, the notion of the size of operators $O(t)$ becomes ambiguous at finite temperature since OTOCs depend on $\rho$. An important question is to how to define the size of operators in the presence of $\rho$ at finite temperature in a physically meaningful manner. In a recent work on the operator growth in the SYK model [379], a possible definition of the finite temperature size of $O(t)$ is proposed. Namely, the authors argued that the size should be defined by subtracting the thermal background, i.e. as the difference between the support of $\rho^{1/4}O(t)\rho^{1/4}$ and $\rho^{1/2}$. They showed that the Lyapunov growth of OTOCs corresponds to the exponential growth of this difference.

On the other hand, in Section 11.3, we have argued that thermal TOC and OTOC measure coarse-grained volumes of the operator spreading where the resolution of the phase space is set by $\rho$. Specifically, we saw that a thermal state $\tilde{\rho}_{n_{\text{th}}}$ induces a Gaussian blurring of the phase space. One concrete open question concerns the connection between their proposal of subtracting the thermal background and our results developing the notion of coarse-graining. Here we present a heuristic argument while postponing rigorous discussions to the future work. Recall that the size of the operator can be counted by the number of qubits. The coarse-grained volume $V$ in the phase space roughly corresponds to a $\sqrt{V}$-state quantum spin, so it can be embedded in $\frac{1}{2}\log V$ qubits. So, one may assign $\frac{1}{2}\log V$ as the size of the operator. The coarse-grained volume $V$ depends on the scale of the resolution set by $\rho$. Letting an approximate radius of the spreading be $R$ and the resolution be $\delta$, the volume is $V = R^2/\delta^2$, the size will be given by $\frac{1}{2}(\log R - \log \delta)$. Hence we may interpret the second term as the thermal background.

A final set of open questions is on the connection to resource theory. It is well known that universal quantum computation requires non-Clifford operations, since Clifford operations admit efficient classical simulations according to the Gottessman-Knill theorem. It would therefore be interesting to assign a resource theoretic interpretation of $\sum_{Q \in \text{Pauli}} |f[Q; P(t)]|^4$, thereby relating quantum computational power and scrambling/decay of OTOCs. Similarly

in CV systems, the preservation of OTOC amplitude by Gaussian unitaries (quasi scramblers) and the decay of OTOC amplitude caused by non-Gaussian unitaries (genuine scramblers) might lead to a new resource theory framework for non-Gaussianity [14, 165, 166, 320, 454, 537]. Finally, there has also been some recent interest in characterizing the complexity of Gaussian CV states relevant for quantum field theory [92]; we speculate that our general approach, based upon frame potentials and OTOC decays, can be applied towards a more broad characterization of the complexity of states in field theories.

# Part IV

# Learning about Quantum Dynamics

# Chapter 12

# Learning quantum systems via out-of-time-order correlators

In this Part, we seek to apply the ideas contained in Part III to questions in quantum learning (introduced in Ch. 2.4). As we saw in Ch. 2.4, quantum learning is particularly interesting since learning properties of quantum systems can pose challenges not present in their classical counterparts [220, 329]. These challenges stem fundamentally from the existence of entanglement. In the context of learning quantum dynamics or quantum processes, these difficulties are most commonly encountered in *strongly-interacting* quantum systems. Strong interactions can introduce non-local entanglement throughout the system at short time scales, and are found to thereby inhibit the learning of system properties (e.g. the Hamiltonian) from physical observables [10, 220, 354, 491, 530].

The ubiquity of strong interactions in experimental applications of quantum learning has spurred a variety of solutions to this problem. For instance, in nuclear magnetic resonance (NMR) spectroscopy, a suite of technologies have been developed to controllably dampen undesired strong interactions between solid-state nuclear spins, which has enabled the identification of hitherto inaccessible molecular structures [279]. In a similar spirit, in quantum device characterization [481] and quantum sensing [530], dynamical decoupling control sequences [145] can effectively eliminate unwanted interactions and improve learning of the residual interactions. Other approaches include learning by transducing quantum data from the system onto a quantum simulator [10, 220, 480, 491], or learning from high-precision local measurements at early times, before entanglement has formed [32, 186]. Nonetheless, owing to incomplete control or limited experimental precision, many physical systems remain unlearnable with existing approaches.

In this Part, we introduce a new paradigm for learning properties of quantum dynamics—learning via out-of-time-order correlators (OTOCs) [46, 113, 114, 194, 213, 277, 315, 316, 340, 436, 442, 476, 506]. In the present Chapter, we utilize the OTOC as a tool for learning properties of *strongly-interacting* quantum Hamiltonians. Our application is motivated by a simple intuition: while time-ordered observables decay quickly as a system becomes entangled, out-of-time-order observables continue to fluctuate up to long times (Fig. 12.1).
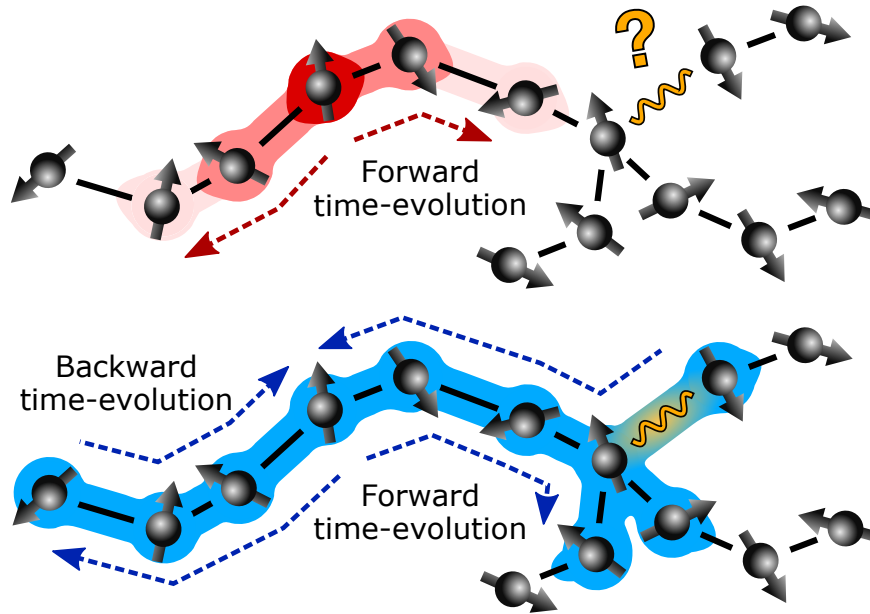
Figure 12.1: Schematic of time-ordered correlators (TOCs) and out-of-time-order correlators (OTOCs) in strongly-interacting systems. TOCs typically decay in $O(1)$ times and distances (top, red), making it hard to learn features (yellow bond) that manifest only at late times. OTOCs utilize backwards time-evolution to "refocus" many-body correlations (bottom, blue), enabling learning of such features.

Guided by this intuition, we demonstrate the power of learning via OTOCs across a range of physical systems, supported by numerical studies, phenomenological estimates, and rigorous information-theoretic proofs. We begin in locally-interacting systems, where we identify two general scenarios in which OTOCs provide a strong learning advantage: ($i$) when experimental access to the system is spatially-restricted, for example via a single "probe" qubit [83, 167, 304], and ($ii$) for detecting weak interactions in an otherwise strongly-interacting system [481, 530]. We characterize these advantages using both information-theoretic measures (the Fisher information) and performance metrics for concrete learning tasks. Moreover, we find that the advantages are robust to experimental read-out error and time-reversal imperfections arising from strong coupling with an environment or decoherence. In Ch. 13, we substantiate these results by introducing a learning task for which OTOCs provide a provable exponential advantage over any time-ordered learning protocol.
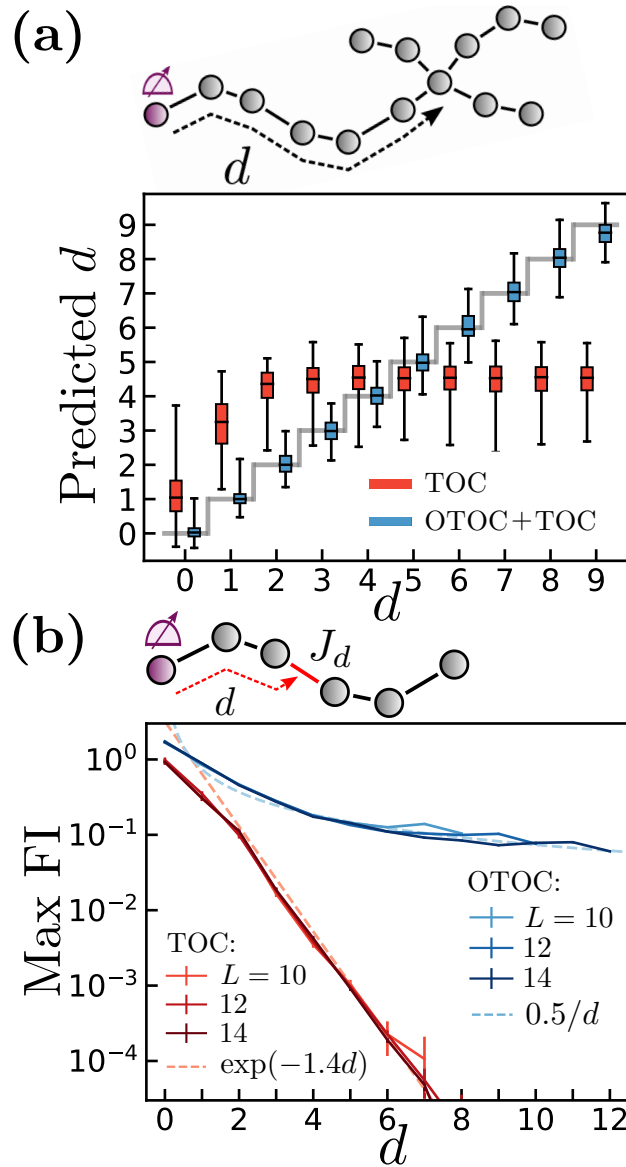
Figure 12.2: Learning with state preparation and read-out restricted to a probe qubit, and local unitary control over the remaining system. **(a)** Results from SVM regression for learning the distance, $d$, in the spin geometry shown, with access to TOCs (red) or both TOCs and OTOCs (blue). Color bars (black ticks) denote 75% (100%) percentiles of predictions on 200 disorder realizations, and grey step function represents the actual $d$. **(b)** Fisher information, $\mathrm{FI}(J_d|C)$, of an interaction, $J_d$ (top; red line), a distance $d$ away from the probe (top; purple circle), maximized over all correlators, $C$, in an $L$-qubit 1D chain. The FI decays exponentially in $d$ when $C$ is time-ordered (red), and algebraically, $\sim 1/d$, when $C$ is out-of-time-order (blue).

## 12.1 Behavior of time-ordered vs. out-of-time-order correlators

We begin by reviewing the phenomenology of time-ordered and out-of-time-order correlators in ergodic locally-interacting systems (Fig. 12.1). A time-ordered correlator (TOC) is defined as any correlation function that takes the following general form:

$$C_{\text{TOC}} = \text{tr}(A_k(t_k) \ldots A_1(t_1) \, \rho \, B_1(t_1') \ldots B_\ell(t_\ell')). \tag{12.1}$$

where the operators $A, B$ increase in time away from the initial density matrix $\rho$, i.e. $t_k > \cdots > t_1$ and $t_\ell' > \cdots > t_1'$. Time-ordered correlators can be measured by evolving the state $\rho$ forward in time (e.g. via Hamiltonian evolution $O(t) = e^{iHt}Oe^{-iHt}$) while applying intermediary quantum operations at each time $t_i, t_j'$[1]. Any correlation function that does not obey this form is called an *out-of-time-order* correlator.

A common example of a time-ordered correlator is the two-point function,

$$C_{\text{TOC}} = \langle V_x(t)W_{x'}(0) \rangle, \tag{12.2}$$

where $\langle \cdot \rangle \equiv \text{tr}(\cdot)/2^L$ denotes the infinite temperature trace for $L$ qubits, and $V_x, W_{x'}$ are local operators at sites $x, x'$. Such correlators measure the spread of local quantities in space and time; for instance, how much spin prepared at site $x'$ at time zero has transferred to site $x$ at time $t$. A wide range of literature on thermalization in strongly-interacting systems has found that local TOCs typically decay quickly, i.e. in $O(1)$ times, to their thermal values [120]. This quick decay can inhibit learning tasks, since no additional information can be acquired from the TOC at times after the decay has occurred [354].

Meanwhile, the prototypical out-of-time-order correlator is the four-point function [506],

$$C_{\text{OTOC}} = \langle V_x(t)W_{x'}(0)V_x^\dagger(t)W_{x'}(0) \rangle, \tag{12.3}$$

with local operators $V_x, W_{x'}$. Unlike time-ordered measurements, OTOCs typically require both forwards *and* backwards time-evolution to measure [111, 506]. (Importantly for our application, nearly all experimental techniques for time-reversal rely only on the type of interaction being reversed and require no knowledge of the specific Hamiltonian, which one might wish to learn. For example, the same pulse sequence reverses an arbitrary dipole-dipole coupling Hamiltonian in an NMR experiment [416].) Physically, the OTOC probes whether information encoded at site $x'$ at time zero is contained in correlations involving site $x$ at time $t$. This is quantified by the squared commutator of a time-evolved operator at $x$ with a local operator at $x'$, $\langle |[V_x(t), W_{x'}(0)]|^2 \rangle = 1 - C_{\text{OTOC}}$. In local strongly-interacting

---

[1]To see this explicitly, we specify to the case that $A_k, B_l$ are Pauli operators. Consider four possible operations on the density matrix $\rho$: the two unitary rotations $\rho \to \frac{1 \pm iA_k}{\sqrt{2}}\rho\frac{1 \mp iA_k}{\sqrt{2}}$ (where the signs are chosen together) and the two projections $\rho \to \frac{1 \pm A_k}{2}\rho\frac{1 \pm A_k}{2}$. By taking linear combinations of these four possibilities in classical post-processing, we can in effect implement the matrix multiplications $\rho \to A_k\rho$ (and similarly $\rho \to \rho B_l$). Applying the appropriate operation at each time $t_k$ (and $t_l'$) gives Eq. (12.1).

systems, operators are expected to spread ballistically according to the connectivity of the system [17, 293, 340]. Crucially, this spread continues for a duration proportional to the system's spatial extent $\sim L$ by which time the information has been delocalized across the entire system.

This phenomenology leads to two central intuitions for learning from OTOCs. First, the dynamics of the OTOC contain information primarily about the connectivity of the system under study. Second, the OTOC continues to reveal such information up to $O(L)$ times, long after TOCs have decayed. Notice that this timescale increases as the system size increases. In what follows, we apply these intuitions to identify two broad regimes where access to OTOCs provides a significant learning advantage.

## 12.2 Learning with restricted access

The first regime we consider is learning in systems with *restricted access*. Specifically, motivated by recent advances in solid-state defects [66, 167, 304] and NMR [58, 81, 459], we focus on the scenario where an experimenter has state preparation and read-out capabilities over only a single "probe" qubit interacting with a larger system that one wishes to learn. We note that high-fidelity OTOC measurements have already been achieved in similar setups by using rapid global pulse sequences to reverse time-evolution [416, 418, 486]. Previous theoretical approaches to learning in this scenario have been limited to non-interacting dynamics [81, 82, 83, 135, 441, 527]. Meanwhile, experiments have found that it is in general difficult to learn features of a system that are distant from the probe qubit [167, 304]. In strongly-interacting systems, this difficulty can be understood from the quick decay of correlation functions in space and time. Here, we provide evidence via phenomenological estimates (Appendix G.2) and numerical simulations (Fig. 12.2) that access to OTOCs can exponentially improve the learnability of distant features.

To be concrete, we will assume for now that the experimenter has local unitary control over the qubits of the larger system[2]. We will also assume that the larger system begins in an infinite temperature (i.e. maximally mixed) state, which is the natural scenario in NMR and solid-state defect setups [66, 459]. Within these assumptions, a simple class of measurement protocols proceeds as follows:

1. Prepare the probe qubit $p$ in an eigenstate of an operator $V_p$, such that the density matrix of the entire system is $\rho = \frac{1}{2}(\mathbb{1}_p + V_p) \otimes \frac{1}{2^{L-1}} \mathbb{1}_{\text{sys}}$.

2. Time-evolve by time $\tau$.

3. Perturb the system by a unitary operation $W_x$ on a qubit $x$.

---

[2]To be precise, our numerical simulations are relevant to the scenario where the experimenter can perform $O(1)$ local spin rotations within the time-scale set by the inverse spin interaction strength. If substantially faster local rotations are possible, then the Hamiltonian interactions could be individually tuned via local dynamical decoupling, which would allow substantially faster learning [481].

4. Time-evolve by a time $\tau'$.

5. Read out the expectation value of $V_p$ on the probe qubit.

Taking $\tau, \tau'$ to be positive (e.g. $\tau = \tau' = t/2$), this allows measurement of time-ordered correlation functions of the form $\langle V_p(t)\, W_x(t/2)\, V_p(0)\, W_x^\dagger(t/2)\rangle$. With access to reversible time-evolution (e.g. $\tau = -\tau' = t$), the above protocol also allows measurement of out-of-time-order correlation functions $\langle V_p(0)\, W_x(t)\, V_p(0)\, W_x^\dagger(t)\rangle$. In Appendix G.2 and G.3, we discuss how learning is modified when $W$ is instead a *global* spin rotation over the larger system.

We begin our exploration of learning via OTOCs by introducing a concrete learning task. We consider the following scenario: one is given access to a quantum system consisting of two spin chains intersecting at a distance $d$ from a probe qubit [Fig. 12.2(b)]. The value of $d$ as well as the specific Hamiltonian parameters of the system are unknown (see below for the specific distribution that the Hamiltonian is drawn from). The goal is to learn the value of $d$, i.e. the geometry of the system, from measurements of the system's correlation functions.

To solve this task, we assume that the experimenter is capable of simulating quantum dynamics on either a classical or quantum computer. Since the task involves high-dimensional input data (i.e. the correlators for every $x, t$), we will approach it using machine learning techniques. Specifically, we envision using the quantum simulator to compute the correlation functions of an ensemble of Hamiltonians for each value of $d$. These ensembles can then be used to train a classical learning model to predict an unknown Hamiltonian's value of $d$ given its correlation functions.

Let us briefly summarize our numerical simulations in more detail (see Appendix G.1 for a complete description). Throughout this work, we consider spin systems with disordered on-site fields, $H_f = \sum_{i,\alpha} h_i^\alpha \sigma_i^\alpha$ with $h_i^\alpha \in [-1, 1]$ and $\alpha = x, y, z$, and dipolar interactions between neighboring spins, $H_c = \sum_{\langle ij\rangle} J_{ij}(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y - 2\sigma_i^z \sigma_j^z)$ with $J_{ij} \in [0.6, 1.4]$. We specify to Floquet dynamics consisting of alternating applications of $H_f$ and $H_c$ for time $T = \pi/2$, and simulate time-evolution via Krylov-subspace methods [1]. We expect that learning Floquet dynamics will be qualitatively similar to learning time-independent Hamiltonian dynamics at moderate times and distances, which we are restricted to in our numerics (see Appendix G.3 for numerical support of this statement). At larger distances we expect Hamiltonian dynamics to be dominated by hydrodynamics of the conserved energy (Appendix G.2) and the two will differ.

Returning to the learning task at hand, we train a support vector machine (SVM) on 3000 randomly drawn Hamiltonians (300 for each value of $d = 0, \ldots, 9$), and test its performance on 2000 additional Hamiltonians. To ensure that learning is not sensitive to fine-tuned features of the correlation functions, we add a Gaussian distributed "read-out error" to all correlation functions, with mean zero and standard deviation $\delta = 3\%$. The model's predictions as a function of the actual value of $d$ are displayed in Fig. 12.2(b), for learning either via TOCs (red) or both TOCs and OTOCs (blue). We find that learning via OTOCs allows accurate predictions of $d$ within $\pm 1$ of its actual value for all distances probed (up to
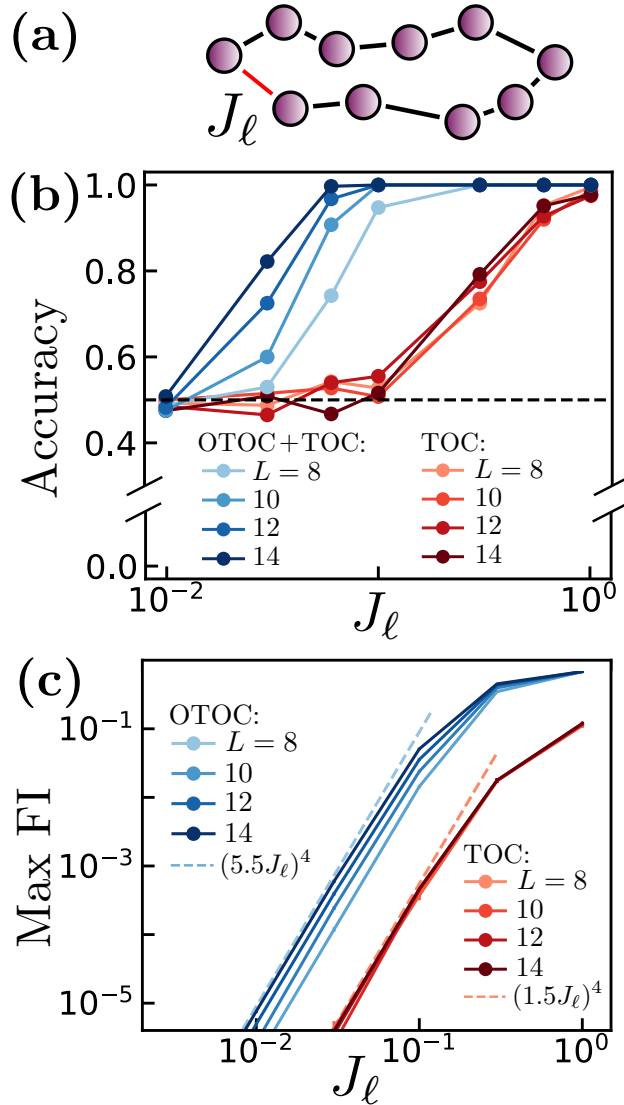
Figure 12.3:   **(a)** Learning a weak "link" interaction (red line) in a 1D spin chain with otherwise strong interactions (black lines). **(b)** Accuracy of binary SVM classification of whether the link is present or absent, as a function of the link strength $J_\ell$ and at fixed readout error $\delta = 3\%$. Learning via OTOCs can detect smaller $J_\ell$ as $L$ increases, while the TOC can only detect relatively large $J_\ell$, independent of $L$. **(c)** The maximum Fisher information $\mathrm{FI}(\log(J_\ell)|C)$ of $J_\ell$ decays $\sim J_\ell^4$ for small $J_\ell$, and is enhanced in OTOCs (blue) compared to TOCs (red) by a factor that increases with $L$.

$d = 9$). In contrast, with access to only TOCs, the model performs significantly worse for all $d$ and resorts to nearly random guessing for $d \gtrsim 3$.

To evaluate the learning advantage of OTOCs independent of a specific learning task, we turn to the Fisher information (FI). The FI quantifies the amount of information that a random variable (e.g. a correlation function $C$, measured within some read-out error $\delta$) carries about an unknown parameter (e.g. a coupling strength, $J$), and thereby bounds the ultimate learnability of the parameter [242]. If one assumes that read-out errors are normally distributed, the FI is simply a squared derivative, $\mathrm{FI}(J|C) \equiv \delta^2 \mathrm{FI}(J|C; \delta) = |\partial C/\partial J|^2$, where we remove the $\delta$-dependence by introducing a factor $\delta^2$.

We numerically compute the FI in ergodic 1D spin chains, where one seeks to learn a coupling $J_d$ lying a distance $d$ away from a probe qubit [Fig. 12.2(b) inset] [81, 82, 83, 135, 441, 527]. We consider the same set of correlation functions as specified for the learning task in Fig. 12.2(a). In Fig. 12.2(b), we plot the maximum Fisher information $\max_C \mathrm{FI}(J|C)$ over all correlation functions (i.e. over all $x, t$), averaged over 200 and 1000 disorder realizations for TOCs and OTOCs respectively. We find that the maximum FI of TOCs (red) decays exponentially in the distance $d$ from the probe qubit. In contrast, the maximum FI of OTOCs (blue) follows a slow algebraic decay, $\sim 1/d$, thereby achieving a multiple-order-of-magnitude advantage over TOCs even at modest distances, $d \gtrsim 3$. This algebraic decay arises from the $\sim \sqrt{t}$ broadening of the OTOC wavefront in time [340], see Appendix G.2 for a full phenomenological derivation.

## 12.3 Learning weak interactions

We now turn to our second learning scenario: characterizing weak interactions in an otherwise strongly-interacting system. Such characterization is notoriously difficult because weak interactions take long times to manifest (of order the inverse interaction strength), at which point TOCs have decayed due to the strong interactions. Previous approaches require either dynamical decoupling of the strong interactions [481, 530] or high-precision measurements at early times [32, 186]. We will now show that access to OTOCs allows one to side-step these requirements when characterizing weak interactions that *change the connectivity* of a strongly-interacting system. Notably, in contrast to the previous learning scenario, this advantage holds when the experimenter is capable of measuring all local correlation functions of the system of interest.

For concreteness, we specialize to 1D spin chains with a single "weak link" interaction, of strength $J_\ell$ much less than the typical interaction strength $J$ [see Fig. 12.3(b) inset]. We consider TOCs and OTOCs of the form Eq. (12.2) and Eq. (12.3), where $x, x'$ run over all qubits in the system. We anticipate that access to more general correlators within a given time-ordering, e.g. via shadow tomography or related techniques [2, 112, 217], will not qualitatively change the observed physics (see Appendix G.1).

We begin as before with a concrete learning task. Specifically, we suppose that one is given access to a spin chain with unknown Hamiltonian parameters and either no link
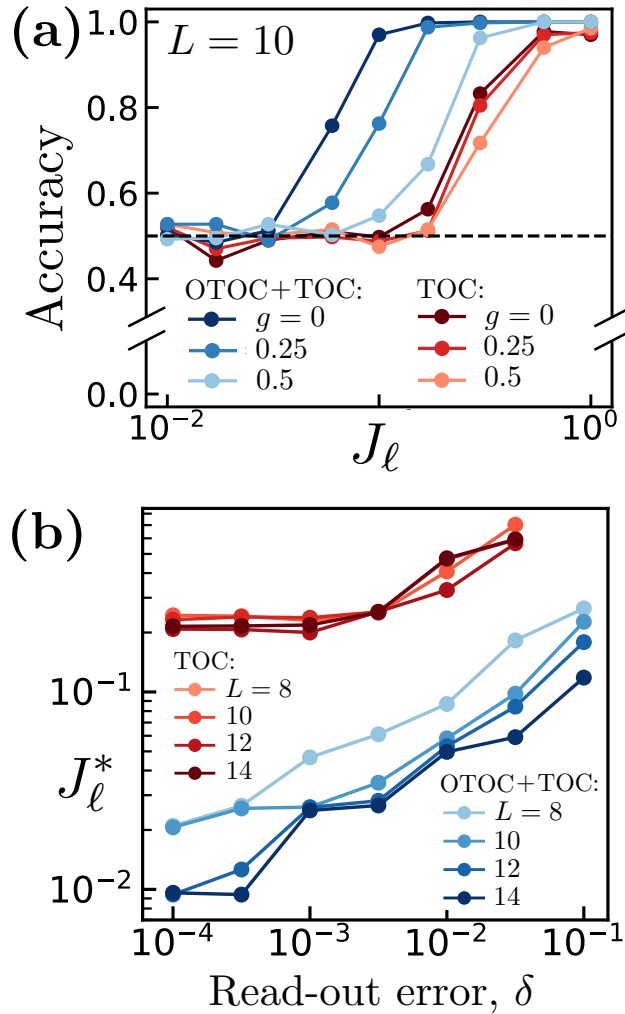
Figure 12.4: Learning as a function of experimental error, in the "weak interaction" learning task of Fig. 12.3(a). **(a)** Accuracy of binary SVM classification as in Fig. 12.3(a), now with a coupling $g$, to an extrinsic cavity mode that is not time-reversed (cavity frequency $\omega = 1.7$). Despite imperfect time-reversal, learning via OTOCs continues to provide an advantage up to large spin-cavity couplings $g \sim 0.5$. **(b)** The minimum link strength $J_\ell^*$ classifiable with $> 90\%$ accuracy as a function of read-out error $\delta$, obtained by repeating Fig. 12.3(a) for each $\delta$. The minimum link strength in general decreases with decreasing $\delta$; for learning via TOCs, this decrease plateaus for $\delta \lesssim 0.1\%$, indicating that learning below this value is not limited by read-out error.

interaction ($J_\ell \to 0$) or a fixed non-zero weak link interaction strength $J_\ell$. For each fixed value of $J_\ell$, we train a binary SVM classifier on the correlation functions [Eqs. (12.2), (12.3)] of 300 disorder samples, again including a read-out error $\delta = 3\%$ in each correlator value. We test model performance on 200 additional samples; the resulting classification accuracies are shown in Fig. 12.3(b). We observe the following general trends: ($i$) the accuracy decreases as $J_\ell$ decreases; ($ii$) learning via both OTOCs and TOCs (blue) allows detection of $\sim 10$ times smaller $J_\ell$ than learning via only TOCs (red); and ($iii$) OTOCs allow detection of increasingly small $J_\ell$ as the size $L$ of the chain increases.

To understand this behavior analytically, we first note that the optimal correlation functions for detecting the link will typically involve operators lying immediately adjacent to that link, on both of its sides. These correlators measure either the transfer of spin polarization (for TOCs) or operator support (for OTOCs) across the link, and will be non-trivial only if the link interaction strength is nonzero. For TOCs, one expects spin polarization to cross the link incoherently, at a rate $\sim J_\ell^2/J$, where $J$ is the typical strong interaction strength. Combined with an overall exponential decay of spin in time (if the system has no conserved quantities), we expect $C_{\text{TOC}} \sim (J_\ell^2/J)\, t\, e^{-Jt}$. For OTOCs, one expects an operator's support to cross the link at a similar rate, $1 - C_{\text{OTOC}} \sim (J_\ell^2/J)t$. Crucially however, this growth persists until much *later* times, $t \sim L/J$, at which information traveling "around" the chain will abruptly cause the OTOC to decay to zero. The optimal time for detecting the link occurs when these correlators are maximized, since each is zero in the absence of the link. The TOC is maximized at an order one time $t \sim 1/J$, at which the correlator magnitude $C_{\text{TOC}} \sim J_\ell^2/J^2$ is suppressed by the square of the weak link interaction strength. In contrast, the OTOC is maximized at a much later time $t \sim L/J$, and thereby features a magnitude $1 - C_{\text{OTOC}} \sim L(J_\ell^2/J^2)$. In both cases we see that detection of the link becomes more difficult as the link strength decreases. Detection via the OTOC is enhanced by a factor of $L$, which captures the connectivity change associated with the link.

We confirm these estimates quantitatively by computing the Fisher information of the link interaction strength. In Fig. 12.3(b), we plot the maximum of the Fisher information, $\max_C \text{FI}(\log(J_\ell)|C)$, over all local correlation functions, averaged over 100 disorder realizations. Here, we consider the logarithm of the link interaction strength in order to appropriately compare the Fisher information over multiple orders of magnitude of the interaction. The Fisher information of $\log(J_\ell)$ bounds the learnability of the interaction strength as a *percentage* of its actual value. Applying our phenomenological estimates, we predict that $\text{FI} \sim J_\ell^4/J^4$ for TOCs, and $\text{FI} \sim L^2 J_\ell^4/J^4$ for OTOCs. Observing Fig. 12.3(b), we indeed find that the FI is suppressed by $\sim J_\ell^4$ (dashed lines) for small $J_\ell$, and displays a multiplicative advantage for OTOCs (blue) compared to TOCs (red), which grows as $L$ increases.

## 12.4   Effect of experimental errors

Let us now address the impact of experimental errors on learning. We begin with errors that accumulate throughout time-evolution. These may occur from extrinsic decoherence or
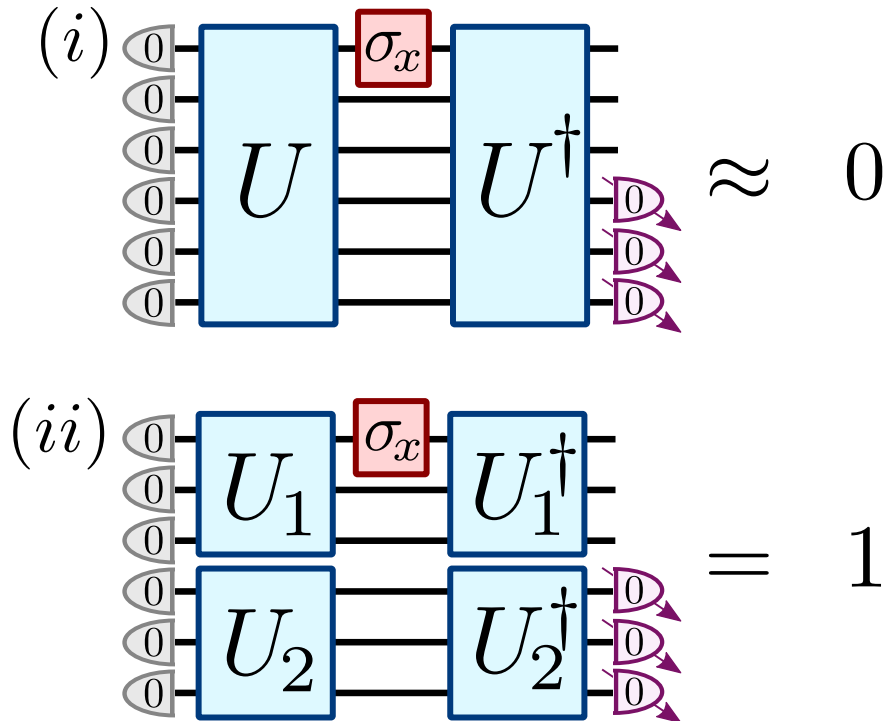
Figure 12.5:   Solution to the disjoint unitary problem with out-of-time-order measurements.
The state $|0\rangle^{\otimes n}$ is prepared and the unknown unitary (either $U$ or $U_1 \otimes U_2$) is applied. Next
$\sigma_x$ is applied to the first qubit, followed by the inverse of the unknown unitary. Finally, it is
checked if the second block of $n/2$ qubits ends up in the all zero state. If so, then the hidden
unknown unitary is $U_1 \otimes U_2$ as per case $(ii)$; if not, then the unknown unitary is $U$ as per
case $(i)$.

imperfect time-reversal dynamics, each of which disrupt the non-local correlations probed
by the OTOC [21, 139, 326, 416, 427, 450]. While this disruption can be mitigated via
independent error estimates [326, 450], for sufficiently large errors these estimates involve
measuring quantities of small magnitude (comparable to the TOC), squandering the OTOC's
learning advantage. In Appendix G.2 we estimate that our previous results are modified in
the presence of a small local error rate $\varepsilon \ll J$ as follows: in the first learning regime, the
OTOC maintains its advantage up to distances $d \lesssim J/\varepsilon$; in the second regime, the $L$-fold
advantage is replaced by a $(\min\{L, \sqrt{J/\varepsilon}\})$-fold advantage.

In practice, we find that learning via OTOCs remains robust even to relatively large
amounts of imperfect time-reversal [Fig. 12.3(d)]. We study this numerically in the "weak
interaction" learning problem of Fig. 12.3(b). As a concrete instance of imperfect time-
reversal, we take the spins to be coupled to an extrinsic cavity mode and assume that the
spin dynamics are perfectly reversed but the cavity dynamics and spin-cavity coupling $g$ are

unreversed. We find that access to OTOCs substantially improves the classification accuracy even for quite large spin-cavity couplings $g \sim 0.5$, up to half the spin-spin interaction strength.

We can also examine the dependence of learning on read-out errors, namely where one measures a correlator $C$ up to additive error. Indeed, we have already incorporated a realistic read-out error $\delta = 3\%$ in our previous numerical studies [Figs. 12.2(a), 12.3(a), 12.4(a)]. Intuitively, we expect larger read-out errors to make learning more difficult; however, we have little reason to expect read-out error to change the *relative* advantage of OTOCs compared to TOCs. We test this numerically by repeating the analysis of Fig. 12.3(b) for various read-out errors, $\delta$. For each $\delta$, we compute the minimum link strength $J_\ell^*$ that can be learned with $> 90\%$ accuracy [Fig. 12.3(c)]. For errors $\delta \gtrsim 10^{-3}$, our results agree well with analytic estimates, which predict $(J_\ell^*/J)^2 \sim \delta$ for TOCs and $(J_\ell^*/J)^2 \sim \delta/L$ for OTOCs. Intriguingly, for sufficiently small errors $\delta \lesssim 10^{-3}$, the minimum link strength detectable with TOCs *saturates* to a finite value $J_\ell^* \sim 0.2$. Below this value, sample-to-sample fluctuations of the TOC cause the learning task to be difficult regardless of the read-out error.

## 12.5   Provable learning advantage

We have so far demonstrated the learning power of OTOCs using phenomenological arguments and numerical simulations, owing to the difficulty of obtaining analytic results for ergodic Hamiltonian systems. Complementary to these results, we now introduce a binary classification task in which the OTOC is provably efficient. The task is as follows:

**Disjoint unitary problem:** One is given oracle access to either: $(i)$ a fixed, $n$-qubit Haar-random unitary $U$, or $(ii)$ a tensor product of two fixed, $n/2$-qubit Haar-random unitaries, $U_1 \otimes U_2$. The task is to determine which of (i) or (ii) is realized.

Qualitatively, this problem resembles the Hamiltonian learning scenarios identified previously. First, the feature we seek to learn—the connectivity of the unitary—directly determines how information spreads through the system, as measured by the OTOC. Second, a Haar-random unitary is inherently "strongly-interacting", which causes time-ordered measurements to decay and thus provide little information.

In Fig. 12.5 we show that the disjoint unitary problem can be solved with a constant number (with respect to $n$) of queries to the oracle *and its time-reverse $U^\dagger$*, by measuring an out-of-time-order observable. Letting $V$ denote the unknown unitary (either $U$ or $U_1 \otimes U_2$), the OTOC is

$$\mathrm{OTOC}(V) = \mathrm{tr}\left(\mathbb{1}_{\frac{n}{2}} \otimes |0\rangle\langle 0|^{\otimes \frac{n}{2}} \left\{ V^\dagger \sigma_x^1 V |0\rangle\langle 0|^{\otimes n} V^\dagger \sigma_x^1 V \right\}\right).$$

In case $(i)$, the OTOC is near zero with probability exponentially close to one [111]. In case $(ii)$, the OTOC is one, since the two subsystems are not coupled by $U_1 \otimes U_2$. Thus, with probability exponentially close to one, the two cases may be distinguished with a single

query to the unknown unitary and its time-reverse. In contrast, in the following Chapter, we prove that any time-ordered learning protocol requires an exponential number $\Omega(2^{n/4})$ of queries of the unknown unitary to solve the disjoint unitary problem. Our proof applies even to adaptive measurement strategies, and leverages novel contemporary techniques from quantum learning theory [10, 96, 220].

## 12.6 Discussion and outlook

In this Chapter, we have shown that quantum information dynamics can provide powerful advantages for learning properties of quantum systems. As we have seen in Part III, the physics of quantum information dynamics is intimately linked with experimental protocols that involve time-reversal. (In the following Chapter we rigorize this connection.) Our results thus highlight the potential gains that can be achieved by quantum experiments if they have sufficient control and coherence to apply time-reversed dynamics. Extraordinary experimental progress has led to an ever-increasing number of such platforms [61, 139, 163, 274, 289, 326, 416], and we envision that learning via OTOCs might find applications across these diverse physical contexts. Specific future directions include learning long-range cross-talk in quantum processors [420], and strongly-interacting problems in NMR [354].

# Chapter 13

# Information-theoretic hardness of out-of-time-order correlators

In this Chapter, we establish that there are properties of quantum many-body dynamics which are efficiently learnable if we are given access to out-of-time-order correlators (OTOCs) [250, 277, 315], but which require exponentially many operations in the system size if we can only measure time-ordered correlators. In the context of quantum learning, this extends the phenomenological and numerical results of Ch. 13 to a rigorous setting. Our result also provides support to a general theme observed in Part III, where experimental protocols to measure quantum information dynamics always required the ability to reverse time mid-experiment [46, 61, 139, 163, 233, 274, 289, 326, 381, 416, 451, 473, 512, 518]. In particular, we show that in certain cases, any experimental protocol which reconstructs OTOCs solely from time-ordered correlators must be, in certain cases, exponentially inefficient.

Our proofs build upon and generalize recent work on quantum algorithmic measurements (QUALMs) and quantum learning theory [10, 95, 96, 219]. The formulation of our results entails a precise definition of the most general time-ordered and out-of-time-order correlators that can be measured; this may be of interest for other applications in quantum learning theory.

The remainder of the Chapter is organized as follows. In Section 13.1 we formulate the most general time-ordered experiments, and the most general out-of-time-order experiments. In Section 13.2 we explain our main results on the hardness of measuring certain properties of quantum many-body systems using time-ordered operations alone. We conclude in Section 13.3 with a discussion.

## 13.1 Time-ordered and out-of-time-order experiments

In this section we mathematically formalize how correlators are extracted from measurements of a system. While correlators are natural objects in quantum systems and field theories, they are often studied abstractly without acknowledgement of how they might be

measured in a physical system. The question of devising a measurement protocol to extract a particular correlator from a system of interest is particularly pressing in the case of OTOCs. Below we will use the framework of learning theory to provide a general definition of how correlators can be obtained via quantum measurements. We leverage the learning tree formalism for quantum channels, developed in [10, 96]. Let us outline an intuitive understanding of how such experiments operate, and then render this into more precise definitions.

Suppose we have some experimental system with time evolution by a unitary $U$, which is not known to or fully characterized by the experimentalist. The experimentalist desires to learn about $U$ by making measurements on the system as it evolves. For a Hamiltonian system, we might have $U = e^{-iH\Delta t}$ for some not fully characterized $H$, where $\Delta t$ is the shortest time scale over which we can control the evolution. So if the experimentalist wants to evolve the system by a time $k\Delta t$, he can simply apply $U^k$. (Our formulation will also work if the experimentalist has continuous control over the time $t$, but this $\Delta t$ discretization will make our definitions simpler to state.) To be explicit, we stipulate that the system in question is composed of $n$ qubits on which the unitary $U$ acts.

An experiment would operate as follows. The experimentalist begins by preparing the system in some initial state $\rho_0$. Thereafter, he can choose to either: (i) apply $U$; (ii) apply some other quantum channel; or (iii) perform a partial or complete measurement, which would confer some classical information about the state of the system which he could store on a classical computer. He can exercise these options again and again in a sequence, each time basing his decision of what to do next on the information collected thus far. That is, the protocol for information collection can be *adaptive*. At the end of the experiment, the classical computer contains the information the experimentalist has gained by performing measurements at any stage throughout the protocol.

We will make the assumption that throughout the protocol, the state of the system is not entangled with any external ancilla system which the experimentalist can manipulate. This choice is made to reflect contemporary experimental realities; for instance, at present, there is no way of entangling a sample of graphene to an external quantum computer. In the Discussion, we will comment further on the possibility of ancilla-assisted protocols.

Next we turn to formalizing the notion of an experiment explained above. First we note that in our setting, the most general operation the experimentalist can perform on a quantum state is a POVM measurement [353]. That is, consider a collection of operators $\{F_i\}_i$ on $n$ qubits, satisfying $\sum_i F_i^\dagger F_i = \mathbb{1}$. Then the POVM measurement with respect to these operators maps

$$\rho \longmapsto \frac{F_i \rho F_i^\dagger}{\mathrm{tr}(F_i \rho F_i^\dagger)} \quad \text{with probability } \mathrm{tr}(F_i \rho F_i^\dagger). \tag{13.1}$$

Here the experimentalists' apparatuses would register that he had measured the $i$th outcome. This generalizes the notion of a projective measurement from elementary quantum mechanics. Note that a special case of a POVM measurement is simply the application of a unitary $V$; if the POVM is the singleton set $\{V\}$ which clearly satisfies $V^\dagger V = \mathbb{1}$, then 'measuring' $\rho$ yields $V \rho V^\dagger$ with probability one.
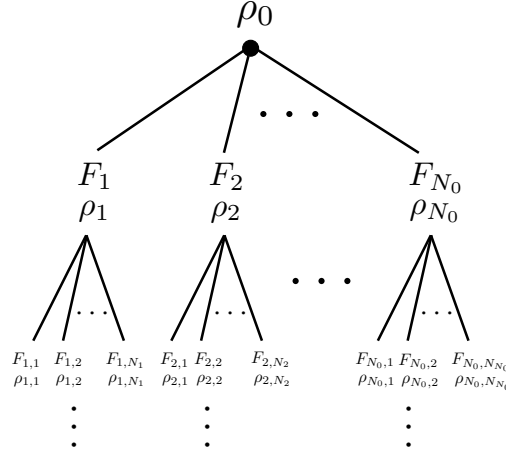
Figure 13.1: *Depiction of a learning tree $\mathcal{T}$.* We begin with a state $\rho_0$ at the root of the tree, and perform successive, adaptive POVM measurements. A root-to-leaf path through the tree corresponds to a sequence of POVM measurement outcomes.

The reason that POVM measurements are so general is encapsulated in the following fact: any composition of quantum channels and POVM measurements can be captured by a *single* new POVM measurement. That is, suppose we have a quantum state and suscept it to a sequence of POVM measurements and quantum channels; then the result of this is the same as having applied some über-POVM measurement.

With the above in mind, we can conceive of an experimental protocol as occurring in a sequence of rounds. The protocol is as follows:

- Initialize $\rho_0$.

- Apply $U$, measure the state using a POVM $\{F_i\}_i$. Suppose the outcome is $i = q$; then store this in the classical memory. The output is the state $\rho_q := \frac{F_q U \rho_0 U^\dagger F_q^\dagger}{\mathrm{tr}(F_q U \rho_0 U^\dagger F_q^\dagger)}$.

- Apply $U$, measure the state using a POVM $\{F_{q,i}\}_i$, which can be contingent on the previous measurement outcome $q$. Suppose the new outcome is $i = r$; then store this in the classical memory. The output is the state $\rho_{q,r} := \frac{F_{q,r} U \rho_q U^\dagger F_{q,r}^\dagger}{\mathrm{tr}(F_{q,r} U \rho_q U^\dagger F_{q,r}^\dagger)}$.

- Apply $U$, measure the state using a POVM $\{F_{q,r,i}\}_i$, which can be contingent on the previous measurement outcomes $q, r$. Suppose the new outcome is $i = s$; then store this in the classical memory. The output is the state $\rho_{q,r,s} := \frac{F_{q,r,s} U \rho_{q,r} U^\dagger F_{q,r,s}^\dagger}{\mathrm{tr}(F_{q,r,s} U \rho_{q,r} U^\dagger F_{q,r,s}^\dagger)}$.

- Repeat this kind of adaptive POVM measurement procedure for $T$ total rounds.

Here, the (adaptive) sequence of POVMs only 'knows' about $U$ via the sequence of measurement outcomes.

Several further comments are in order. First, note that a POVM measurement is not applied between the initial preparation of $\rho_0$ and the initial application $U$; this would be superfluous since it is equivalent to having prepared a different initial state. Second, this protocol is clearly adaptive, since the choice of each POVM measurement can be contingent on all previous measurement outcomes. Indeed, the learning tree specifies an adaptive strategy since it prescribes how the experimentalist makes his adaptive choices. Third, observe that if the experimentalist wanted to apply $k$ $U$'s in a row, i.e. $U^k$, he could simply choose for the POVMs to be $\{\mathbb{1}\}$ for $k$ rounds in a row. Finally, we observe that the protocol outlined above is only directly capturing *time-ordered* correlations, since the experimentalist can apply $U$ but not $U^\dagger$.

This class of protocols can be fruitfully organized into a tree, as per Figure 13.1. We start at the root of the tree (i.e. the top-most vertex), and traverse down the tree by successively performing POVM measurements in an adaptive fashion. We see, then, that a particular instantiation of the protocol is a root-to-leaf (i.e. top-to-bottom) path through the tree. A tree of depth $T$ corresponds to applying $U$ a total of $T$ times, i.e. once per round. The classical information that the experimentalist obtains is the sequence of POVM measurement outcomes, which corresponds to a root-to-leaf path through the tree. A path is labelled by a sequence of vertices $v_0, v_1, ..., v_T = \ell$, or more simply by $\ell$ since the leaf node specifies the entire root-to-leaf path.

Let us denote such a learning tree by $\mathcal{T}$. It represents a specification of an adaptive experimental protocol that an experimentalist can perform. We provide more formal details in Appendix H. Now if $v_0, v_1, ..., v_T = \ell$ is a root-to-leaf path through $\mathcal{T}$, then the probability of taking that path is

$$p^U(\{v_t\}) := \prod_{t=1}^{T} \mathrm{tr}(F_{v_t} U \rho_{v_{t-1}} U^\dagger F_{v_t}^\dagger), \qquad (13.2)$$

which can be more conveniently notated by $p^U(\ell)$. In other words, this is the probability of the experimentalist obtaining the sequence of measurement outcomes given by the root-to-leaf path through the tree terminating in $\ell$. The way that information is extracted from an experiment is via a function $G(\ell)$ which maps the sequence of measurement outcomes to the value of some desired quantity, e.g. a time-ordered correlator. The empirical expectation value of $G(\ell)$ is then $\hat{G} := \mathbb{E}_{p^U(\ell)}[G(\ell)] = \sum_\ell p^U(\ell) G(\ell)$.

The above motivates the following definition of a time-ordered experiment for learning properties of $U$, which we further detail in Appendix H:

**Definition 1** (Time-ordered experiment)**.** *A time-ordered experiment is any learning tree protocol $\mathcal{T}$ which queries $U$.*

The definition of an out-of-time-order experiment follows in a similar fashion:

**Definition 2** (Out-of-time-order experiment)**.** *An out-of-time-order experiment is any learning tree protocol $\mathcal{T}'$ which queries both $U$ and $U^\dagger$, where the choice of which one is to be queried in each round can be determined adaptively.*

Now suppose we want to measure an OTOC such as $\text{tr}(\rho_0 U^\dagger W U V U^\dagger W U V)$. Clearly this is most accessible with an out-of-time-order experiment. However, we emphasize that we can obtain this OTOC using the data of a time-ordered experiment, although we might require many more rounds of the experiment to obtain the answer to within the desired precision.

Indeed, our goal in next section is to establish that if we do not fully know $U$ (or $U^\dagger$), then there are certain OTOCs which are readily and efficiently attained by an out-of-time-order experiment, but which require exponentially many operations if the experiments are time-ordered.

## 13.2 Information-theoretic hardness of OTOCs

In this section we explain our main result, namely that for quantum many-body systems with partially unknown dynamics, there can be OTOCs which are easy to measure with out-of-time-order experiments but which are exponentially hard to measure with only time-ordered experiments. Said differently, any experimental protocol that reconstructs OTOCs from only time-ordered experiments must in some cases be exponentially inefficient. In this manner, our results elucidate fundamental differences between OTOCs and time-ordered correlators.

Our proof strategy is to construct an explicit example for which measuring an OTOC to within constant error has an exponential disparity between the time-ordered and out-of-time-ordered experimental settings. Concretely, consider again an $n$ qubit system, here for $n$ even, equipped with a partially uncharacterized unitary. Suppose that it is either: (i) a fixed, Haar-random unitary $U$ on $n$ qubits, or (ii) a product $U_1 \otimes U_2$ of two fixed, Haar-random unitaries $U_1, U_2$, each on $n/2$ qubits. Here $U_1$ is to act on the first $n/2$ qubits, and $U_2$ is to act on the remaining $n/2$ qubits. The experimentalist will not know which of these two possibilities (i) or (ii) is the case, and is tasked with performing an experimental protocol to determine which one is instantiated.

The two possibilities are physically rather different. In (i) all of the qubits interact with one another, whereas in (ii) only blocks of half of the qubits mutually interact. This suggests that if the experimentalist can perform an out-of-time-order experiment, it is quite easy to distinguish between (i) and (ii) by measuring a single OTOC. This works in the following way. The experimentalist prepares the system in the all zero state $|0\rangle^{\otimes n}$, and then applies the unknown unitary. Thereafter, the experimentalist applies $\sigma_x$ on the first qubit to flip it, followed by applying the inverse of the unknown unitary. Then the experimentalist checks if the second block of $n/2$ qubits is again in the all zero state. This corresponds to measuring the OTOC

$$\mathsf{OTOC}(V) = \text{tr}\left(\mathbb{1}_{\frac{n}{2}} \otimes |0\rangle\langle 0|^{\otimes \frac{n}{2}} \left\{ V^\dagger \sigma_x^1 V |0\rangle\langle 0|^{\otimes n} V^\dagger \sigma_x^1 V \right\} \right), \tag{13.3}$$

where $V$ is a placeholder for the unknown unitary. In case (i), the final output state will be complicated, having little overlap with the all zero state. Indeed, on average we have

$$\mathbb{E}_{U \sim \text{Haar}(2^n)}[\mathsf{OTOC}(U)] = \frac{2^{\frac{3n}{2}} - 1}{2^{2n} - 1} \leq O(1/2^{n/2}). \tag{13.4}$$
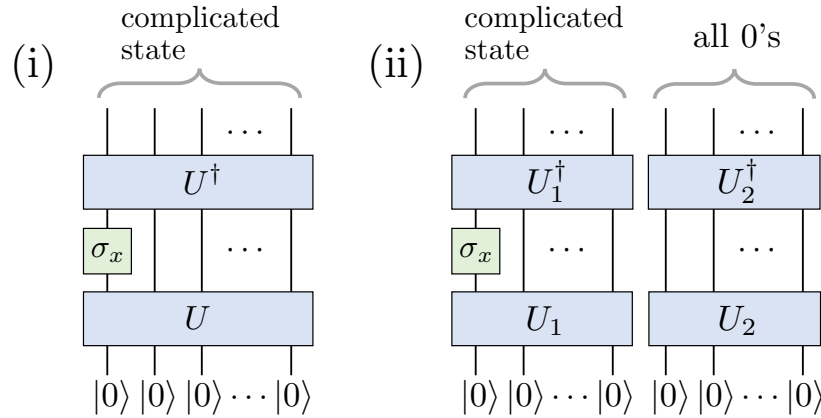
Figure 13.2: *Schematic of out-of-time-order experiment.* (i) In the first case, applying $U$, then $\sigma_x^1$, and then $U^\dagger$ results in a complicated state. (ii) In the second case, applying $U_1 \otimes U_2$, then $\sigma_x^1$, and then $U_1^\dagger \otimes U_2^\dagger$ leads to a product of two pure states, each on $n/2$ qubits; the first is complicated, the second is the all zero state.

However, in case (ii) the $\sigma_x^1$ operator still allows $U_2$ to cancel with $U_2^\dagger$, and so the second block of $n/2$ qubits ends up precisely in the all zero state. In terms of the OTOC correlator, we have

$$\mathsf{OTOC}(U_1 \otimes U_2) = 1 \quad \text{for all} \quad U_1, U_2. \tag{13.5}$$

These results are illustrated in Figure 13.2.

More formally, these results have the following consequence:

**Theorem 1** (Easiness of task with out-of-time-order experiment). *If the experimentalist can perform an out-of-time-order experiment, then with probability exponentially close to one the cases (i) and (ii) can be distinguished using only a single application of the unknown unitary, and a single application of its inverse.*

*Proof.* In case (i), the probability that the OTOC is less than or equal to a small constant $\varepsilon$ is bounded by Markov's inequality, namely

$$\mathrm{Prob}[\mathsf{OTOC}(U) \leq \varepsilon] \geq 1 - \frac{\mathbb{E}_{U \sim \mathrm{Haar}(2^n)}[\mathsf{OTOC}(U)]}{\varepsilon}$$
$$\geq 1 - O(1/(\varepsilon \, 2^{\frac{n}{2}})), \tag{13.6}$$

where we have used (13.4). In case (ii), the probability that $\mathsf{OTOC}(U_1 \otimes U_2)$ is greater than $\varepsilon$ is one, on account of (13.5). Thus the two possibilities can be distinguished with probability exponentially close to unity, and the protocol only requires a single query of the unknown unitary and a single query of its inverse. $\qquad \square$

This result lies in contrast to the following, more difficult theorem:

**Theorem 2** (Exponential hardness of task with time-ordered experiment). *Any time-ordered experiment which can distinguish cases (i) and (ii) requires at least $\Omega(2^{n/4})$ queries of the unknown unitary, and so is exponentially inefficient.*

An equivalent rephrasing is our promised result:

**Theorem 2′ 1** (Exponential hardness OTOCs with time-ordered experiments). *Any time-ordered experimental protocol to determine an OTOC to within constant error must in certain instances require accessing the time evolution exponentially many times. However, in some of these instances, an out-of-time-order experiment can determine the OTOC to within constant error by accessing the time evolution only a constant number of times.*

The theorems should be regarded as quantifying a form of information-theoretic hardness, since they bound the number of applications of the unknown unitary that we need to succeed in the time-ordered setting. While the proof of hardness is given in Appendix H, we sketch its high-level strategy here.

The idea, coming from previous work [10, 96], is to upper bound the sum

$$\sum_{\ell \in \text{leaf}(\mathcal{T})} \left| \mathop{\mathbb{E}}_{U \sim \text{Haar}(2^n)} [p^U(\ell)] - \mathop{\mathbb{E}}_{U_1, U_2 \sim \text{Haar}(2^{n/2})} [p^{U_1 \otimes U_2}(\ell)] \right| \tag{13.7}$$

for any time-ordered experiment $\mathcal{T}$. We would like to show that this quantity is $o(1)$ if the number of applications of the unknown unitary is less than $o(2^{n/4})$. This would imply that the probability of distinguishing (i) and (ii) can only reach a constant value (i.e. one that is not suppressed in $n$) if we apply the oracle exponentially many times in $n$. Intuitively, upper bounding (13.7) by a small number means that the probability distribution over measurement outcomes looks extremely similar regardless of whether case (i) or case (ii) is instantiated; this means that the two cases cannot be distinguished.

Operationally, we show that $\mathbb{E}_{U \sim \text{Haar}(2^n)}[p^U(\ell)]$ and $\mathbb{E}_{U_1, U_2 \sim \text{Haar}(2^{n/2})}[p^{U_1 \otimes U_2}(\ell)]$ are each close to the uniform distribution $\frac{1}{2^{nT}}$, and hence close to one another via the triangle inequality. To establish closeness to the uniform distribution, we rewrite the Haar averages in terms of sums of correlators involving permutation operators via the Weingarten calculus. In each case, one of the correlators is exponentially close to $\frac{1}{2^{nT}}$, and intricate algebraic manipulations establish that the remaining correlators are further suppressed by factors of the Hilbert space dimension. The required technical tools including Haar integration, Weingarten functions, and the learning tree formalism are provided in Appendix H. Our main proofs are presented in Appendix H.

## 13.3 Discussion and outlook

In this Chapter we have given a precise framework for defining and analyzing time-ordered versus out-of-time-order experiments, and established that the latter have an exponential ad-

vantage over the former for measuring certain OTOCs. Our methods advance recent developments in quantum learning theory [10, 95, 96, 216, 219], and extend the phenomenological results presented in Ch. 12.

An interesting future direction is to generalize our results to more realistic settings, e.g. when the unitaries in question are not constructed from Haar-random ensembles (which are exponentially complex to realize physically). There has been progress in this vein for a related class of learning problems involving states instead of unitaries, e.g. [96, 219]. To this end, it may be fruitful to borrow intuition from the learning problems introduced in Ch. 12.

We also emphasize that in this Chapter we have made the physically reasonable assumption that we cannot entangle our system to ancillas which could act as a quantum memory. Indeed, in certain cases adding ancillas could equalize the power balance between time-ordered and out-of-time-order experiments for certain OTOC learning tasks. While we do not expect this equalization to hold generally, the tradeoff between quantum memory and the ability to reverse time is worthy of further investigation.

There has been previous work on the difficulty of simulating the Hermitian conjugate of a unitary $U$ given only black box access to $U$ [382, 383]; our approach in the present work is different, since we instead consider experiments for learning *properties* of $U$. However, our results and techniques may interface in interesting ways with this line of previous work, for instance establishing new hardness results. We note that our Theorem 2 implies that in the worst case it is exponentially hard to construct the inverse of a unknown unitary $U$ for which one has query access; this is consistent with [382, 383].

As a conceptual coda to our results, we remark that in our own universe we do not have the ability to reverse the direction of time. As such, there may be physically interesting features of nature, such as ones pertaining to quantum chaos, which are effectively inaccessible to us. This is also true of experimental systems in which we cannot, in practice, reverse the direction of their time evolution. In the latter case, we may one day be able to exercise the option of simulating that physical system on a quantum computer, and so time-reversal becomes available. Thus the ability to control the flow of time evolution in a quantum computer may ultimately allow us to unlock hidden properties of natural systems around us.

# Part V

# Appendices

# Appendix A

# Details on Floquet Hopf insulators

In this Appendix we provide additional theoretical and numerical details on the Floquet Hopf insulator (discussed in Ch. 5) and its behavior during adiabatic transformations.

## A.1   Twisting interpretation of the topological invariants

We elaborate on the twisting interpretations of the static and Floquet Hopf invariants. Formally, these derive from the framed-cobordism understanding of the homotopy groups $\pi_3(S^2)$ and $\pi_4(S^3)$ in mathematics [372]. For the static Hopf invariant, this understanding was related to the conventional linking interpretation by Ref. [243], which we briefly summarize. Consider two infinitesimally close points $\mathbf{n}$ and $\mathbf{n} + \delta\mathbf{n}$ on the Bloch sphere. Their pre-images will also be infinitesimally close, and we may parameterize them by the periodic parameter $\theta$, $\mathbf{k}(\theta)$ and $\mathbf{k}(\theta) + \delta\mathbf{k}(\theta)$ [where $\theta$ interpolates between 0 and $2\pi$ along the pre-image]. By calculus, the difference $\delta\mathbf{k}(\theta)$ between the two pre-images is determined by the Jacobian $J_{ij}(\theta) \equiv \partial_{k_i} n_j(\mathbf{k})|_\theta$ along the first pre-image, according to $J(\theta)\delta\mathbf{k}(\theta) = \delta\mathbf{n}$. Crucially, this implies that all information about the pre-images' linking, and thus the Hopf invariant, can be found from the Jacobian along a single pre-image! The Jacobian is a $2 \times 3$ matrix (as $\mathbf{n}$ is normalized, only derivatives with respect to its two orthogonal components are nonzero), which generically has two independent non-parallel column vectors. As shown in Fig. 5.1, we may choose one of these vectors to point along the direction to the second pre-image, such that the vector 'traces out' the second pre-image as it varies along the first pre-image. As seen in the figure, the total rotation, or *twisting*, of the column vectors along the pre-image is precisely the linking number of the pre-images. To formalize this, we view the two vectors as those of a real orthogonal matrix in $SO(2)$ (the reduction of the $2 \times 3$ matrix $J_{ij}(\theta)$ to a $2 \times 2$ matrix is proper, because the derivative $\partial_{k_\parallel} n_j(\mathbf{k})|_\theta$ along the pre-image is always zero). The twisting of this matrix along the 1-dimensional pre-image is classified by the homotopy group $\pi_1(SO(2)) = \mathbb{Z}$, agreeing with the original $\pi_3(S^2)$ classification.

This twisting interpretation generalizes to the Floquet invariant much better than the linking number interpretation. Although the pre-images of $SU(2)$ in the 4D Floquet Bril-

louin zone are again 1D loops (because $SU(2)$ is three-dimensional), there is no concept of linking number in 4D, as any two loops may be smoothly unlinked by moving in the fourth dimension. Nevertheless, one may consider the Jacobian along the pre-image loop, now a $3 \times 4$ matrix with three independent non-parallel column vectors. To analyze the twisting of the Jacobian along the loop, we view these three vectors as the columns of a real orthogonal matrix in $SO(3)$. The twisting of this matrix along the 1D loop is characterized by the group $\pi_1(SO(3)) = \mathbb{Z}_2$, reproducing the Floquet invariant's classification. The reduced classification indicates that any even twisting of the Jacobian may be smoothly deformed to a constant Jacobian, and is familiar from characterizing particle statistics (bosons vs. fermions) in three-dimensions. We see that the reduced $\mathbb{Z}_2$ invariant is the natural result of extending the static Hopf invariant to higher dimensions.

We can verify the twisting interpretation for the stroboscopic evolution described in Ch. 5. To do so, we first write down the micromotion operator for the stroboscopic evolution. The micromotion operator closes the bands of the Floquet unitary, and is homotopy equivalent to the flat band unitary:

$$U_m(\mathbf{k}, t) = \begin{cases} e^{-\frac{2\pi i t}{T} H_{h_S - h_F}(\mathbf{k})} & 0 \leq t < T/2 \\ -e^{\frac{2\pi i (t - T/2)}{T} H_{h_S}(\mathbf{k})} & T/2 \leq t < T \end{cases}. \tag{A.1}$$

We'll take $h_F = 1 \bmod 2$, $h_S = 0$, and choose the trivial insulator to be the product state, $H_{h_S=0}(\mathbf{k}) = \sigma_z$. Consider the pre-image of fixed $\hat{\mathbf{n}} \neq \hat{\mathbf{z}}$ and $\phi$. This occurs at fixed time-slice $t = \phi T/2\pi$, with $\mathbf{k}$ determined by the pre-image of $\hat{\mathbf{n}}$ under the static Hamiltonian $H_{h_S - h_F = 1}(\mathbf{k})$. Compared to the static case, the Jacobian along a pre-image has one new row, corresponding to derivatives of the quasienergy $\phi$, and one new column, for derivatives with respect to time $t$. In the flat band micromotion operator $\partial_{k_i} \phi = 0$ and $\partial_t \hat{\mathbf{n}} = 0$, and so the time column is nonzero only on the quasienergy row and is orthogonal to both other columns. The Jacobian twisting is therefore determined entirely by the columns corresponding to $\hat{\mathbf{n}}$, and so is the same as the twisting for the static Hamiltonian $H_{h_S - h_F}(\mathbf{k}) = H_1(\mathbf{k})$: equal to the Floquet invariant $h_F = 1$.

## A.2 Derivation of the Floquet invariant

In this section, we derive an explicit formula for the Floquet Hopf invariant in terms of the eigenvalues and eigenvectors of the micromotion operator. This allows for straightforward numerical calculation of the invariant, and, at the end of this section, allows us to rigorously derive the equality between the Floquet invariant and the quasienergy-$\pi$ topological defect charge.

The Floquet Hopf invariant characterizes elements of the homotopy group $\pi_4(SU(2)) = \mathbb{Z}_2$. Witten introduced a method to compute this invariant [499]. Unfortunately, this method can be somewhat opaque, and unwieldy for numerics, in its original formulation. Our goal here is to re-express Witten's formula for the invariant in a more convenient, and explicit, form.

Witten's method relies on the fact that the fourth homotopy group of three-band unitaries, $\pi_4(U(3))$, is trivial, in contrast to the homotopy group of interest, of $SU(2)$ [or, equivalently, $U(2)$]. The prescription to compute the invariant proceeds as follows. First, embed the micromotion unitary $U_m(\mathbf{k}, t)$ in $U(3)$ by adding a third, trivial band (i.e. with no $\mathbf{k}, t$ dependence):

$$\tilde{U}_m(\mathbf{k}, t) \equiv \begin{pmatrix} U_m(\mathbf{k}, t) & \begin{matrix} 0 \\ 0 \end{matrix} \\ 0 \quad 0 & 1 \end{pmatrix}. \tag{A.2}$$

Next, contract the three-band unitary $\tilde{U}_m(\mathbf{k}, t)$ smoothly to the identity. Specifically, find a continuous family of unitaries $\tilde{U}(\mathbf{k}, t, \rho)$, labelled by contraction parameter $\rho$, such that $\tilde{U}(\mathbf{k}, t, \rho = 0) = \tilde{U}_m(\mathbf{k}, t)$ and $\tilde{U}(\mathbf{k}, t, \rho = \pi/2) = \mathbb{1}, \forall \mathbf{k}, t$. Prior to embedding in $U(3)$, such a contraction is forbidden if the Hopf defect at $\pi$-quasienergy has a nontrivial topological charge. After embedding, the contraction becomes possible due to the triviality of the Hopf insulator in three-band systems [or, equivalently, due to the trivial homotopy group $\pi_4(U(3))$]. Remarkably, the Floquet Hopf invariant of the *original* two-band micromotion operator can be computed from the entire family of three-band unitaries along the contraction, $\tilde{U}(\mathbf{k}, t, \rho)$, via the five-dimensional integral

$$h_F = \frac{-i}{240\pi^3} \int_0^{\pi/2} d\rho \int dt \, d^3\mathbf{k} \sum_{\substack{ijkln \in \\ k_x, k_y, k_z, t, \rho}} \epsilon^{ijkln} \text{Tr}\left(\tilde{U}^{-1}\partial_i\tilde{U}\tilde{U}^{-1}\partial_j\tilde{U}\tilde{U}^{-1}\partial_k\tilde{U}\tilde{U}^{-1}\partial_l\tilde{U}\tilde{U}^{-1}\partial_m\tilde{U}\right).$$

$$\tag{A.3}$$

The invariant $h_F$ is quantized to an integer modulo 2, and is independent of the deformation chosen.

To see that this integral is a topological invariant of the original unitary, we review an argument from Ref. [499]. Consider two micromotion operators $U_m(\mathbf{k}, t)$ and $U'_m(\mathbf{k}, t)$. If the two operators are homotopy equivalent, there exists a smooth deformation of one into the other. To contract $\tilde{U}'_m(\mathbf{k}, t)$ to the identity after embedding in $SU(3)$, one can therefore first deform the embedded $\tilde{U}'_m(\mathbf{k}, t)$ to the embedded $\tilde{U}_m(\mathbf{k}, t)$, and then contract $\tilde{U}_m(\mathbf{k}, t)$ to the identity. Crucially, the first step necessitates only matrices of the form

$$\begin{pmatrix} U(\mathbf{k}, t; \rho) & \begin{matrix} 0 \\ 0 \end{matrix} \\ 0 \quad 0 & 1 \end{pmatrix}, \tag{A.4}$$

since the two unitaries are homotopy equivalent in $SU(2)$. However, one can show that the integrand of Eq. (A.3) is identically zero for $\tilde{U}$ of this form. Therefore the second part of the contraction is the sole contributor to the Floquet Hopf invariant, and so the two unitaries have the same invariant. To show that the integrand indeed vanishes for all $SU(2)$ matrices, one can express the integral in terms of eigenvectors and eigenvalues of the unitary, as we do for $SU(3)$ unitaries below, and use the fact that the eigenvectors have normalization 1.

We can also show that the parity of invariant calculated is independent of the contraction chosen. Consider two different contractions, $\tilde{U}_1(\mathbf{k}, t, \rho)$ and $\tilde{U}_2(\mathbf{k}, t, \rho)$, of the same micromotion operator. Each contraction is a map from the domain $T^4 \times [0, 1]$ to $SU(3)$. One can 'sandwich' together the two contractions to define a new map,

$$\tilde{U}_{12}(\mathbf{k}, t, \rho) = \begin{cases} \tilde{U}_1(\mathbf{k}, t, -\rho) & -\pi/2 \leq \rho \leq 0 \\ \tilde{U}_2(\mathbf{k}, t, \rho) & 0 \leq \rho \leq \pi/2 \end{cases}, \tag{A.5}$$

which is now periodic in $\rho$ with period $\pi$, and therefore maps instead from the domain $T^5$ to $SU(3)$. Since the domain is now closed, we can consider the homotopy equivalence classes of such maps. Neglecting weak invariants, these are characterized by the homotopy group $\pi_5(SU(3)) = \mathbb{Z}$. This integer invariant $\Gamma \in \mathbb{Z}$ can be calculated [67]:

$$\Gamma = \frac{-i}{480\pi^3} \int_{-\pi/2}^{\pi/2} d\rho \int dt \, d^3\mathbf{k} \sum_{\substack{ijkln \in \\ k_x, k_y, k_z, t, \rho}} \epsilon^{ijkln} \text{Tr}\big(\tilde{U}^{-1} \partial_i \tilde{U} \tilde{U}^{-1} \partial_j \tilde{U} \tilde{U}^{-1} \partial_k \tilde{U} \tilde{U}^{-1} \partial_l \tilde{U} \tilde{U}^{-1} \partial_n \tilde{U}\big).$$
$$\tag{A.6}$$

This integral is easily related to the Floquet Hopf invariants of the two contractions,

$$2\Gamma = h_{F,1} + h_{F,2}. \tag{A.7}$$

Taking both sides modulo 2 gives our desired equivalence,

$$h_{F,1} \bmod 2 = h_{F,2} \bmod 2, \tag{A.8}$$

proving that the Floquet Hopf invariant calculated is independent of the contraction. One can also see from Eq. (A.7) that $h_F$ is quantized to be an integer: set the second contraction to be the same as the first one, i.e. $h_{F,1} = h_{F,2} \equiv h_F$, which then gives $h_F = \Gamma$, an integer.

We now turn to recasting the above formula in terms of the eigenvectors $|z_m\rangle$ and quasienergies $\phi_m$ of the *original* micromotion operator $U_m(\mathbf{k}, t)$. We do this by writing down an explicit contraction (valid for an arbitrary micromotion operator), and using it to simplify Eq. (A.3). The embedding and contraction are depicted in Fig. A.1. The embedded micromotion operator has three bands: an upper band with eigenvector $\big(z_{m,1}(\mathbf{k}, t), z_{m,2}(\mathbf{k}, t), 0\big)^T$, a lower band with orthogonal eigenvector $\big(-z_{m,2}^*(\mathbf{k}, t), z_{m,1}^*(\mathbf{k}, t), 0\big)^T$, and the third, trivial band with eigenvector $\big(0, 0, 1\big)^T$. Here, the upper and lower bands are written in terms of the original two-band micromotion operator's upper eigenvector $|z_m(\mathbf{k}, t)\rangle = \big(z_{m,1}(\mathbf{k}, t), z_{m,2}(\mathbf{k}, t)\big)^T$, and are labelled in Fig. A.1 by the original Bloch spin vectors $+\hat{\mathbf{n}}(\mathbf{k}, t), -\hat{\mathbf{n}}(\mathbf{k}, t)$. The contraction consists of three steps. First, the quasienergy of the trivial band is shifted to be degenerate with the lower $SU(2)$ band. At this point the unitary takes the form

$$\tilde{U}(\mathbf{k}, t) = e^{i\phi_m(\mathbf{k},t)} |\psi(\mathbf{k}, t)\rangle\langle\psi(\mathbf{k}, t)| + e^{-i\phi_m(\mathbf{k},t)}(1 - |\psi(\mathbf{k}, t)\rangle\langle\psi(\mathbf{k}, t)|), \tag{A.9}$$
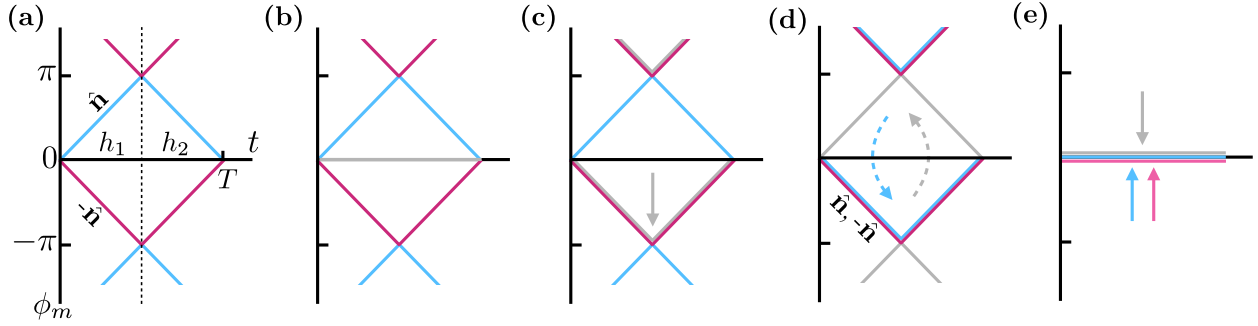
Figure A.1: Depiction of the embedding of the micromotion operator in $U(3)$, and its contraction to the identity operator. (**a**) The original, two-band micromotion operator. Its upper (blue) and lower band (magenta) are labelled by their Bloch spins $\hat{\mathbf{n}}, -\hat{\mathbf{n}}$, and possess some instantaneous Hopf invariant. Here we depict a flat band micromotion operator, where the invariant changes from $h_1$ to $h_2$ across a $\pi$-defect at $t = T/2$. If $h_1 \neq h_2$ mod 2, the $\pi$-defect is topologically protected and the unitary cannot be contracted to the identity. (**b**) After embedding, the unitary gains an additional, trivial band (gray) at zero quasienergy. (**c**) The first step of the contraction changes the quasienergy of the trivial band to be degenerate with the lower band (solid arrow). (**d**) Second, the eigenvectors of the upper band and the trivial are smoothly exchanged via Eq. (A.10), without changing either band's quasienergy (dashed arrows). After this exchange, the original upper and lower band are degenerate, and do not possess any nontrivial topology. (**e**) This allows the quasienergy of all bands to be deformed continuously to zero, completing the contraction.

where $|\psi(\mathbf{k}, t)\rangle = \big(z_1(\mathbf{k}, t), z_2(\mathbf{k}, t), 0\big)^T$ is the upper band's three-component eigenvector. In the second step, we use the triviality of the Hopf insulator in three-band systems to deform the upper band to a trivial band (no $\mathbf{k}$-dependence), while leaving the quasienergies $\phi_m(\mathbf{k}, t)$ unchanged. Specifically, the unitary maintains the above form, but with upper band eigenvector

$$|\psi(\mathbf{k}, t, \rho)\rangle = \begin{pmatrix} \cos(\rho)\, z_{m,1}(\mathbf{k}, t) \\ \cos(\rho)\, z_{m,2}(\mathbf{k}, t) \\ \sin(\rho) \end{pmatrix}, \tag{A.10}$$

for $\rho \in [0, \pi/2]$. As the lower bands' eigenvectors must remain orthogonal to the upper band, this step effectively swaps the eigenvectors of the upper band and the trivial lower band [Fig. A.1(d)]. After this step, the two $\mathbf{k}$-dependent bands are degenerate (see Fig. A.1), and are therefore equivalently labelled by $\mathbf{k}$-*independent* eigenvectors $\big(1, 0, 0\big)$, $\big(0, 1, 0\big)$. This relabelling makes it clear that the 'defect' at $\pi$-quasienergy is no longer topologically protected. In the third step, we complete the contraction by continuously take the quasienergies of all three bands to 0 for all $\mathbf{k}, t$.

Contraction in hand, we turn to computing the Floquet Hopf invariant via Eq. (A.3).

The first and final steps of the contraction involve changing the quasienergies of bands with no $\mathbf{k}$-dependence, and as such can be shown to not contribute to Eq. (A.3). We thus only need to consider the second step, defined via Eqs. (A.9) and (A.10). Using these, we compute

$$\tilde{U}^{-1}\partial_i\tilde{U} = (i\partial_i\Delta\phi_m)\,|\psi\rangle\langle\psi| + |\alpha|^2\,\langle\psi|\partial_i\psi\rangle\,|\psi\rangle\langle\psi| - \alpha^*\,|\psi\rangle\langle\partial_i\psi| + \alpha\,|\partial_i\psi\rangle\langle\psi| - \frac{i}{2}\partial_i\Delta\phi_m\mathbb{1},$$

(A.11)

where we define the quasienergy difference $\Delta\phi_m \equiv 2\phi_m$, and $\alpha(\mathbf{k},t) \equiv e^{i\Delta\phi_m(\mathbf{k},t)} - 1$ for convenience. From this, we see that the integrand of Eq. (A.3) consists of a sum of terms, each of which are products of expressions of the form $i\partial_i\Delta\phi_m$, $\langle\psi|\partial_i\psi\rangle$, $\langle\partial_i\psi|\psi\rangle$, or $\langle\partial_i\psi|\partial_j\psi\rangle$, with each $\partial_i$ occurring exactly once in each term. We note that the final term $\sim \mathbb{1}$ in this expression can be eliminated by adding a global phase to $\tilde{U}(\mathbf{k},t,\rho)$, and thus does not contribute the topological invariant.

We now argue that only terms of the form $\partial_i\phi\,\langle\partial_j\psi|\partial_k\psi\rangle\,\langle\partial_l\psi|\partial_n\psi\rangle$ are nonzero after antisymmetrization by $\epsilon^{ijklm}$. It is clear that any term with greater than one derivative acting on $\Delta\phi_m$ antisymmetrizes to zero. Similarly, any term that contains greater than one expression $\langle\psi|\partial_i\psi\rangle$ also antisymmetrizes to zero. A term containing $\langle\psi|\partial_i\psi\rangle\,\langle\partial_j\psi|\psi\rangle$ also gives zero, as $\langle\psi|\partial_i\psi\rangle$ and $\langle\psi|\partial_j\psi\rangle$ are both pure imaginary (because $|\psi\rangle$ is normalized), and so the index-flipped $\langle\psi|\partial_j\psi\rangle\,\langle\partial_i\psi|\psi\rangle = \langle\psi|\partial_i\psi\rangle^*\,\langle\partial_j\psi|\psi\rangle^*$ is equal to the original expression and they cancel after antisymmetrization. The only terms left given these restrictions are of the form $\partial_i\phi\,\langle\partial_j\psi|\partial_k\psi\rangle\,\langle\partial_l\psi|\partial_n\psi\rangle$ and $\langle\psi|\partial_i\psi\rangle\,\langle\partial_j\psi|\partial_k\psi\rangle\,\langle\partial_l\psi|\partial_n\psi\rangle$. The latter is ruled out by our choice of deformation: recalling our expression for $|\psi\rangle$ in terms of $|z\rangle$ in Eq. (A.10), we compute the inner products

$$\begin{aligned}
\langle\partial_\rho\psi|\psi\rangle &= 0 \\
\langle\partial_\rho\psi|\partial_j\psi\rangle &= -\cos(\rho)\sin(\rho)\,\langle z(\mathbf{k},t)|\partial_j z(\mathbf{k},t)\rangle \\
\langle\partial_l\psi|\partial_n\psi\rangle &= \cos^2(\rho)\,\langle\partial_l z(\mathbf{k},t)|\partial_n z(\mathbf{k},t)\rangle,
\end{aligned}$$

(A.12)

for $j,l,m \in \{k_x,k_y,k_z,t\}$. As $\partial_\rho\phi$ and $\langle\psi|\partial_\rho\psi\rangle$ are both zero, the $\rho$-derivative must appear in a term of the form $\langle\partial_\rho\psi|\partial_j\psi\rangle \sim \langle z(\mathbf{k},t)|\partial_j z(\mathbf{k},t)\rangle$. However, the first expression in the term $\langle\psi|\partial_i\psi\rangle\,\langle\partial_j\psi|\partial_k\psi\rangle\,\langle\partial_l\psi|\partial_n\psi\rangle$ is similarly proportional to some $\langle z(\mathbf{k},t)|\partial_i z(\mathbf{k},t)\rangle$, and so the term will again vanish upon antisymmetrization.

Keeping only the aforementioned nonzero terms, we have

$$\begin{aligned}
h_F &= \frac{1}{48\pi^3}\int_0^{\pi/2} d\rho \int dt\, d^3\mathbf{k}\,|\alpha|^4 \sum_{\substack{ijkln\in\\k_x,k_y,k_z,t,\rho}} \epsilon^{ijkln}\partial_i\Delta\phi_m\,\langle\partial_j\psi|\partial_k\psi\rangle\,\langle\partial_l\psi|\partial_n\psi\rangle \\
&= \frac{1}{24\pi^3}\int d\rho\, dt\, d^3\mathbf{k}\,|\alpha|^4 \sum_{\substack{ijkl\in\\k_x,k_y,k_z,t}} \epsilon^{ijkl}\partial_i\Delta\phi_m\,\langle\partial_j\psi|\partial_k\psi\rangle\,\big(\,\langle\partial_l\psi|\partial_\rho\psi\rangle - \langle\partial_\rho\psi|\partial_l\psi\rangle\,\big).
\end{aligned}$$

(A.13)

A factor of 5 was gained from the cyclicity of the trace. We proceed further by using Eqs. (A.12) and performing the $\rho$ integral. We find

$$
h_F = -\frac{1}{24\pi^3} \left( \int d\rho \cos^3(\rho) \sin(\rho) \right) \int dt \, d^3\mathbf{k} \, |\alpha|^4 \times
$$

$$
\sum_{\substack{ijkl \in \\ k_x, k_y, k_z, t}} \epsilon^{ijkl} \partial_i \Delta\phi_m \, \langle \partial_j z(\mathbf{k},t) | \partial_k z(\mathbf{k},t) \rangle \left( \langle z(\mathbf{k},t) | \partial_l z(\mathbf{k},t) \rangle - \langle \partial_l z(\mathbf{k},t) | z(\mathbf{k},t) \rangle \right)
$$

$$
= -\frac{1}{96\pi^3} \int dt \, d^3\mathbf{k} \, |\alpha|^4 \times
$$

$$
\sum_{\substack{ijln \in \\ k_x, k_y, k_z, t}} \epsilon^{ijln} \partial_i \Delta\phi_m \, \langle \partial_j z(\mathbf{k},t) | \partial_k z(\mathbf{k},t) \rangle \left( \langle z(\mathbf{k},t) | \partial_l z(\mathbf{k},t) \rangle - \langle \partial_l z(\mathbf{k},t) | z(\mathbf{k},t) \rangle \right),
$$

$$
= \frac{1}{24\pi} \int dt \, d^3\mathbf{k} \left[ 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m) \right] \sum_{\substack{ijkl \in \\ k_x, k_y, k_z, t}} \epsilon^{ijkl} \partial_i \Delta\phi_m \, \mathcal{F}_{jk} \, \mathcal{A}_l,
$$

$$(A.14)$$

defining $\mathcal{A}_l = \frac{-i}{4\pi}(\langle z | \partial_l z \rangle - \langle \partial_l z | z \rangle)$ and $\mathcal{F}_{jk} = \frac{-i}{2\pi}(\langle \partial_j z(\mathbf{k},t) | \partial_k z(\mathbf{k},t) \rangle - \langle \partial_k z(\mathbf{k},t) | \partial_j z(\mathbf{k},t) \rangle)$, and recalling $|\alpha|^4 = 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m)$. We will soon show that only the first, non-oscillating term of $|\alpha|^4$ is nonzero after integration, in which case the above expression replicates the Floquet Hopf invariant formula presented in Ch. 5. We have also confirmed that these terms are zero numerically.

We are now ready to confirm the relation between the Floquet invariant and the defect charges. We begin with the special case of flat band micromotion operators, to build intuition. Consider the micromotion operator depicted in Fig. A.1(a) (and written in Eq. (A.1), replacing $h_S - h_F \to h_1, h_S \to h_2$). Here $\Delta\phi_m(\mathbf{k},t) = \Delta\phi_m(t)$ and so we have

$$
h_F = \frac{1}{24\pi} \left( \int_0^T dt \, \partial_t \Delta\phi_m \left[ 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m) \right] \right) \left( \int d^3\mathbf{k} \sum_{\substack{jkl \in \\ k_x, k_y, k_z}} \epsilon^{jkl} \mathcal{A}_j \, \mathcal{F}_{kl} \right)
$$

$$
= \frac{1}{12\pi} \int_0^T dt \, (\partial_t \Delta\phi_m) \left[ 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m) \right] h(t)
$$

$$
= \frac{1}{12\pi} \int_0^{2\pi} d(\Delta\phi_m) \left[ 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m) \right] h_1 +
$$

$$
\frac{1}{12\pi} \int_{2\pi}^0 d(\Delta\phi_m) \left[ 6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m) \right] h_2 \Big)
$$

$$
= \frac{6 \times 2\pi}{12\pi} \left( h_1 - h_2 \right)
$$

$$
= h_\pi \bmod 2.
$$

$$(A.15)$$

In the first step, we use the definition of the Hopf invariant. In the second to last step, we separately compute the integral for the two regions in which $h(t)$ is constant: $t \in [0, T/2]$, where $\Delta\phi_m$ increases from 0 to $2\pi$, and $t \in [T/2, T]$, where $\Delta\phi_m$ decreases from $2\pi$ to 0. The integral receives a positive contribution from the first region, and a negative contribution from the second; this subtraction is precisely the change in Hopf invariant between the two regions, at $\Delta\phi_m = 2\pi$ ($\phi_m = \pi$). It is straightforward to generalize this to the case of multiple defects, and one replicates the above formula for the total $\pi$-defect charge $h_\pi$.

Inspired by the flat band case, we can relate the invariant to the defect charges for general micromotion operators with the help of a coordinate change. We transform from the space-time coordinates $(k_x, k_y, k_z, t)$ to coordinates $(x_1, x_2, x_3, \Delta\phi_m)$, where $\Delta\phi_m$ is again the micromotion's relative phase and the $x_i$ are some coordinates that parameterize the three-dimensional submanifolds of constant $\Delta\phi_m$. In general this coordinate transformation is not possible (not 1-to-1) on the entire space, and we will have to divide the space into a union of patches $P$, each with their own coordinates $(x_1^P, x_2^P, x_3^P, \Delta\phi_m^P)$. This is similar to the flat band case, where $\Delta\phi_m \sim t$ is increasing on the first patch ($t < T/2$) and $\Delta\phi_m \sim -(t - T/2)$ is decreasing on the second ($t > T/2$). The Floquet Hopf invariant then takes the form

$$h_F = \frac{1}{24\pi} \sum_P \int d(\Delta\phi_m^P) \, |\alpha(\Delta\phi_m^P)|^4 \int d^3\mathbf{x}^P \sum_{jkl \in \{x_1^P, x_2^P, x_3^P\}} \epsilon^{jkl} \, \mathcal{A}_j(\Delta\phi_m^P, \mathbf{x}^P) \, \mathcal{F}_{kl}(\Delta\phi_m^P, \mathbf{x}^P),$$
(A.16)

where the change of coordinates has little effect because the integral is a topological invariant. The $\mathbf{x}^P$-integrals are performed over three-dimensional manifolds, defined by the surface of constant $\Delta\phi_m^P$. Contributions from different patches of the same constant $\Delta\phi_m$ are added together, each with the same prefactor $|\alpha(\Delta\phi_m)|^4$. The sum over $P$ thus amounts to changing the domain of the $\mathbf{x}$-integral to be over the *entire* submanifold of constant $\Delta\phi_m$ (i.e. over all patches), which we will call $\mathcal{M}_{\Delta\phi_m}$. We have,

$$h_F = \frac{1}{12\pi} \int d(\Delta\phi_m) \, |\alpha(\Delta\phi_m)|^4 \left[ \frac{1}{2} \int_{\mathcal{M}_{\Delta\phi_m}} d^3\mathbf{x} \sum_{jkl \in \{x_1, x_2, x_3\}} \epsilon^{jkl} \, \mathcal{A}_j(\Delta\phi_m, \mathbf{x}) \, \mathcal{F}_{kl}(\Delta\phi_m, \mathbf{x}) \right].$$
(A.17)

The submanifold $\mathcal{M}_{\Delta\phi_m}$ is closed, and may consist of multiple disjoint regions. The sign of the contribution from each region is equivalent to the orientation of that region, and is fixed by sign of the original derivative $(\partial_i \Delta\phi_m)$ to point in the direction of increasing $\Delta\phi_m$. (For instance, in the flat band example there are two regions for each $\Delta\phi_m$, the increasing region $[t < T/2]$ and the decreasing region $[t > T/2]$, and the contribution from the decreasing region is subtracted from that of the increasing region.) Looking at Eq. (A.18), the integral over $\mathcal{M}_{\Delta\phi_m}$ is nothing more than the (signed) sum of the static Hopf invariants of each closed region of the submanifold, $h(\mathcal{M}_{\Delta\phi_m})$. As a topological invariant, $h(\mathcal{M}_{\Delta\phi_m})$ cannot change with the relative phase $\Delta\phi_m$, and is therefore fixed to be a constant for all $\mathcal{M}_{\Delta\phi_m}$. As one might guess from the flat band example, this constant is precisely the Floquet Hopf

invariant,

$$
\begin{aligned}
h_F &= \frac{1}{12\pi} \int_0^{2\pi} d(\Delta\phi_m) \, |\alpha(\Delta\phi_m)|^4 h(\mathcal{M}_{\Delta\phi_m}) \\
&= \frac{1}{12\pi} \int_0^{2\pi} d(\Delta\phi_m) \left[6 - 8\cos(\Delta\phi_m) - 2\cos(2\Delta\phi_m)\right] h(\mathcal{M}_{\Delta\phi_m}) \\
&= h(\mathcal{M}_{\Delta\phi_m}).
\end{aligned}
\tag{A.18}
$$

where the constancy of $h(\mathcal{M}_{\Delta\phi_m})$ causes all oscillatory terms $\sim e^{in\Delta\phi_m(\mathbf{k},t)}$ (for nonzero $n$) to integrate to zero.

It remains only to identify the constant $h(\mathcal{M}_{\Delta\phi_m}) = h_F$ with the defect charges. Consider the submanifold $\mathcal{M}_{\Delta\phi_m}$ for $\Delta\phi_m$ infinitesimally close to $2\pi$. For a micromotion operator with $N$ point Hopf defects at $\pi$-quasienergy, of charges $\Delta h_\pi^i, i = 1, \dots, N$, this manifold consists of $N$ 3-spheres, each enclosing one of the point Hopf defects. The Hopf invariant of each 3-sphere is exactly the charge of its enclosed defect, and each 3-sphere contributes with the same sign. The total Hopf invariant of the submanifold then gives

$$
h_F = h(\mathcal{M}_{2\pi-\epsilon}) = \sum_i \Delta h_\pi^i = h_\pi.
\tag{A.19}
$$

This completes our identification of the parity of $h_\pi$ with the Floquet invariant.

## A.3  Examples of point and loop Hopf defects

A simple example of a point Hopf defect, of charge $\Delta h = 1$, is introduced in Ch. 5, taking $m = 1$, $\mathbf{k} = 0$. Replacing $m$ with $t$, the eigenvector near the defect varies as

$$
|z(\mathbf{k}, t)\rangle = \begin{pmatrix} \delta k_x + i\,\delta k_y \\ \delta k_z + i\,\delta t \end{pmatrix}.
\tag{A.20}
$$

To derive the quadratic energy degeneracy of this defect, we consider the instantaneous Hamiltonian giving rise to the defect, $H_{\text{inst}}(\mathbf{k}, t) = \mathbf{n}(\mathbf{k}, t)\cdot\boldsymbol{\sigma}$, with $\mathbf{n}(\mathbf{k}, t) = \langle z(\mathbf{k}, t)|\,\boldsymbol{\sigma}\,|z(\mathbf{k}, t)\rangle$. Note that $|z(\mathbf{k}, t)\rangle$, and therefore $\mathbf{n}(\mathbf{k}, t)$, are not normalized here, in contrast to Ch. 5. [There, we considered a gapped Hamiltonian, which can always be deformed to a flat band Hamiltonian with normalized $\mathbf{n}(\mathbf{k}, t)$.] Their normalization gives us the quasienergy in the vicinity of the defect, $\Delta\phi(\mathbf{k}, t) = |H_{\text{inst}}(\mathbf{k}, t)| = |\langle z(\mathbf{k}, t)|\,\boldsymbol{\sigma}\,|z(\mathbf{k}, t)\rangle| = \delta\mathbf{k}^2 + \delta t^2$.

We now turn to loop Hopf defects. After demonstrating an example of such a defect, we will use the example to show: first, that the *point* Hopf defect can be viewed as a contraction of the loop defect; and second, that each point along the loop defect appears as a Weyl cone, with the frame of the Weyl cone twisting by $2\pi$ as one circles the defect. Our example loop defect is also found in Ch. 5, this time by adding an on-site splitting $\mu\,\sigma_z$ to the associated instantaneous Hamiltonian. This splitting changes the Bloch sphere

vector $\mathbf{n}(\mathbf{k}) \to \mathbf{n}(\mathbf{k}) + \mu \hat{\mathbf{z}}$. In the Hopf insulator phase, the pre-image of $\hat{\mathbf{n}}(\mathbf{k}) = -\hat{\mathbf{z}}$ is a loop in the Brillouin zone defined by $k_z = 0$, $\cos(k_x) + \cos(k_y) = m - 1$, with energy $E(\mathbf{k}) = |\mathbf{n}(\mathbf{k})| = \sqrt{||z_1(\mathbf{k})|^2 - \mu|} = \sqrt{|\sin^2(k_x) + \sin^2(k_y) - \mu|}$ along the loop. For each point along the pre-image, there is a value of $\mu$ at which the energy equals zero, and above which $n_z$ is positive. As $\mu$ is increased this change propagates along the pre-image, until all points have transitioned to being pre-images of $\hat{\mathbf{n}}(\mathbf{k}) = +\hat{\mathbf{z}}$. Envisioning $\mu$ as varying in time, this realizes a loop defect with some finite extent in the time-direction.

It is illuminating and convenient to analyze the loop defect in the limit where it occurs at small $k_x, k_y$. The smallness of the pre-image is controlled by the parameter $\epsilon = 3 - m$, where for small $\epsilon$ the pre-image occurs along the circle $k_z = 0$, $k_x^2 + k_y^2 = 2\epsilon$, with energy $\sqrt{|2\epsilon - \mu|}$ [neglecting corrections $O(k^3)$]. Since the energy is now constant along the pre-image, the defect occurs at a single value of the splitting (i.e. a single instant in time), $\mu = 2\epsilon$. Near the defect, the Bloch sphere vector varies as

$$
\begin{aligned}
n_x &= k_x k_z + k_y \delta \\
n_y &= -k_x \delta - k_y k_z \\
n_z &= k_x^2 + k_y^2 - k_z^2 - \delta^2 - \mu
\end{aligned}
\qquad , \qquad (\text{A.21})
$$

defining $k_r^2 \equiv k_x^2 + k_y^2$, $\delta \equiv \epsilon - \frac{1}{2}k_r^2$, and $\mu = 2\epsilon + \delta\mu$. Keeping only first order terms, and defining $\delta k_r \equiv -\delta/\sqrt{2\epsilon}$ and $k_x = k_r \cos(\theta)$, $k_y = k_r \sin(\theta)$ we have

$$
\begin{aligned}
n_x &= \sqrt{2\epsilon}\, k_z \cos(\theta) - 2\epsilon\, \delta k_r \sin(\theta) \\
n_y &= -2\epsilon\, \delta k_r \cos(\theta) + \sqrt{2\epsilon}\, k_z \sin(\theta). \\
n_z &= -\delta\mu
\end{aligned}
\qquad (\text{A.22})
$$

From this we see that the energy is indeed linearly dispersing about the loop defect. We note that in the limit $\epsilon \to 0$ the linearly-dispersing loop defect contracts to a single quadratically-dispersing point at $\mathbf{k} = 0$, showing that the point Hopf defect can be understood as a contraction of the loop defect.

At each point along the loop, the three vectors $\partial_{k_z}\mathbf{n}$, $\partial_\mu\mathbf{n}$, and $\partial_{k_r}\mathbf{n}$ define the frame of a Weyl cone extending in the directions perpendicular to the loop. To analyze the twisting of the Weyl cone about the loop, we compute these vectors

$$
\begin{aligned}
\partial_{k_z}\mathbf{n} &= \big(\sqrt{2\epsilon}\cos(\theta),\, -\sqrt{2\epsilon}\sin(\theta),\, 0\big) \\
\partial_{k_r}\mathbf{n} &= \big(-2\epsilon\sin(\theta),\, -2\epsilon\cos(\theta),\, 0\big) \\
\partial_\mu\mathbf{n} &= \big(0,\, 0,\, -1\big).
\end{aligned}
\qquad (\text{A.23})
$$

These vectors are linearly independent and twist by $2\pi$ about the $\mu$-axis along the loop parameterized by $\theta$, verifying the Weyl cone interpretation.

## A.4  Defects during the deformation

Here we describe the topological defects' trajectories throughout the smooth deformation in more detail. We denote the initial ($\lambda = 0$) times of the 0- and $\pi$-defects as $t_0$ and $t_\pi$ respectively. Expressing $U_\lambda$ in the 3-sphere coordinates $(\xi, \mathbf{n})$, the gapless points at a given $\lambda$ are by definition the pre-images of $U_\lambda = (\pm 1, 0, 0, 0)$. From Ch. 5, these are identical to the pre-images of the points $R^{-1}_{\xi n_z}(\lambda)(\pm 1, 0, 0, 0) = (\pm \cos(\pi\lambda), 0, 0, \pm \sin(\lambda))$ in the *unrotated* evolution $U_{\lambda=0}$. For simplicity, we assume that $U_0$ has flat bands for some range in between $t_0$ and $t_\pi$, such that in this range the quasienergies linearly interpolate between their defect values of 0 and $\pi$, $\xi(\mathbf{k}, t) = \xi(t) = \cos\left(\frac{t-t_0}{t_\pi - t_0}\pi\right)$. With this assumption, the defects occur at fixed time-slices $t = \xi^{-1}\left(\pm\cos(\pi\lambda)\right) = \pm(\lambda - \frac{1}{2} \pm \frac{1}{2})(t_\pi - t_0) + t_0$. (If we relax the flat band assumption the defects will have some extent in the $t$-direction.) Further, within these time-slices, the pre-images in the Brillouin zone are precisely the pre-images of $\hat{\mathbf{n}} = \pm\hat{\mathbf{z}}$ in the instantaneous *static* Hamiltonian defined by the evolutions' eigenvectors at this time. Thus the 0-defect is a loop of gapless points moving forward in time (as $\lambda$ increases) and the $\pi$-defect is a different loop moving backwards. At the intermediate value $\lambda = 1/2$, the two defects exist in the same time-slice $t = (t_0 + t_\pi)/2$ and are *linked* in the Brillouin zone, by virtue of the linking of the $\hat{\mathbf{n}} = \pm\hat{\mathbf{z}}$ pre-images in the instantaneous Hamiltonian, which has Hopf invariant 1. At this $\lambda$, only the total charge $h_0 + h_\pi$ is well-defined. The analysis of Ch. 5 shows that as the pre-images pass each other the individual charges $h_0$ and $h_\pi$ are indeed changed from their initial values, and only the total charge and the Floquet invariant $h_\pi$ mod 2 are preserved.

## A.5  Edge modes

In this section, we first review the theoretical arguments made in Ch. 5 on the bulk-boundary correspondence in more detail, and second, numerically demonstrate that edge modes at certain sharp terminations of the lattice are not protected against generic perturbations. Details on the numerical simulations of edge modes are found in the following section.

### Expanded theoretical arguments for edge modes

We briefly review the properties of edge modes in the static Hopf insulator, and use this to establish the edge modes of the Floquet Hopf insulator in more detail. Unlike topological insulators classified by K-theory, the edge modes of the Hopf insulator are protected only for sufficiently smooth edges [335, 430]. This is a consequence of the two-band requirement: the notion of bands relies on translational invariance, which is broken at a sharply terminated edge. At edges that are smooth relative to the lattice length, we may define the adiabatic Hamiltonian $H(\mathbf{k}; \mathbf{r})$. At a boundary where this Hamiltonian's topological invariant changes,

the Hopf insulator's nontrivial homotopy classification requires the Hamiltonian's gap to close, which we identify as the gapless edge mode.

A similar argument holds for the Floquet Hopf insulator, if we replace the adiabatic Hamiltonian with an adiabatic unitary evolution $U(\mathbf{k}, t; \mathbf{r})$. The defect charges $h_0/h_\pi$ of the adiabatic evolution can change by two mechanisms: a $0/\pi$-gap closing, respectively; or the smooth deformation that takes $(h_0, h_\pi) \to (h_0 + 2n, h_\pi - 2n)$. The latter cannot change the charges' parities, which leads to a simple rule: if the parity of $h_0/h_\pi$ changes across a smooth boundary, there will be a gapless edge mode at $0/\pi$-quasienergy. For instance, in Fig. 5.4(a), the parity of both $h_0$ and $h_\pi$ change across the edge region, leading to gapless edge modes in each edge. In contrast, in Fig. 5.4(b) neither defect charge parity changes, nor does the static invariant, and so the two systems are in fact the same phase and no edge modes are observed. In both of these examples, the behavior is identical to that of a stable $\mathbb{Z}_2 \times \mathbb{Z}_2$ Floquet topological insulator.

The more unique case is that described in Ch. 5 and displayed in Fig. 5.4(c-d), at a boundary where the static invariant changes, but the defect charges' parities do not (i.e. the Floquet invariant does not change, but the static invariant changes by an even amount). While the change in static invariant necessitates a gap closing, neither gap individually requires one, and so the quasienergy of the gapless edge mode will be determined by details of the edge region. This is illuminated by a specific example: consider a boundary between the trivial phase, $(h_S, h_F) = (0, 0)$ [e.g. $(h_0, h_\pi) = (0, 0)$], and the phase $(h_S, h_F) = (2, 0)$ [e.g. $(h_0, h_\pi) = (0, 2)$]. A gap closing at $\pi$-quasienergy suffices to change $h_\pi$ by $+2$ and connect the two phases. However, the latter phase is equivalently described by defect charges $(h_0, h_\pi) = (-2, 0)$, which is connected to the $(0, 0)$ phase by a closing at $0$-quasienergy. The 'natural' quasienergy of the gapless edge modes depends on the specific defect charges of the two phases, which is a non-universal property of the adiabatic edge region. This behavior arises precisely because of the different classifications of the static and Floquet invariants, which prevents the integer static invariant $h_0 + h_\pi$ from being decomposed into a sum of two independently conserved invariants $h_0$, $h_\pi$, since $h_\pi$ is only conserved modulo 2.

## Edge modes at sharp terminations

In this section we briefly investigate edge modes at sharp edges of the Floquet Hopf insulator, in contrast to the smooth edges used for the numerics of Ch. 5. As previously mentioned, owing to the triviality of the Hopf insulator in systems with band number greater than 2, the Hopf insulator only protects gapless edge modes at smooth edges. This was first pointed out by Ref. [335], and observed numerically in Ref. [430]. Past works have nevertheless observed gapless edge modes even at 'sharp' edges – specifically, for open boundary conditions, which can be viewed as a sharp boundary between the system and the trivial phase [131, 335]. In Ref. [430], it was shown that such sharp edge modes arise as a result of unintentional crystalline symmetries of the Hamiltonians considered, and are not protected against perturbations that break these symmetries. For instance, one such crystalline symmetry, pointed out in Ref. [299], stabilizes the Hopf invariant to a $\mathbb{Z}_2$ invariant in higher band systems. In
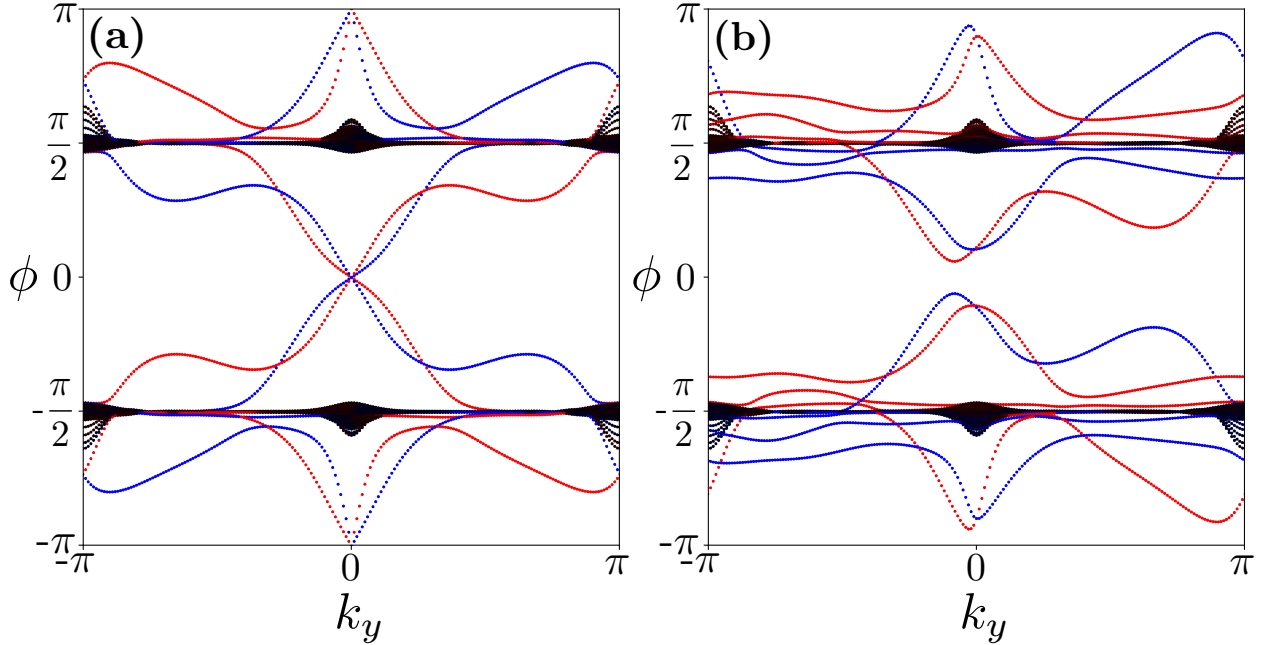
Figure A.2: Quasienergy spectra of the Floquet unitary for the $(h_S, h_F) = (0, 1)$ phase with open boundary conditions, solved via exact diagonalization. Quasienergies are colored according to their eigenstates' average position, from left edge (blue) to bulk (black) to right edge (red). **(a)** Without any additional perturbation to the edge, gapless edge modes are observed across both band gaps at both the left and right edges. **(b)** Adding perturbations localized at each edge leads to gaps in both edges' spectra, suggesting that gapless sharp edge modes are protected by unintentional crystalline symmetries of the model.

numerics, one must therefore be careful to work with only smooth edges, if one is interested in pure Hopf insulator edge modes and not those of related topological crystalline insulators.

Like the static Hopf invariant, the Floquet Hopf invariant is only non-trivial in two-band systems. One therefore expects gapless edge modes at sharp terminations of the Floquet Hopf insulator, as seen in Ref. [207], to only exist via the same 'unintentional' crystalline symmetries that stabilize the Hopf insulator edge modes in non-driven systems. We do not aim to provide an exhaustive classification of such symmetries, but instead only to demonstrate one particular case of these edge modes' instability to generic perturbations, at the (100) edge of our model. In Fig. A.2(a), we compute the spectrum of the Floquet Hopf insulator $(h_S, h_F) = (0, 1)$ with naive open boundary conditions (i.e. hoppings from sites within the lattice to sites terminated from it are set to zero, but hoppings remaining within the lattice are unchanged). We indeed observe gapless edge modes, with linear dispersion, across both band gaps, and on both the left and right edges of the system. However, as expected, these edge modes are gapped by adding a perturbation localized to each lattice

edge, as shown in Fig. A.2(b). This indicates that they are *not* in correspondence with the bulk Floquet Hopf insulator phase, and instead must be protected some unintentional symmetry of the model.

## A.6 Details on numerics

### Fig. 5.2 and Fig. 5.3

The evolution depicted in Fig. 5.2, and those used the numerics of Fig. 5.3, are calculated similarly. The evolution consists of five components:

1. For times $0 \leq t < t_1$ the quasienergy linearly increases, $\phi = (t/t_1)\, \phi_m$. For the first half of this $0 \leq t < t_1/2$ the eigenvector are in the product state $\hat{\mathbf{n}}(\mathbf{k}, t) = \hat{\mathbf{z}}$. For the second half $t_1/2 \leq t < t_1$ the eigenvectors are smoothly deformed to those of the trivial insulator described in Ch. 5 at $m = 4$. This done by multiplying the first component $z_1(\mathbf{k}, t)$ of the eigenvector by constant that increases from 0 to 1 from $t_1/2$ to $t_1$, and is possible while maintaining the eigenvectors normalization because $z_2(\mathbf{k}, t) \neq 0$ for all $\mathbf{k}$ in the trivial phase.

2. For times $t_1 \leq t < t_2$, the eigenvectors continue to be described by those in Ch. 5, but with $m$ varying with time as $m(t) = 4 - 2(t - t_1)/(t_2 - t_1)$. During this the instantaneous Hopf invariant changes from 0 at $m = 4$ to 1 at $m = 2$ through a gap closing at quasienergy 0, corresponding to a point 0-defect with charge $+1$. We take the quasienergy to be

$$\phi = \begin{cases} (1 - d) + d|z(\mathbf{k}, t)|^2 & (1 - d) + d|z(\mathbf{k}, t)|^2 < \phi_m \\ \phi_m & \text{else} \end{cases} \tag{A.24}$$

where $d \equiv e^{-(k_x^2 + k_x^2 + k_x^2 + (m-3)^2)/w^2}$ is 1 at the gapless point and asymptotically 0. This describes a gapless region of approximately quadratic dispersion and finite-extent, surrounded by an otherwise flat band with quasienergy $\phi_m$. We choose width $w = 1$, small enough such that the quasienergy is constant with $\mathbf{k}$ at the beginning and end times $t_1$, $t_2$, so as to match the preceding and proceeding evolution.

3. For times $t_1 + t_2 \leq t < T - t_1 + t_2$ the eigenvectors remain described by Ch. 5 at $m = 2$, and the quasienergy increases linearly from $\phi_m$ to $\pi - \phi_m$. This has instantaneous Hopf invariant 1 at all times.

4. For times $(T - t_2) \leq t < (T - t_1)$ the process of the second step is replicated, but with the reversed quasienergy $\phi \to \pi - \phi$ and time $t \to T - t$. This gives a gap closing at $\pi$ with charge $-1$.

5. For times $T - t_1 \leq t$ the evolution of Fig. 5.2 is constant. For the micromotion operator, we take the eigenvectors to remain constant (described by Ch. 5 at $m = 2$), while the quasienergy linearly decreases from $\pi - \phi_m$ to 0. This is homotopy equivalent to the usual micromotion operator $U_m(\mathbf{k}, t) \equiv U(\mathbf{k}, t) \left( U(\mathbf{k}, T) \right)^{-t/T}$ via the deformation $U_m(\mathbf{k}, t) \equiv U(\mathbf{k}, t) \left( U(\mathbf{k}, T) \right)^{-p(t;\tau)}$, where the power $p(t; \tau)$ varies as

$$p(t; \tau) = \begin{cases} 0 & t < \tau \\ (t - T + \tau)/\tau & t \geq \tau \end{cases}, \tag{A.25}$$

which interpolates between the usual micromotion operator at $\tau = 0$ and our micromotion operator at $\tau = t_1$.

Fig. 5.2 depicts an evolution with parameters $t_1 = .2T, t_2 = .4T, \phi_m = 3\pi/8$. The phase transition and the smooth deformation were implemented with $t_1 = .3T, t_2 = .35T, \phi_m = .025\pi$ and $t_1 = .3T, t_2 = .35T, \phi_m = .025\pi$, respectively. This evolution has a 0-defect with charge 1 at time $t_0 = (t_1 + t_2)/2$ and a $\pi$-defect with charge $-1$ at time $t_\pi = T - (t_1 + t_2)/2$. A phase transition is naturally found by truncating the evolution at times $T(\lambda) < T$. If the truncation time $T(\lambda)$ occurs after the time of the $\pi$-defect, the system is in the initial phase $(h_S, h_F) = (0, 1)$. However, if the truncation time is *before* the $\pi$-defect, the evolution consists only of the charge 1 0-defect, and thus is in the $(h_S, h_F) = (1, 0)$ phase. The phase transition between the two is at $T(\lambda) = t_\pi$. For Fig. 5.3, we take $T(\lambda) = T - 2\lambda(T - t_\pi)$, which has a phase transition at $\lambda = 1/2$. The smooth deformation is implemented as described in Ch. 5, with $\theta(t)$ chosen to be the piecewise linear ramp

$$\theta(t) = \begin{cases} (t/t_1)\,\pi & t < t_1 \\ \pi & t \geq t_1 \end{cases}. \tag{A.26}$$

The integral for the Floquet Hopf invariant is evaluated by summing the integrand on $N^4$ points in the Floquet Brillouin zone. The points are chosen by first forming an evenly-spaced $N^4$ grid in the Floquet Brillouin zone, and then randomizing the position of each point within the width$-1/N$ hypercube surrounding each grid point. The randomization improves convergence of the integral, and reproduces the non-randomized result in the limit of large $N$. The eigenvectors and the phases $\phi(\mathbf{k}, t)$ obey a number of reflection symmetries such that the Floquet invariant integrand is invariant under $k_i \rightarrow -k_i$ for any $k_i$. This is used to reduce the domain of the integral to the $0 \leq k_i < \pi$ octant of the Brillouin zone. The $N^4$ grid is shifted by half a discretization in each dimension to avoid over-counting the contribution of the integration domain's boundary. At $N = 90$, the numerical invariants are within 0.6% of their theoretical values on both sides of the phase transition, and within 1.5% of the theoretical value throughout the smooth deformation. As expected, the error is seen to decrease with increasing $N$.

The defect charges, and the static Hopf invariant, are calculated by evaluating the instantaneous Hopf invariants $h(T_{\text{mid}})$ and $h(T)$ at two time-slices of the evolution: an intermediate

time $T_{\mathrm{mid}}$ between the two gapless points, and the final time $T$. If the 0-defect occurs be-fore the $\pi$-defect, the defect charges follow as $h_0 = h(T_{\mathrm{mid}})$ and $h_\pi = h(T) - h(T_{\mathrm{mid}})$, and vice versa if the $\pi$-defect occurs before the 0-defect. The static invariant is by definition $h_S = h(T)$. The instantaneous Hopf invariants are numerically calculated similar to the Floquet Hopf invariant, now on a $N^3$ grid in the (non-Floquet) Brillouin zone.

## Fig. 5.4 and Fig. A.2

All spectra are calculated via exact diagonalization of the hybrid real/momentum-space Floquet unitary $U_{xx'}(k_y, k_z, T)$, which is obtained by time-evolving under the Hamiltonian $H_{x,x'}(k_y, k_z, t)$. All numerics are done for $N = 60$ lattice sites in the $x$-direction. Time-evolution is computed by repeatedly applying the matrix $\exp(-i\delta t H_{x,x'}(k_y, k_z, t))$, with $\delta t/T = 500$ for Fig. 5.4(b) and $\delta t/T = 50$ for all other figures. The evolutions for Fig. 5.4(a,c,d) and Fig. A.2 are generated from the stroboscopic flat-band Hamiltonian construction in Ch. 5. The flat-band Hamiltonians are obtained by normalizing the Hamiltonians of Ref. [132] with $p = q = 1$ for $h = 1$ and $p = 2, q = 1$ for $h = 2$. Hoppings are obtained by Fourier transforming in the $x$-direction and truncating to a distance of at most 5 lattice sites. To form the boundary between two such phases for Fig. 5.4(a,c-d), we linearly interpolate between the two hoppings over a distance of 20 lattice sites in the center of the lattice. For Fig. 5.4(b), we begin with the stroboscopic evolution for $(h_S, h_F) = (0, 1)$, and perform the deformation of Fig. 5.2 'in real space' over a distance of 20 lattice sites in the center of the lattice. That is, at each of the sites in the edge region, we use time-dependent hoppings derived from the time-evolution $U(\mathbf{k}, t; \lambda)$, where $\lambda$ interpolates from 0 to 1 across the edge region. In all panels Fig. 5.4, we only plot quasienergies from eigenstates that have an average $x$-position outside a region of length 6 bordering either edge. This eliminates the sharp edge modes studied in Fig. A.2 from the spectra. In Fig. A.2(b), the symmetry-breaking perturbation is chosen as an $x$-dependent chemical potential $\mu_x \mathbb{1}$ localized on the two layers nearest each edge.

# Appendix B

# Details on quantum information scrambling on a trapped ion quantum processor

In this Appendix we provide additional details on the experiments and numerics presented in Ch. 7.

## B.1   Experimental details

*Trapped ion qubits.*—We perform the experiment on a quantum computer consisting of a chain of nine $^{171}\text{Yb}^+$ ions confined in a Paul trap and laser cooled near the motional ground state. The hyperfine-split $^2S_{1/2}$ ground level with an energy difference of 12.642821 GHz provides a pair of qubit states, $|0\rangle = |0, 0\rangle$ and $|1\rangle = |1, 0\rangle$ with quantum numbers $|F, m_F\rangle$, that are magnetic field independent to first order. The 1/e-coherence time of this so-called "atomic clock" qubit is 1.5(5) s in our system, limited by magnetic field noise. Optical pumping is used to initialize the state of all ions, and the final states are measured collectively via state-dependent fluorescence detection [356]. Each ion is mapped to a distinct channel of a photomultiplier tube (PMT) array. The average state detection fidelity is 99.4(1)% for a single qubit, while a 7-qubit state is typically read out with 92(1)% average fidelity, limited by channel-to-channel crosstalk. These state detection and measurement (SPAM) errors are characterized in detail by measuring the state-to-state error matrix.

*Gate operations.*—Quantum operations are achieved by applying two Raman beams from a single 355 nm mode-locked laser, which form beat notes near the qubit frequency. The first Raman beam is a global beam applied to the entire chain, while the second is split into individual addressing beams to target each ion qubit [126], controlled by a set of Arbitrary Waveform Generators (AWGs). Single qubit gates are generated by driving resonant Rabi rotations (R-gates) of defined phase, amplitude, and duration. Single-qubit Z-rotations are applied efficiently as classical phase advances. Two-qubit gates (XX-gates) are realized by
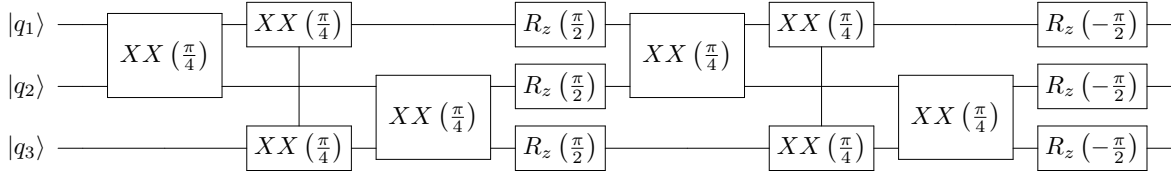
Figure B.1: Circuit representation of the scrambling unitary used for the probabilistic teleportation scheme [see Eq. (B.2)], consisting of six two-qubit entangling XX-gates and individual Z-rotations.

illuminating two ions with beat-note frequencies near the motional sidebands and creating an effective spin-spin Ising interaction via transient entanglement between the state of two ions and all modes of motion [330]. To ensure that the motion is left disentangled from the qubit states at the end of the interaction, we employ a pulse shaping scheme [101, 535]. We use only the middle seven ions in the chain as qubits, in order to ensure a higher uniformity in the ion spacing, matching the equally-spaced individual addressing beams. The two edge ion qubits are neither manipulated nor measured, but their contribution to the collective motion is included when creating the entangling operations.

*Bell state preparation and measurement.*—In Fig. 7.3(b), we depict Bell state pairs initially in the state $\frac{1}{\sqrt{2}} |00\rangle + |11\rangle$. This entangled state is created using the following circuit:

$$
\boxed{XX(\pm\tfrac{\pi}{4})} \; \boxed{R_z(\pm\tfrac{\pi}{2})} \tag{B.1}
$$

The sign of the XX gate depends on the pulse shape solution used for the particular gate, and is stored in a table in the control software, which then determines the appropriate Z rotation to create the Bell state. We use additional rotations to create the other Bell states from this circuit. We measure in the Bell basis by applying a simple CNOT gate followed by a Hadamard gate, and subsequent measurement of the two qubits.

## B.2 Implementing and optimizing scrambling operators

The scrambling unitary used for the probabilistic teleportation scheme can be represented in the computational basis by the following matrix:

$$\hat{U}_s = \frac{1}{2} \begin{pmatrix} -1 & 0 & 0 & -1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & -1 & -1 & 0 & -1 & 0 & 0 & -1 \end{pmatrix} \tag{B.2}$$

Since this unitary is real, $\hat{U}_s^* = \hat{U}_s$, simplifying the experimental sequence since $\hat{U}_d = \hat{U}_s^* = \hat{U}_s$. The unitary fulfills the following set of equations which verify its scrambling property by showing that it disperses all single-qubit operators into three-qubit operators.

$$\begin{aligned} U^\dagger (X \otimes I \otimes I) U &= X \otimes Z \otimes Z \\ U^\dagger (I \otimes X \otimes I) U &= Z \otimes X \otimes Z \\ U^\dagger (I \otimes I \otimes X) U &= Z \otimes Z \otimes X \\ U^\dagger (Y \otimes I \otimes I) U &= Y \otimes X \otimes X \\ U^\dagger (I \otimes Y \otimes I) U &= X \otimes Y \otimes X \\ U^\dagger (I \otimes I \otimes Y) U &= X \otimes X \otimes Y \\ U^\dagger (Z \otimes I \otimes I) U &= Z \otimes Y \otimes Y \\ U^\dagger (I \otimes Z \otimes I) U &= Y \otimes Z \otimes Y \\ U^\dagger (I \otimes I \otimes Z) U &= Y \otimes Y \otimes Z \end{aligned} \tag{B.3}$$

Here, X,Y,Z, and I are the Pauli operators, and the identity operator, respectively. $\hat{U}_s$ is implemented experimentally using six two-qubit entangling gates (see Fig. B.1).

In the experimental implementation, we take advantage of the following identity to reduce the number of two-qubit gates needed:



$$\tag{B.4}$$

where the vertical line indicates that the qubits are in an EPR state.

In this way, we can combine the first two XX gates on ions 2 and 3 in $\hat{U}_s$ (and similarly on ions 4 and 5 in $\hat{U}_d$) into two single-qubit X-rotations, as shown in Fig. B.2.

In Fig. 7.2 we vary the amount of scrambling in $U$ parametrized by $\alpha$. This is achieved by changing the angles of the Z-rotations depicted in Fig. B.4 according to $\theta = \pm \frac{\alpha \pi}{2}$. When
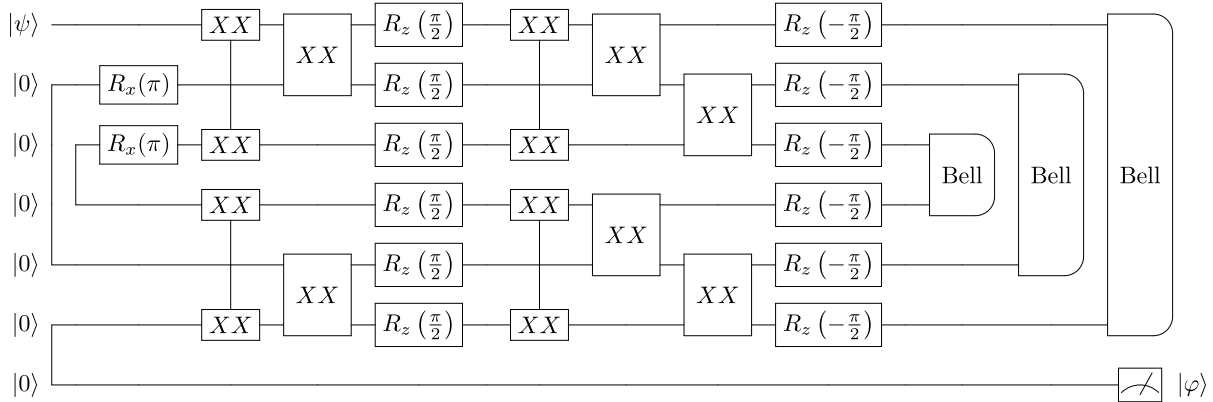
Figure B.2: The experimental sequence used for the probabilistic teleportation scheme. Any one of the 3 Bell measurements can be used. The scrambling unitary has been simplified using the identity given in Eq. (B.4).
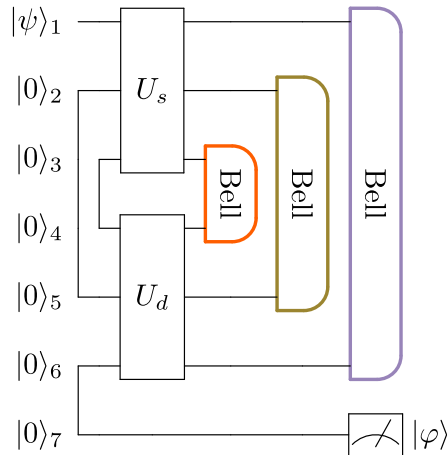


Figure B.3: Circuit depicting the Bell measurement pairs used in Fig. 7.3(a,b).

$\alpha = 0$, the XX-gates combine to create the identity matrix, and when $\alpha = 1$, the unitary corresponds to the maximally scrambling case shown in Fig B.1. Additional Z-rotations are applied around the XX-gates to ensure that $\hat{U}_d = \hat{U}_s^*$. In Fig. 7.3, we measure the delocalization of information throughout the seven-qubit system in the presence of a maximally scrambling unitary $\hat{U}_s(\alpha = 1)$. The Bell measurements used are depicted in Fig. B.3 with the same color scheme.

(a)



(b)
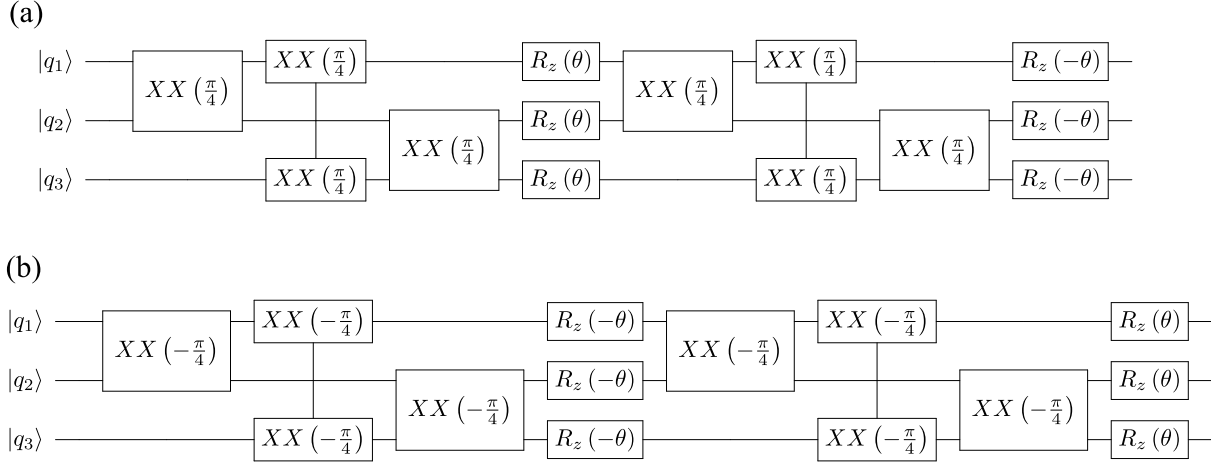


Figure B.4: (a) Circuit for the unitary used in Fig. 7.3 as well as (b) the same unitary with varying degrees of scrambling for the data in Fig. 7.2. The angles of the the Z-rotations are changed according to $\theta = \pm\frac{\alpha\pi}{2}$ to continuously scan between not scrambling ($\alpha = 0$) and maximally scrambling ($\alpha = 1$).

For the data in Fig. 7.4, we compose the scrambling unitary depicted in Fig. B.5. This unitary has the following matrix representation in the computational basis:

$$
U_{CZ} = \frac{1}{2\sqrt{2}}
\begin{pmatrix}
1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 \\
1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
-1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\
-1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 \\
-1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\
-1 & 1 & 1 & 1 & 1 & 1 & 1 & -1
\end{pmatrix}
\tag{B.5}
$$

We can confirm that this is indeed a maximally scrambling unitary via a set of equations analogous to Eq. B.3. The optimized circuit used to implement this unitary experimentally is shown in Fig. B.6. We use a similar circuit to effect the classical scrambling unitary, $\hat{U}_c$, by implementing only the first three controlled-Z gates.

Lastly, the Grover search operator labeled $G$ in Fig. 7.4 is realized using the following circuit:



$$
\tag{B.6}
$$
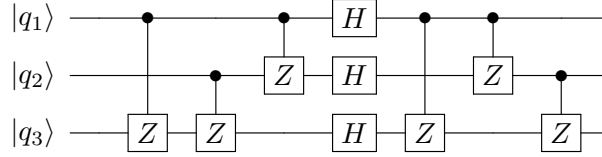
Figure B.5: Circuit representation of the scrambling unitary from Eq. (B.5), used for the data in Fig. 7.4. The breakdown into native gates for the experimental implementation is shown in Fig. B.6. A reduced circuit, made up of only the first three controlled-Z gates, is used to create the classical scrambling unitary $\hat{U}_c$.
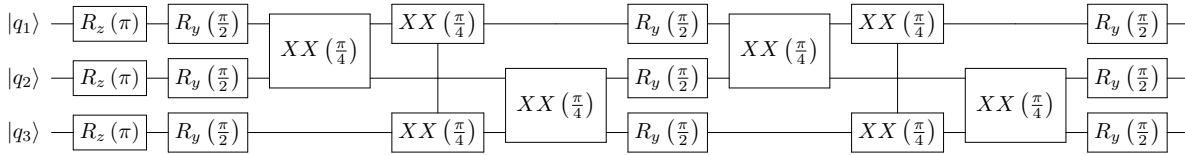


Figure B.6: The scrambling unitary from Eq. (B.5) compiled into native gates. This circuit was used for the measurements in Fig. 7.4.

The SWAP gate (line connecting two X's) is implemented classically by reassigning qubit labels.

## B.3  Numerical simulations

Theory curves were obtained through numerical simulation of the circuits in Appendix B using a simple one-parameter coherent error model. To simulate coherent errors, a random single-(two-)qubit unitary close to the identity was applied following each single-(two-)qubit gate. Single-qubit rotations about the z-axis were performed classically with negligible experimental error and were therefore omitted from this procedure. Random single-(two-)qubit unitary errors were taken as the exponential of a linear combination of the 3 single-(15 two-)qubit traceless Hermitian matrices, with coefficients sampled from a normal distribution with mean 0 and standard deviation $\epsilon/\sqrt{3}$ ($\epsilon/\sqrt{15}$). The resulting observables $\langle P_\psi \rangle$ and $\langle F_\psi \rangle$ were averaged over $N = 10$ realizations of random error, and the same random errors were used for simulation at each experimental parameter. The error strength $\epsilon = .184$ was chosen to minimize the sum of squared errors in $\langle P_\psi \rangle$ and $\langle F_\psi \rangle$.

# Appendix C

# Details on quantum information scrambling on a qutrit superconducting quantum processor

In this Appendix we provide additional details on the experiments and numerics presented in Ch. 7.

## C.1   Processor and fabrication details

The processor we use features five fixed-frequency (single junction) transmon qutrits on a chip with an eight-transmon ring geometry. The readout and coupling resonators, Purcell filter, transmon capacitors, microwave drive lines and ground plane are composed of niobium, while the transmon junctions are aluminum with an aluminum oxide barrier (Fig. 8.1).

The processor is fabricated on intrinsic >8000 ohm-cm silicon <100> wafers. Initial cleaning of the silicon wafer occurs in piranha solution—a mixture of sulfuric acid and hydrogen peroxide—at 120°C for 10 minutes, followed by 5:1 buffered oxide etch (BOE) for 30 seconds to remove surface contaminants and native oxides. A 200-nm thick film of niobium is then sputtered onto the wafer, with deposition pressures optimized to yield a slightly compressive film. Following this, junctions and all other structures are patterned using 3 rounds of electron-beam lithography. We use MicroChem MMA EL-13 copolymer as a resist, developing it in a 3:1 mixture of IPA:MIBK (isopropyl alcohol and methyl isobutyl ketone) at room temperature for 8 minutes. We then etch the niobium with chlorine chemistry in an inductively coupled reactive ion etcher with about 50 nm overetch into the silicon. After etching, resist is removed with Microposit 1165 at 80°C for 60 min. The Josephson junction fabrication process begins by stripping native Nb and Si oxides with 30 seconds in BOE. Resist is then spun: we use 500 nm of MMA EL-13 and 150 nm of AllResist AR-P 6200.9, both baked at 150°C for 60 and 90 seconds, respectively. We write "Manhattan style" junction patterns [110, 374] (proximity-effect-corrected with Beamer by Genisys software)

at 100 keV in a Raith EBPG 5150 using a 200 pA beam current and 200 $\mu$m aperture. After writing, the exposed AR-P is first developed in n-amyl acetate chilled to 0°C; after this the AR-P development is halted with 10s immersion in IPA; finally MMA is developed in 3:1 IPA:MIBK for 10 min. We then dry the resulting structure with $N_2$ and descum it with an 80W, 200 mbar oxygen plasma etch. This etching step is split into 4 separate substeps, with 90 degree substrate rotations between each substep for improved junction uniformity. Newly-formed oxides at the bottom of the developed structure are then removed with a 15s dip in BOE. The wafer is then loaded into a Plassys MEB550s evaporator and pumped overnight before the junction evaporation steps: first, an Al base electrode is evaporated and the tunnel barrier then formed by thermal oxidiation, introducing a 95%/5% Ar/O mix into the chamber at 10 mbar for 10 min. A second aluminum electrode is then evaporated to complete the junction and a third evaporation is necessary to climb the second 250 nm capacitor step edge. The junction pattern includes a 6 x 8 $\mu$m Al wire on top of the Nb for electrical contact between the junction and capacitor. After liftoff for 2 hours in acetone at 67°C, the same resist stack is spun, and 10 x 15 $\mu$m rectangles are opened over the Al/Nb overlap region. The exposed metals are then ion milled and Al is subsequently e-beam evaporated to ensure a low loss galvanic connection between Nb and Al [140]. More details on junction fabrication, including the steps leading to higher uniformity, can be found in [269]. After fabrication, the wafer is diced into 1x1 cm dies; cleaned in Microposit 1165 for 12 hours at 80C; sonicated in DI water, acetone, and IPA; descummed in 100 W oxygen plasma for 1 min and then wirebonded into a gold plated copper cryopackage on a 300 $\mu$m air gap.

Each transmon is coupled to ($i$) a linear readout resonator to enable multiplexed dispersive measurement, ($ii$) two coupling resonators to enable entangling interactions with nearest neighbors, and ($iii$) a microwave drive line. Readout resonators are loaded so that their effective linewidth $\kappa_{\mathrm{ext}} \approx 1$ MHz. All readout resonators on the chip are coupled to a common $\lambda/2$ resonator, a Purcell filter with an external Q $\approx 10$ [434]. The Purcell filter's passband overlaps with all readout resonator frequencies, allowing fast readout; all qutrit frequencies lie outside the passband, suppressing qutrit relaxation through this channel. Slotline modes of all structures are suppressed using wirebonds; a wirebond also enables the readout bus to overlap a coupling resonator.

## C.2 Experimental setup

The processor is installed in the 10 mK stage of a BlueFors XLD dilution refrigerator. Room-temperature and cryogenic electronics for performing control and measurement of the qutrit chip are shown in Fig. C.2. A Holzworth multi-channel synthesizer generates three local oscillator tones: a qubit control LO at 4.72 GHz, a readout LO at 6.483 GHz, and a pump at 7.618 GHz for the traveling-wave parametric amplifier (TWPAs). Qutrit control pulses are formed by IQ modulating the amplified qubit LO (split six ways) with IF signals from a Tektronix AWG (sample rate 2.5 GS/s) with frequencies between 0.5 and 1.1 GHz. We
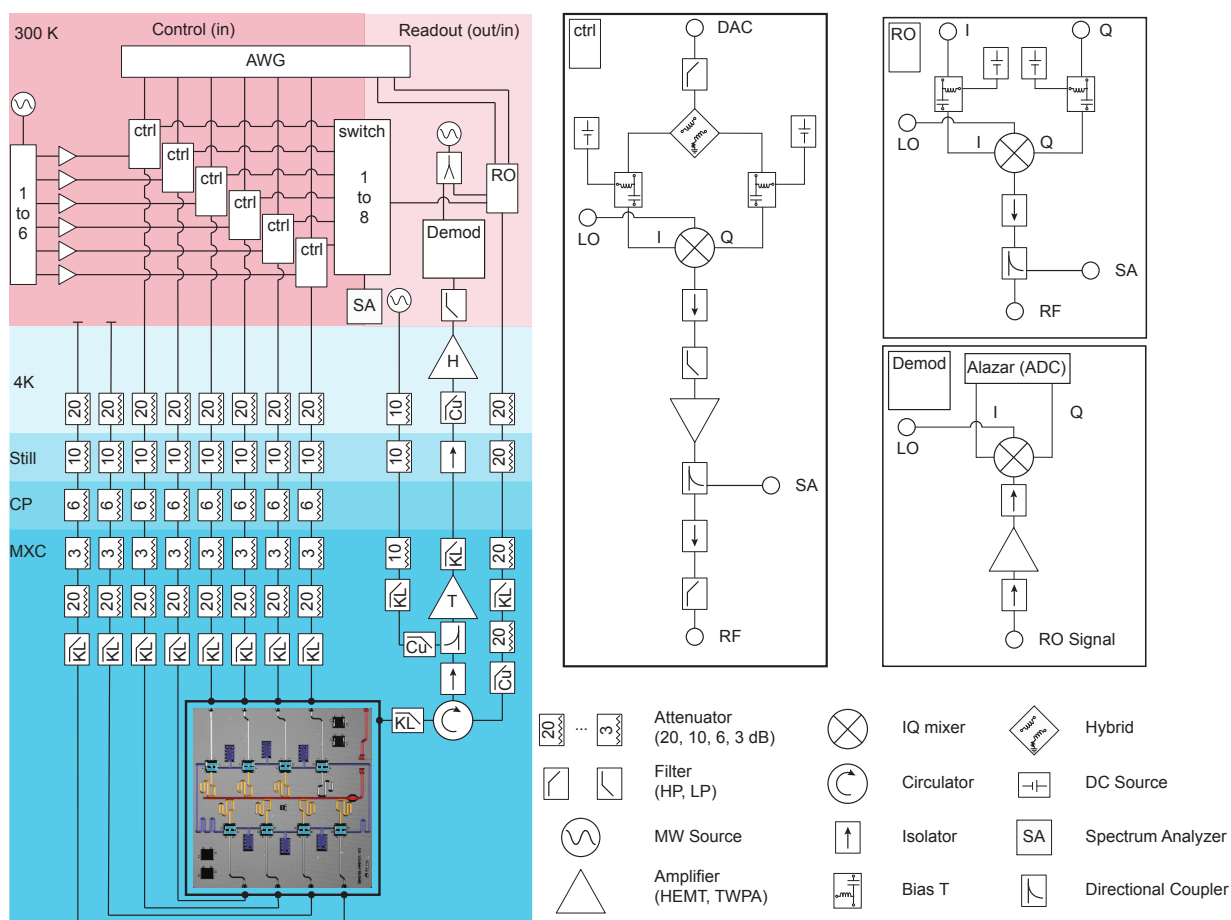
Figure C.1: Experimental setup described in detail in the Ch. 8.

use both single-channel, hybrid-enabled SSB modulation and high-pass filtering to eliminate the lower sideband of the pulse, with additional band-pass filtering at room-temperature to eliminate noise from the AWG itself. Readout signals are generated using two-channel SSB modulation with IF signals from the same Tektronix AWG. All input signals are further attenuated in the cryostat.

Readout signals are amplified by the TWPA at 10 mK, high-electron mobility transistor amplifiers (HEMT) at 4K, and further amplification at room-temperature before being digitized at 1.25 GSa/s and demodulated in software.

| | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ |
|---|---|---|---|---|---|
| Qutrit $|0\rangle \leftrightarrow |1\rangle$ frequency, $\omega_{01}/2\pi$ (GHz) | 5.447 | 5.634 | 5.776 | 5.619 | 5.431 |
| Qutrit $|1\rangle \leftrightarrow |2\rangle$ frequency, $\omega_{12}/2\pi$ (GHz) | 5.177 | 5.368 | 5.512 | 5.351 | 5.160 |
| Readout frequency, $\omega_{\mathrm{RO}}/2\pi$ (GHz) | 6.384 | 6.324 | 6.731 | 6.673 | 6.618 |
| Lifetime $T_1^{|1\rangle \to |0\rangle}$ ($\mu$s) | 70 | 49 | 43 | 55 | 63 |
| Lifetime $T_1^{|2\rangle \to |1\rangle}$ ($\mu$s) | 38 | 29 | 39 | 32 | 36 |
| Ramsey decay time $T_2^*$, $|1\rangle/|0\rangle$ ($\mu$s) | 73 | 13 | 41 | 48 | 20 |
| Ramsey decay time $T_2^*$, $|2\rangle/|1\rangle$ ($\mu$s) | 13 | 10 | 16 | 23 | 10 |
| Ramsey decay time $T_2^*$, $|2\rangle/|0\rangle$ ($\mu$s) | 16 | 6 | 15 | 26 | 11 |
| Echo time $T_{2\mathrm{Echo}}$, $|1\rangle/|0\rangle$ ($\mu$s) | 71 | 51 | 46 | 64 | 74 |
| Echo time $T_{2\mathrm{Echo}}$, $|2\rangle/|1\rangle$ ($\mu$s) | 29 | 22 | 22 | 35 | 32 |
| Echo time $T_{2\mathrm{Echo}}$, $|2\rangle/|0\rangle$ ($\mu$s) | 39 | 26 | 34 | 45 | 39 |
| Readout fidelity, $|0\rangle$ | 0.99 | 0.99 | 0.97 | 0.98 | 0.99 |
| Readout fidelity, $|1\rangle$ | 0.97 | 0.95 | 0.94 | 0.95 | 0.96 |
| Readout fidelity, $|2\rangle$ | 0.95 | 0.94 | 0.92 | 0.95 | 0.96 |
| Per-Clifford error, $|1\rangle/|0\rangle$ subspace | 3.6e-4 | 3.9e-4 | 5.5e-4 | 2.7e-4 | - |
| Per-Clifford error, $|2\rangle/|1\rangle$ subspace | 3.6e-4 | 6.0e-4 | 5.0e-4 | 7.5e-4 | - |

Table C.1: Measured properties of the five qutrits

## C.3 Chip characterization

Transmon parameters are given in Table C.1. The frequencies are extracted using standard spectroscopy methods. Lifetimes are extracted by fitting decay curves to a single model with five parameters: two energy-relaxation times ($T_1^{1\to0}$ and $T_1^{2\to1}$) and a dephasing time for each basis state. We perform randomized benchmarking to measure pulse errors of single qubit operations in the different subspaces, shown in the table.

We further measure the coefficients of the cross-Kerr (or 'ZZ') interaction by performing a Ramsey measurement with neighboring qutrits in the $|1\rangle$ or $|2\rangle$ states. The cross-Kerr Hamiltonian between neighboring qutrits is

$$H_{\mathrm{Kerr}}/\hbar = \alpha_{11}|11\rangle\langle 11| + \alpha_{12}|12\rangle\langle 12| + \alpha_{21}|21\rangle\langle 21| + \alpha_{22}|22\rangle\langle 22| \tag{C.1}$$

Table C.2 gives the value of these coefficients. Residual cross-Kerr interaction coefficients between non-nearest-neighbor transmons were found to be negligible.

The cross-Kerr interaction is the dispersive limit of an exchange interaction between nearest-neighbor transmons mediated by the coupling resonators, governed by the Hamiltonian

$$H_{\mathrm{int}}/\hbar = g\left(a^\dagger b + b^\dagger a\right) \tag{C.2}$$

|              | $Q_1/Q_2$ | $Q_2/Q_3$ | $Q_3/Q_4$ | $Q_4/Q_5$ |
|--------------|-----------|-----------|-----------|-----------|
| $\alpha_{11}$ | -279      | -138      | -276      | -262      |
| $\alpha_{12}$ | 160       | 158       | -631      | -495      |
| $\alpha_{21}$ | -528      | -335      | 243       | -528      |
| $\alpha_{22}$ | -743      | -342      | -748      | -708      |

Table C.2: Cross-Kerr interaction coefficients between nearest-neighbor transmons, given in kilohertz

We measure the value of $g$ on a chip with tunable transmons, but otherwise identical to the one used in the experiment. Spectroscopy of the avoided crossing reveals an interaction amplitude $g$ of roughly 3 MHz.

## C.4 Coherence of third transmon level

The dominant noise processes affecting transmons tend to worsen for states higher up the transmon ladder. For our qutrit-based processor there are two salient manifestations of this:

- Due to bosonic enhancement, amplitude damping (spontaneous emission) decay from state $|2\rangle$ to $|1\rangle$ proceeds roughly twice as fast as the decay from $|1\rangle$ to $|0\rangle$.

- Dephasing due to charge noise, which randomizes the relative phases between the $|0\rangle$, $|1\rangle$, and $|2\rangle$ states, occurs roughly an order of magnitude faster for each state up the transmon ladder: in particular, for the $|2\rangle$ state relative to the $|1\rangle$ state.

As stated in Ch. 8, careful fabrication, microwave engineering, and parameter selection were required to obtain high coherence in the transmon qutrit. The fabrication and microwave engineering are detailed in the following section, and served to mitigate the $T_1$ decay. Here we describe the parameter selection – specifically, the choice of the transmon $E_J/E_C$ ratio – which was chosen to minimize the effect of dephasing.

Transmons are characterized by two parameters: the Josephson energy $E_J$, and the capacitive energy $E_C$ [258]. Increasing the $E_J/E_C$ ratio exponentially decreases the sensitivity of all transmon eigenstates to charge noise, at the expense of also lowering the transmon's anharmonicity. Specifically, the charge dispersion $\epsilon_m$ of the $m^{\text{th}}$ level is given by

$$\epsilon_m \approx (-1)^m E_C \frac{2^{4m+5}}{m!} \sqrt{\frac{2}{\pi}} \left( \frac{E_J}{2E_C} \right)^{\frac{m}{2}+\frac{3}{4}} e^{-\sqrt{8E_J/E_C}}, \tag{C.3}$$

while the relative anharmonicity $\alpha_r$ is given by

$$\alpha_r \approx -(8E_J/E_C)^{-1/2}. \tag{C.4}$$

Typical transmon qubit designs use ratios $E_J/E_C \approx 50$. We initially used such a ratio, which resulted in charge dispersion of 102 kHz and <10 kHz of $|2\rangle$ and $|1\rangle$ states, respectively. However, with these parameters, charge-parity fluctuations [396] dephase the coherence between the $|2\rangle$ and $|1\rangle$ states within 5 $\mu$s, making high-fidelity gates impossible to implement. To mitigate this dephasing, we switched to a design with $E_J/E_C \approx 73$, which resulted in charge dispersions of 12 kHz and 261 Hz for the $|2\rangle$ and $|1\rangle$ states, respectively. This also reduced the anharmonicity from roughly 300 MHz to roughly 250 MHz.

*Crosstalk.*—As discussed in Ch. 8, each transmon features a dedicated microwave control line through which we drive single- and two-qutrit gates. However, we find significant (order unity compared to intended coupling) crosstalk between the microwave drive lines for each qutrit. This crosstalk is non-local, not confined to nearest or next-nearest neighbors. When driving Rabi oscillations on a given qutrit, it produces two unwanted effects:

1. All other qutrits will be off-resonantly Rabi driven. Depending on the relative frequencies between the qutrits, this can either manifest as an unwanted change in a qutrit's state populations (if the frequencies are relatively close) or a Stark shift (if the frequency difference is large compared to the Rabi frequency).

2. Microwave field leaking onto one or more neighboring qutrits will result in an unwanted cross-resonance interaction, making the desired Rabi oscillation frequency vary with the state of the neighboring qutrit(s). This effect was anticipated in [394].

We observed no indications of nonlinearity in the cross-talk at the drive powers we used. That is, for a given drive frequency, the cross-talk can be characterized in terms of a five-by-five complex-valued matrix $C(\omega)$ relating the field amplitudes $\mathbf{o}(\omega)$ seen by each of the five qutrits to the input field amplitudes $\mathbf{i}(\omega)$ on each drive line: $\mathbf{o}(\omega) = C(\omega)\mathbf{i}(\omega)$. We did observe a strong frequency-dependence of the cross-talk matrix.

The linearity of the cross-talk enabled us to compensate for it by inverting the matrix $C(\omega)$ at each drive frequency, yielding combinations of microwave drive lines which would route the drive field to only a single qutrit. The main challenge in this scheme was the measurement of $C(\omega)$. Our strategy was to focus on two drive lines at a time, and find for each line the relative amplitudes and phases which exactly cancelled the field at the location of all of the qutrits on our chip—depending on the relative frequencies, we used either a Stark shift or a Rabi oscillation as a symptom of unwanted microwave field. This measurement was repeated for each of ten drive frequencies of interest (i.e. the $|0\rangle \leftrightarrow |1\rangle$ and $|1\rangle \leftrightarrow |2\rangle$ transition frequencies of all five qutrits), each pair of lines, and each qutrit on the chip.

Our crosstalk cancellation method is extremely measurement-intensive and was feasible only because of the relatively few qutrits in this work. On future quantum processors with tens or hundreds of quantum systems, the number of measurements required for our cancellation scheme would be prohibitively expensive. In addition, the strong frequency dependence of the cross-talk matrix limits the speed at which one can apply single-qudit pulses in this manner: for pulses approximately 10 ns in length, we observed cross-talk which we could not compensate for using our method, likely because of this frequency dependence combined

with Fourier broadening of the pulses. Going forward, it is thus important to pinpoint the source of the microwave cross-talk, in order to develop scalable solutions at the hardware level.

# C.5 Qutrit Rotations and gate-set

A convenient set of generators to describe qutrit rotations are the Gell-Mann matrices:

$$\lambda_1 \equiv s_x^{01} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 \equiv s_y^{01} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_3 \equiv s_z^{01} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 \equiv s_x^{02} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda_5 \equiv s_y^{02} = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda_6 \equiv s_x^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 \equiv s_y^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

They are the generators of the Lie algebra of the special unitary group SU(3) and can be thought of as the natural extension of Pauli matrices (generators of the Lie algebra of the SU(2) group). For each qutrit (with basis states $|0\rangle, |1\rangle, |2\rangle$), we calibrate a set of microwave pulses that resonantly drive the $|0\rangle \leftrightarrow |1\rangle$-transition and a separate set of pulses to address the $|1\rangle \leftrightarrow |2\rangle$-transition, providing universal control over the qubit subspaces $\{|0\rangle, |1\rangle\}$ and $\{|1\rangle, |2\rangle\}$. Our microwave control pulses directly perform rotations that correspond to exponentiating Gell-Mann matrices $s_x^{01}, s_y^{01}, s_z^{01}, s_x^{12}, s_y^{12}$. The Z-rotation ($s_z^{01}$) is implemented as a virtual Z gate in software by adjusting the phases of subsequent microwave pulses in that subspace [324]. We extend this technique to the 12 subspace to also obtain the following rotation that is not one of the Gell-Mann matrices but that is very useful for single qutrit control since it is a virtual rotation with negligible error:

$$s_z^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{C.5}$$

In principle one could also drive the $|0\rangle \leftrightarrow |2\rangle$-transition to directly implement rotations corresponding to $s_x^{02}, s_y^{02}$. While it would be worthwhile to add these rotations to the available gate-set to compile circuits with lower depth, these two-photon transitions are more challenging to address as they require high-power. Luckily, all rotations generated by the remaining Gell-Mann matrices can be constructed from our available operations, for example

$$e^{-i\frac{\theta}{2}s_{x/y}^{02}} = e^{-i\frac{\pi}{2}s_x^{12}} \cdot e^{-i\frac{\theta}{2}s_{x/y}^{01}} \cdot e^{i\frac{\pi}{2}s_x^{12}} \tag{C.6}$$

where the right most operator is the first to act on the state, so time goes from right to left. Similarly, $\lambda_8$ can be constructed from $s_z^{01}$ and $s_z^{12}$

We write a rotation in one of these subspaces as $\theta_j^k = e^{-i\frac{\theta}{2}s_j^k}$, where $k = \{01, 12\}$ defines the subspace of the rotation, $j = \{x, y, z\}$ the rotation axis and $\theta$ the rotation angle. As an example, the two available x-rotations and their corresponding rotation matrices are

$$\theta_x^{01} = \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 & 0 \\ -i\sin\theta/2 & \cos\theta/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \theta_x^{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta/2 & -i\sin\theta/2 \\ 0 & -i\sin\theta/2 & \cos\theta/2 \end{pmatrix}, \tag{C.7}$$

This notation, combined with some useful qutrit and experiment specific operations, is also adopted in circuit diagrams displayed in Ch. 8 and in this document.

Our gate-set consists of all Z-rotations along an arbitrary angle, combined with the Clifford operations operations in the 01 and 12 subspace :

$$\theta_j^k \text{ with } \theta = \{\pi, -\pi, \frac{\pi}{2}, \frac{-\pi}{2}\}, k = \{01, 12\} \text{ and } j = \{x, y, z\}. \tag{C.8}$$

Three convenient gates to describe qudit-logic, which can be constructed from our universal gate-set, are the $X$ and $Z$ gates

$$X|i\rangle = |i + 1 \mod d\rangle \tag{C.9}$$

$$Z|i\rangle = \omega^i |i\rangle, \tag{C.10}$$

where $\omega = \exp(i2\pi/d)$.

and the Hadamard gate

$$H = \frac{1}{\sqrt{d}} \sum_{i,j} \omega^{ij} |i\rangle\langle j| \tag{C.11}$$

# C.6 Controlled-SUM Gate

In general $d$-dimensional qudits, two-qudit controlled-SUM and controlled-phase gates can be defined using the Pauli $X$ and $Z$ gates We have:

$$U_{\text{CSUM}}^{01} = \sum_{n=1}^{d} |n\rangle\langle n| \otimes X^n \text{ and } U_{\text{C}\phi} = \sum_{n=1}^{d} |n\rangle\langle n| \otimes Z^n \tag{C.12}$$

Here the superscript 01 indicates that the controlled-SUM is applied with $Q_0$ as the control qudit and $Q_1$ as the target; such a label is not necessary for $U_{\text{C}\phi}$, which is symmetric between the two qudits. The two gates are equivalent up to a single-qudit Hadamard gate $H$ on the second qudit:

$$(I \otimes H^\dagger)U_{\text{C}\phi}(I \otimes H) = U_{\text{CSUM}}^{01}, \tag{C.13}$$

where the qudit Hadamard gate is defined to transform the $Z$ gate into the $X$ gate under conjugation. Reversing the order of $H$ and $H^\dagger$ yields a controlled-MINUS gate, and changing
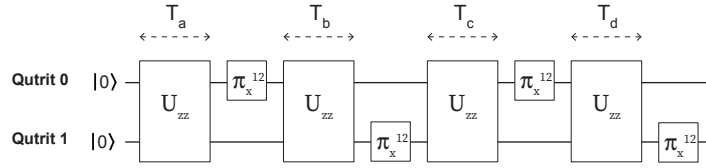
Figure C.2: Four-segment pulse sequence implementing $U_{C\phi}$. Four local $\pi$-pulses that effectively permute the eigenstates are interspersed with periods of free evolution $T_A, T_B, T_C, T_D$. The times depend on the cross-Kerr interaction parameters and are chosen such that the pulse sequence implements the desired diagonal unitary operation.

which qubit receives the conjugation interchanges the control and target. The entangling gates $U_{C\phi}$, $U_{\text{CSUM}}^{01}$, $U_{\text{CSUM}}^{10}$, $U_{\text{CMIN}}^{01}$, and $U_{\text{CMIN}}^{10}$ are therefore all equivalent up to local (single-qudit) operations.

In our system, we directly implement the two-qutrit $U_{C\phi}$ gate by interspersing periods of evolution under the cross-Kerr Hamiltonian (Eq. C.1) with single-qutrit gates. Intuitively, evolution under the cross-Kerr Hamiltonian imparts phases to the two-qutrit states $|11\rangle$, $|12\rangle$, $|21\rangle$, and $|22\rangle$, with values determined by the coefficients $\alpha_{ij}$. By interspersing this phase accumulation with single-qutrit pulses exchanging the various states, we can ensure that each state accumulates exactly the phase required for the controlled-phase gate.

We present two methods for implementing the controlled-phase gate in the manner described above. The first uses fewer single-qutrit pulses and is conceptually simpler, but is not dynamically decoupled from the cross-Kerr interaction with neighboring qutrits. The second is dynamically decoupled and is the one used in the teleportation experiment.

**First method:** As depicted in Fig. C.2, here we use four periods of cross-Kerr evolution, separated by pulses swapping the $|1\rangle$ and $|2\rangle$ states of a single qutrit. Denoting this swap pulse as $\Pi_q^{12}$, where $q$ is the qutrit number, and evolution under the cross-Kerr Hamiltonian for a time $T$ as $ZZ_T$, the total pulse sequence is

$$ZZ_{T_A} \cdot \Pi_0^{12} \cdot ZZ_{T_B} \cdot \Pi_1^{12} \cdot ZZ_{T_C} \cdot \Pi_0^{12} \cdot ZZ_{T_D} \cdot \Pi_1^{12}. \tag{C.14}$$

where the times $T_A$, $T_B$, $T_C$, $T_D$ depend on the cross-Kerr interaction parameters $\alpha_{ij}$. For any choice of times, this operation imparts zero phase to the states $|00\rangle$, $|01\rangle$, $|02\rangle$, $|20\rangle$, $|10\rangle$, and non-zero relative phases $\phi_{11}$, $\phi_{12}$, $\phi_{21}$, and $\phi_{22}$ to the other basis states. These phases are linear combinations of the delay times $T_A$, $T_B$, $T_C$, and $T_D$. The transformation from delay times to induced phases is full rank (except for pathological values of the cross-Kerr coefficients), meaning that, given enough total delay time, this method can in principle generate an arbitrary two-qudit phase gate (the states that receive zero phase above can be made to gain arbitrary phase using only single-qutrit phase gates). On our particular chip, the coefficients $\alpha_{ij}$ allow us to implement the controlled-phase in this manner in roughly 1.5 $\mu$s for qutrit pairs $Q_1/Q_2$ and $Q_3/Q_4$.

The drawbacks of this method become apparent when one tries to use it in a multi-qutrit algorithm. If the two qutrits undergoing the controlled-phase are coupled to other qutrits via the same cross-Kerr Hamiltonian (as is the case for our chip), the above method will not work when the other qutrits are in superpositions of basis states, in which case entanglement between them and the desired qutrits will be generated. The second method addresses this problem.

**Second method:** As depicted in Fig. C.3 a, this method uses six *equal* time periods of cross-Kerr evolution. These are interspersed with single-qutrit pulses swapping the $|0\rangle/|1\rangle$ and $|1\rangle/|2\rangle$ subspaces, denoted $\Pi_q^{12}$ and $\Pi_q^{01}$, respectively. The total pulse sequence consists of three repetitions of:

$$\left[ ZZ_T \cdot \left( \Pi_0^{12} \otimes \Pi_1^{12} \right) \cdot ZZ_T \cdot \left( \Pi_0^{01} \otimes \Pi_1^{01} \right) \right] \tag{C.15}$$

For specificity, we have parameterized this pulse sequence with a single delay time, $T$; an appropriately chosen $T$ realizes the controlled-phase gate. The delay time $T$ is determined by the values of the cross-Kerr coefficients $\alpha_{ij}$ for each pair $Q_1/Q_2$ and $Q_3/Q_4$, and thus differs between the pairs; however, in practice, we find that a delay of 192 ns works well for both.

This pulse sequence constitutes a dynamically-decoupled implementation of the $U_{C\phi}$ unitary, as its operation is successful regardless of the states of the neighboring qutrits. The dynamical decoupling arises because the single-qutrit pulses shuffle each qutrit's states $|0\rangle$, $|1\rangle$, and $|2\rangle$ such that an equal amount of time is spent in each state, regardless of the initial state of either qutrit. This shuffling 'averages out' the cross-Kerr interaction with neighboring qutrits, such that no entanglement is generated.

The particular teleportation algorithm we implement requires applying $U_{C\phi}$ on two pairs of qutrits $Q_1/Q_2$ and $Q_3/Q_4$. We use this dynamically decoupled pulse sequence for both pairs, and apply the gates simultaneously to reduce decoherence associated with a longer total gate time. Naively, the dynamical decoupling effect is weakened by this simultaneity, since the 'neighboring qutrits', with respect to each individual pair, are no longer static. Fortunately, we verify both theoretically and empirically that we can nevertheless decouple the unwanted interaction by reversing the order of the $\Pi^{12}$ and $\Pi^{01}$ gates between the two pairs (Fig. C.3 b).

## C.7 Dynamically-decoupled EPR preparation

We prepare the two intial EPR pairs of the teleportation algorithm using the controlled-$\pi$ gate as discussed in Ch. 8. The basic sequence is presented in Fig. 8.3, and serves to prepare an EPR pair on either $Q_2/Q_3$ or $Q_4/Q_5$ individually, while all other qutrits are in the ground state $|0\rangle$. Simultaneous EPR pair preparation, as required by the algorithm, necessitates a more complicated sequence that incorporates dynamical decoupling. This necessity is demonstrated by Fig. C.4(a-b), which compares the result of individual EPR preparation to joint EPR preparation without dynamical decoupling. Joint preparation fidelities are much
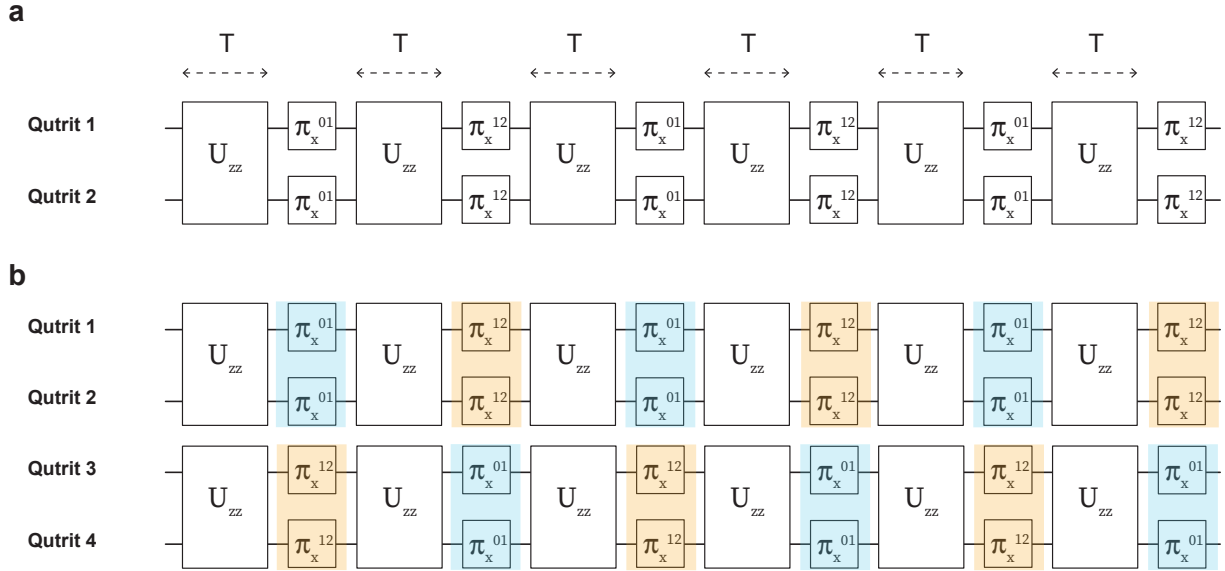
Figure C.3: Six-segment pulse sequence implementing $U_{C\phi}$, dynamically-decoupled from static neighbors. **(a)** By alternating six periods of free evolution with local permutation pulses, a diagonal phase gate can be implemented while also protecting the two qutrits from dephasing and static interactions with neighbors provided that they are static. **(b)** To perform two of these gates in parallel and maintain the decoupling, in this case between qutrit 2 and qutrit 3, the order of local permutations is swapped between pairs: qutrit 1 and qutrit 2 first are permuted in the $|0\rangle/|1\rangle$ subspace (blue) while the first local operation on qutrit 3 and 4 is in the $|1\rangle/|2\rangle$ subspace (orange). This effectively decouples qutrit 2 and 3 in addition to performing the desired gates and decoupling qutrit 1 and 4 from their other neighbors (0 and 5).

lower than those of individual preparation. From the measured density matrices, this loss seems to be largely due to a decrease in the off-diagonal elements (i.e. the coherences).

To understand the source of this decrease in coherence, we measured the density matrix of the $Q_2/Q_3$ EPR pair while projecting the neighboring qutrit, $Q_4$, into each of its basis states. Each of the three conditional density matrices we obtained was much purer (i.e. had much higher coherence) than the unconditional density matrix; however, the phases of each coherence differed depending on the state of $Q_4$. These measurements suggest that the source of the decoherence was indeed unwanted entanglement between $Q_3$ and $Q_4$ arising from the cross-Kerr interaction.

Qutrit State Tomography [56] after each step of the EPR preparation sequence allows us to pinpoint the portions of the sequence that contribute most strongly to the unwanted entanglement. The cross-Kerr interaction affects the $|2\rangle$ states most strongly, and we find correspondingly that most of the entanglement occurs after the $|2\rangle$ state of $Q_3$ gets populated.

Figure C.4: Dynamically-decoupling the EPR pair preparation. **(a)** Density matrices of the $Q_2/Q_3$ (left, purple) and $Q_4/Q_5$ (right, orange) EPR pairs, prepared individually (i.e. with all other qutrits in the ground state). State fidelities for this dataset are $0.94 \pm 0.002$ and $0.98 \pm 0.002$ respectively. **(b)** Density matrices of the same EPR pairs when prepared simultaneously without any dynamical decoupling. Fidelites are markedly lower in this case, $0.81 \pm 0.002$ and $0.82 \pm 0.002$ respectively for the $Q_2/Q_3$ and $Q_4/Q_5$ pairs. As discussed in the text, the loss of fidelity is due to unwanted entanglement arising from the cross-Kerr interaction between the two EPR pairs. **(c)** EPR pairs prepared simultaneously using dynamical decoupling, with fidelities $0.88 \pm 0.002$ and $0.92 \pm 0.002$ , respectively.
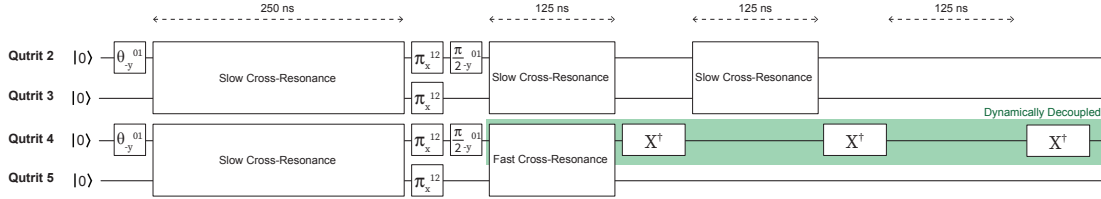
Figure C.5: Dynamically-decoupled gate sequence. The initial step in joint EPR preparation is simultaneous preparation of Bell pairs on both pairs of qutrits. Since the cross-Kerr interaction in the $|0\rangle/|1\rangle$ subspace is relatively weak compared to the interaction time, little unwanted entanglement occurs during this operation. Following Bell state creation, the remaining evolution is divided into three equal periods of 125 ns, during which the cross-resonance interaction completes the creation of the EPR pairs. In between each period, qutrit $X$ gates applied on $Q_4$ serve to dynamically decouple the unwanted cross-Kerr interaction.

We take advantage of this by only dynamical decoupling the cross-Kerr interaction after this point. As shown in Fig. C.5, the initial preparation of Bell states $(|00\rangle + |11\rangle)/\sqrt{2}$, which does not involve the state $|2\rangle$, is performed without dynamical decoupling to reduce the error associated with additional single-qutrit gates.

The mechanism underlying our decoupling sequence is most easily understood by first considering a simpler problem, of decoupling an unwanted cross-Kerr interaction between two qutrits during an idling period. This can be accomplished by splitting the idling time into three equal time periods, and applying single-qutrit $X$ gates to one of the qutrits between each of the periods. This shuffling of the populations decouples the entangling interaction into a product of local $Z$ interactions. Using the same principle, we divide the controlled-$\pi$ operations in the relevant portion of simultaneous EPR preparation into three equal periods of 125 ns, and apply qutrit $X$ gates on $Q_4$ in between the periods. This sequence enables simultaneous EPR pair preparation with fidelities $0.88 \pm 0.002$ and $0.92 \pm 0.002$ on $Q_2/Q_3$ and $Q_4/Q_5$, respectively.

# Appendix D

# Details on many-body quantum teleportation

In this Appendix we provide additional details on many-body quantum teleportation in the traversable wormhole protocol, discussed in Ch. 9. In particular, we provide: rigorous lower bounds on the teleportation fidelity from the operator size distribution; a detailed overview of the connection between the traversable wormhole protocol and the Hayden-Preskill recovery protocol (the subject of Ch. 7 and 8); expressions for the fidelity of quantum state teleportation; details on the numerical results of Ch. 9; and many-body quantum teleportation of fermions.

## D.1 Precise bound for the peaked size regime

Here we provide a precise mathematical bound guaranteeing that the teleportation correlator obeys the peaked-size prediction [Eq. (9.33), Section 9.4] when the size distribution is sufficiently tightly peaked. We apply this bound to two examples where the size distribution is known exactly: late times in all scrambling systems (Section 9.6), and the large-$q$ SYK model (Sections 9.7 and 9.8). Notably, in the latter we find that our bound applies *only* at infinite temperature, despite the profile of the size distribution (e.g. its ratio of size width to average size) behaving similarly at all temperatures. The discrepancy arises instead because the correlator magnitude, $(G_\beta)^p$, decreases exponentially in the encoding size $p$ at all finite temperatures.

### Precise bound

As in the Ch. 9, we decompose a time-evolved finite temperature operator into a sum of Pauli strings:

$$Q_A(t)\rho^{1/2} = \sum_R c_R(t)S \tag{D.1}$$

In this basis, for qubit systems the correlator takes the form

$$
\begin{aligned}
C_Q &= \langle \text{TFD} | \, \tilde{Q}_{A,r}^\dagger(-t) e^{igV} Q_{A,l}(t) \, | \text{TFD} \rangle \\
&= e^{ig + i\pi \mathcal{S}[Q_A(t=0)]} \sum_R e^{-i\eta_d g \mathcal{S}[R]/N} c_R^2(t) \\
&= e^{ig + i\pi \mathcal{S}[Q_A(t=0)]} \sum_n e^{i\eta_d g n/N} f(n)
\end{aligned}
\tag{D.2}
$$

where again $\tilde{Q}_{A,r}^\dagger = D Q_{A,r}^\dagger D^\dagger$ for the decoding operation $D = Y \otimes \ldots \otimes Y$, and we use $\langle \text{TFD} | \, \tilde{Q}_{A,r}^\dagger(-t) = e^{i\pi \mathcal{S}[Q_A]} \langle \text{EPR} | \, Q_{A,l}(t) \rho^{1/2}$ for qubit Pauli operators $Q_A$. Here we define the winding size distribution [71, 350]

$$
f(n) \equiv \sum_{S:\mathcal{S}[R]=n} c_R^2(t).
\tag{D.3}
$$

At finite temperature, this size wavefunction is distinct from the size distribution:

$$
P(n) \equiv \sum_{S:\mathcal{S}[R]=n} |c_R(t)|^2,
\tag{D.4}
$$

which is a real, normalized probability distribution probed by the *one*-sided correlator [379]

$$
\langle \text{TFD} | \, Q_{A,l}^\dagger(t) e^{igV} Q_{A,l}(t) \, | \text{TFD} \rangle = e^{ig} \sum_R e^{-i\eta_d g \mathcal{S}[R]/N} |c_R|^2(t) = \sum_n e^{i\eta_d g n/N} P(n).
\tag{D.5}
$$

Nevertheless, the size distribution bounds the size wavefunction magnitude via the triangle inequality:

$$
|f(n)| \leq P(n),
\tag{D.6}
$$

with equality achieved when all Pauli operators of size $n$ contribute the same phase to $f(n)$.

The average size and size variance are easily found from the size distribution as

$$
\mathcal{S} = \int_0^\infty dn \, n \, P(n), \quad \delta \mathcal{S}^2 + \mathcal{S}^2 = \int_0^\infty dn \, n^2 \, P(n)
\tag{D.7}
$$

where we work in the continuum limit replacing sums over the size by integrals for simplicity. We now define the *asymptotic size width with error* $\varepsilon$ as the minimal width $W_\varepsilon$ about the average size such that

$$
1 - \int_{\mathcal{S}-W_\varepsilon}^{\mathcal{S}+W_\varepsilon} dn \, P(n) \leq \varepsilon,
\tag{D.8}
$$

i.e. a fraction $1 - \varepsilon$ of the size distribution's support is contained in the interval $I = [\mathcal{S} - W_\varepsilon, \mathcal{S} + W_\varepsilon]$ (the lower limit of the integral should be bounded by zero; for simpler notation we'll deal with this by instead defining $P(n) = f(n) = 0$ for $n < 0$). We can now

separate the correlator into two pieces, one arising from sizes in the interval $I$ and the other from the interval's complement $\bar{I} = [-\infty, \mathcal{S} - W_\varepsilon] \cup [\mathcal{S} + W_\varepsilon, \infty]$:

$$C_Q = \int_I dn\, f(n) e^{i\eta_d g n/N} + R \tag{D.9}$$

where the remainder $R = \int_{\bar{I}} dn\, f(n) e^{i\eta_d g n/N}$ is strictly smaller than $\varepsilon$:

$$
\begin{aligned}
|R| &= \left| \int_{\bar{I}} dn\, f(n) e^{i\eta_d g n/N} \right| \\
&\leq \int_{\bar{I}} dn\, \left| f(n) e^{i\eta_d g n/N} \right| \\
&\leq \int_{\bar{I}} dn\, |P(n)| \\
&\leq \varepsilon
\end{aligned}
\tag{D.10}
$$

Peaked size teleportation occurs in the regime where $gW_\varepsilon/N \ll 1$. In this limit, we can expand

$$e^{i\eta_d g n/N} = e^{i\eta_d g \mathcal{S}/N}\left[1 + E(n)\right] \tag{D.11}$$

where the deviation for $n \in I$ is bounded by

$$|E(n)| \leq \max_{n \in I}\left| 1 - e^{i\eta_d g(n - \mathcal{S})/N} \right| = \left| \sin(\eta_d g W_\varepsilon/N) \right|, \tag{D.12}$$

which holds as long as $gW_\varepsilon/N \leq \pi/2$. We then have

$$
\begin{aligned}
C_Q &= \int_I dn\, f(n) e^{i\eta_d g \mathcal{S}/N}\left[1 + E(n)\right] + R \\
&= e^{i\eta_d g \mathcal{S}/N} G_\beta(Q_A) + R + R' + R''
\end{aligned}
\tag{D.13}
$$

where $G_\beta(Q_A) = \int_0^\infty dn\, f(n) = \mathrm{tr}(Q_A^\dagger \rho^{1/2} Q_A \rho^{1/2})$ is the imaginary time two-point function, and the error $R' = e^{ig\mathcal{S}/N} \int_I dn\, f(n) E(n)$ is bounded by

$$
\begin{aligned}
|R'| &= \left| \int_I dn\, f(n) E(n) \right| \\
&\leq \int_I dn\, |f(n)||E(n)| \\
&\leq \left| \sin(\eta_d g W_\varepsilon/N) \right| \int_I dn\, |f(n)| \\
&\leq \left| \sin(\eta_d g W_\varepsilon/N) \right|
\end{aligned}
\tag{D.14}
$$

and the second error $R'' = G_\beta(Q_A) - \int_I dn\, f(n)$ is bounded by

$$|R''| = \left| G_\beta(Q_A) - \int_I dn\, f(n) \right| = \left| \int_{\bar{I}} dn\, f(n) \right| \le \varepsilon. \tag{D.15}$$

We therefore conclude that whenever $\eta_d g W_\varepsilon / N \le \pi/2$, the deviation of $C_Q$ from the peaked size value is controlled by the upper bound

$$\left| C_Q - e^{i\eta_d g \mathcal{S}/N} G_\beta(Q_A) \right| \le 2\varepsilon + \left| \sin(\eta_d g W_\varepsilon / N) \right| \equiv \mathcal{B}. \tag{D.16}$$

Practically speaking, the lowest value of $g$ for successful peaked-size teleportation is $\eta_d g \mathcal{S}/N = \pi$. Therefore, for a given size distribution, we can guarantee that peaked-size teleportation is possible if we find $\varepsilon$ such that $\mathcal{B} \ll G_\beta(Q_A)$, i.e. the error in the correlator is small compared to the correlator magnitude.

## Application to late times

We illustrate this with some examples, in the few cases where we can exactly solve for operators' full size distribution. First, consider a thermalized system at late times, which we will approximate by setting the size distribution of $Q_A(t)$ to be that of a random Pauli string. For large $n, N$ is a Gaussian distribution with mean $\mathcal{S} = 3N/4$ and variance $\delta\mathcal{S}^2 = 3N/16$:

$$P(n) = (3/4)^n (1/4)^{N-n} \approx \frac{1}{\sqrt{2\pi}\delta\mathcal{S}} \exp\left( -(n - \mathcal{S})^2 / 2\delta\mathcal{S}^2 \right). \tag{D.17}$$

We therefore have

$$1 - \int_{\mathcal{S}-W_\varepsilon}^{\mathcal{S}+W_\varepsilon} dn\, P(n) = 2\,\mathrm{erfc}\left( \frac{W_\varepsilon}{\sqrt{2}\delta\mathcal{S}} \right) = \varepsilon. \tag{D.18}$$

The error function decays exponentially in its argument, so even for exponentially small $\varepsilon$ we require only $W_\varepsilon = A\delta\mathcal{S}$ for some constant $A \sim \mathcal{O}(1)$. Setting $g$ equal to its minimal value, $\eta_d g \mathcal{S}/N = \pi$, we have both $\varepsilon \ll 1$ and $\left| \sin(\eta_d g W_\varepsilon / N) \right| \approx A\delta\mathcal{S}/\mathcal{S} \sim 1/\sqrt{N} \ll 1$, and so peaked size teleportation is guaranteed.

## Application to the large-$q$ SYK model

We can also use this method to guarantee peaked-size teleportation in the large-$q$ SYK model at infinite temperature.

We begin by writing down the size distribution for the large-$q$ SYK model in detail, quoting the results of Ref. [379]. The generating function for the size distribution is:

$$\sum_n P(n) e^{-\mu n} = \frac{e^{-\mu p}}{(1 + (1 - e^{-\mu q})\sinh^2 Jt)^{2p/q}} = \sum_n \frac{\Delta_n}{n!} x^n (1 - x)^\Delta e^{-\mu(qn+p)} \tag{D.19}$$

where we define

$$\Delta_n \equiv \frac{\Gamma(\Delta + n)}{\Gamma(\Delta)}, \quad x \equiv \frac{\sinh^2 Jt}{1 + \sinh^2 Jt}, \quad \Delta \equiv 2p/q. \tag{D.20}$$

From this, we can identify the size distribution:

$$P(qn + p) = \frac{\Delta_n}{n!} x^n (1 - x)^\Delta. \tag{D.21}$$

The size and size width are

$$\mathcal{S} = \bar{n} = \sum_n n \frac{\Delta_n}{n!} x^n (1 - x)^\Delta = \frac{\Delta x}{1 - x}, \qquad \delta \mathcal{S} = \sqrt{\overline{n^2} - \bar{n}^2} = \frac{\sqrt{\Delta x}}{1 - x}. \tag{D.22}$$

Therefore, the ratio of size width to average size is

$$\delta \mathcal{S}/\mathcal{S} = \sqrt{\frac{x}{\Delta} \frac{1}{1 + x}}, \tag{D.23}$$

which approaches zero when $p \to \infty$ ($\Delta \to \infty$).

To apply the upper bound Eq. (D.16), we need to integrate (i.e. sum) the tail of the size distribution in order to compute its asymptotic width [Eq. (D.8)]. In this example, the discrete tail can be summed explicitly and we define

$$I(k) \equiv \sum_{n=k}^\infty P(qk + p) = \sum_{n=k}^\infty \frac{\Delta_n}{n!} x^n (1 - x)^\Delta = \frac{B_x(k, \Delta)}{B(k, \Delta)} \tag{D.24}$$

where $B_x(a, b)$ and $B(a, b)$ are incomplete and ordinary beta function respectively. Let us take $k = \bar{n}(1 \pm \zeta)$ for some small $\zeta$ representing the asymptotic width

$$W_\varepsilon = \bar{n}\zeta q. \tag{D.25}$$

This width corresponds to an error

$$\varepsilon = 1 - I(\bar{n}(1 - \zeta)) + I(\bar{n}(1 + \zeta)). \tag{D.26}$$

Taking $g\mathcal{S}/N = \pi$, the upper bound is

$$\begin{aligned}
\mathcal{B} &= 2[1 - I(\bar{n}(1 - \zeta)) + I(\bar{n}(1 + \zeta))] + \sin \frac{2\pi\zeta x}{1 + x} \\
&= 2 \left( 1 - \frac{B_x(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)} + \frac{B_x(\frac{\Delta x(1+\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1+\zeta)}{1-x}, \Delta)} \right) + \sin \frac{2\pi\zeta x}{1 + x}.
\end{aligned} \tag{D.27}$$

At infinite temperature $G_\beta(Q_A) = 1$, we need to show that the minimum of $\mathcal{B}$ tends to zero when $\Delta \to \infty$.

For early time $\sinh Jt \sim \mathcal{O}(1)$, $1 - x$ is an order 1 number, and we take $\Delta \to \infty$ limit to get

$$\frac{B_x(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)} \to 1, \qquad \frac{B_x(\frac{\Delta x(1+\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1+\zeta)}{1-x}, \Delta)} \to 0 \tag{D.28}$$

The bound becomes

$$\mathcal{B} \to \sin \frac{2\pi\zeta x}{1 + x} \tag{D.29}$$

This basically means that the integrated probability between $\bar{n}(1 - \zeta)$ and $\bar{n}(1 + \zeta)$ for any finite $\zeta$ is 1. One can thus take $\zeta \to 0$ with speed slower than $1/\Delta \to 0$ in order to have the bound vanish. This computation applies for $x \in (0, 1)$, which means that the peaked size always holds for early time. This is physically reasonable as the operator has not yet been scrambled extensively. However, since the size is small at such early times, in order for teleportation to work we must choose $g \sim N$.

For intermediate times, such that $\sinh^2 Jt \sim N$ and $\Delta \ll N \sim 1/(1 - x)$, we must take the $x \to 1$ limit first. Using the fact that

$$\frac{B_x(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)} = 1 - \frac{(1-x)^\Delta x^{\frac{\Delta x(1-\zeta)}{1-x}} \Gamma(\frac{\Delta(1-x\zeta)}{1-x})}{\Gamma(\frac{\Delta x(1-\zeta)}{1-x})\Gamma(1+\Delta)} F(1, \frac{\Delta(1-x\zeta)}{1-x}; \Delta + 1; 1 - x) \tag{D.30}$$

where $F$ is Gauss hypergeometric function, in $x \to 1$ limit the right portion of Eq. (D.30) tends to

$$F(1, \frac{\Delta(1-x\zeta)}{1-x}; \Delta + 1; 1 - x) \to$$
$$F_1(1; \Delta + 1; \Delta(1 - \zeta)) = \Delta^{1-\Delta} e^{\Delta(1-\zeta)}(1 - \zeta)^{-\Delta}(\Gamma(\Delta) - \Gamma(\Delta, \Delta(1 - \zeta))) \tag{D.31}$$

where $\Gamma(x, a)$ is incomplete gamma function. Meanwhile, the left portion of the second term of Eq. (D.30) gives

$$\frac{(1-x)^\Delta x^{\frac{\Delta x(1-\zeta)}{1-x}} \Gamma(\frac{\Delta(1-x\zeta)}{1-x})}{\Gamma(\frac{\Delta x(1-\zeta)}{1-x})\Gamma(1+\Delta)} \to \frac{\Delta^\Delta (1-\zeta)^\Delta e^{-\Delta(1-\zeta)}}{\Gamma(1+\Delta)} \tag{D.32}$$

under $x \to 1$. Combining the two, we have

$$\lim_{x\to1} \frac{B_x(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)}{B(\frac{\Delta x(1-\zeta)}{1-x}, \Delta)} = \frac{\Gamma(\Delta, \Delta(1 - \zeta))}{\Gamma(\Delta)}. \tag{D.33}$$

It follows that the upper bound is

$$\mathcal{B} = 2\left(1 - \frac{\Gamma(\Delta, \Delta(1 - \zeta))}{\Gamma(\Delta)} + \frac{\Gamma(\Delta, \Delta(1 + \zeta))}{\Gamma(\Delta)}\right) + \sin \pi\zeta \tag{D.34}$$

This function has a unique minimum for $\zeta \in [0, 1/2]$ and this minimum decreases as $\Delta$ increases. Taking derivative with respect to $\zeta$, we get

$$\partial_\zeta \mathcal{B} = \pi \cos \pi \zeta - \frac{2\Delta^\Delta}{\Gamma(\Delta)} \left[ (1+\zeta)^{\Delta-1} e^{-\Delta(1+\zeta)} + (1-\zeta)^{\Delta-1} e^{-\Delta(1-\zeta)} \right]$$

$$\rightarrow \pi \cos \pi \zeta - \sqrt{\frac{2\Delta}{\pi}} \left[ (1+\zeta)^{\Delta-1} e^{-\Delta\zeta} + (1-\zeta)^{\Delta-1} e^{\Delta\zeta} \right] \tag{D.35}$$

where in the second step we have taken large $\Delta$ limit. Solving $\partial_\zeta \mathcal{B} = 0$ in this limit, we find the minimum at

$$\zeta \approx \sqrt{\frac{1}{\Delta} \log \frac{8\Delta}{\pi^3}} \rightarrow 0 \tag{D.36}$$

which in turn gives the limit value of $\mathcal{B}$ to be zero. This proves that at infinite temperature, teleportation exactly matches the peaked-size prediction for both early and intermediate times. For late times $t \gg \frac{1}{2J} \log N$ the size distribution above breaks down, as can be seen since $P(n)$ is dominated by some $n > N$, which is unphysical since $N$ is the total number of fermions.

In contrast, we can also show that the above bound does *not* apply at low temperatures for large-$q$ SYK, as expected from Ch. 9. At low temperature, the upper bound $\mathcal{B}$ needs to be much smaller than the two-sided correlation function $G_\beta(Q_A) \sim (\beta J)^{-2\Delta}$ in order to guarantee peaked-size teleportation. The low temperature size distribution is essentially the same as at infinite temperature, requiring only the replacement [379]:

$$x \rightarrow \frac{\sinh^2 \pi t/\beta}{(\pi/\beta J)^2 + \sinh^2 \pi t/\beta} \in [0, 1] \tag{D.37}$$

and adding $e^{-\mu N \delta_\beta}$ to the distribution, which shifts the initial size by a constant amount $N\delta_\beta$ (accounting for the size of the thermal density matrix [379]). Following a similar computation to above, one can show that $\mathcal{B}$ still asymptotes to zero, but now with a *slower* speed than $G_\beta(Q_A)$. For example, in the early time and large $\Delta$ limits, $\mathcal{B} \sim \exp(-\Delta C(x, \zeta))/\sqrt{\Delta}$ where $C(x, \zeta)$ is order 1, while $G_\beta(Q_A) \sim \exp(-2\Delta \log(\beta J))$ is exponentially smaller for large $\beta J$. Therefore, the upper bound $\mathcal{B}$ fails to guarantee peaked-size teleportation. This is consistent with the fact that the correlation function $C_Q(t)$ in Eq. (9.87) in low temperature is far from being a pure phase.

## D.2 The Hayden-Preskill recovery protocol

Here we review the HPR protocol following Refs. [517, 518] and derive its equivalence to the TW protocol in the case of infinite temperature teleportation of a single qubit (introduced in Section 9.6). This single-qubit variant of the HPR protocol was experimentally implemented in Ref. [274], although an explicit derivation of its quantum circuit was not provided.

There are two variants of the HPR protocol: a probabilistic variant, which teleports successfully only with some finite probability, and a deterministic variant, which uses an analog of Grover's search algorithm and succeeds with unit probability, but involves a more complex decoding operation. Both protocols take the general form,

$$\text{(D.38)}$$
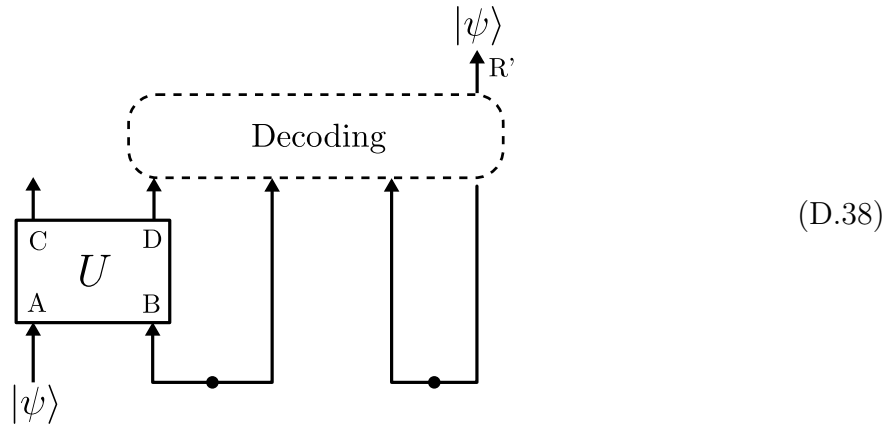
shown for teleportation of a quantum state $|\psi\rangle$ (the generalization to EPR teleportation is straightforward). We now outline the interpretation of each aspect of the above protocol in the context of the Hayden-Preskill thought experiment. For consistency with past literature, we have used *different* subsystem labels than introduced in Ch. 9—most notably, subsystem D now denotes the coupled qubits, and subsystem C denotes its complement. Subsystem B represents an eternal black hole that is maximally entangled with its past Hawking radiation subsystem B', as represented by a dimension $d_B = d'_B$ EPR pair between the two subsystems. Subsystem A contains the initial state $|\psi\rangle$ of an observer Alice's diary. Upon falling into the black hole, the diary's information is scrambled by the unitary time-evolution $U$ acting on the left subsystem $l \equiv AB = CD$. Far from destroying the information of Alice's diary, scrambling by $U$ in fact allows an outside observer Bob to decode the diary if he has access to *any* few qubits of new Hawking radiation D, along with the past Hawking radiation B' and an ancillary EPR pair between A' and R', where $d'_A = d_A$. This decoding relies on OTOCs between subsystem A and D being minimal, a general feature of thermalizing time-evolution after the scrambling time. We describe each of the decoding protocols of Ref. [517] in detail below.

## Probabilistic decoding: intuition

Although our main focus will be on the deterministic teleportation protocol, we review the probabilistic protocol here for completeness, and as a convenient platform to introduce the intuition connecting operator spreading to the success of teleportation. The decoding operation of the probabilistic HPR protocol consists of projection onto EPR pairs on a

subsystems D, D':



$$(D.39)$$
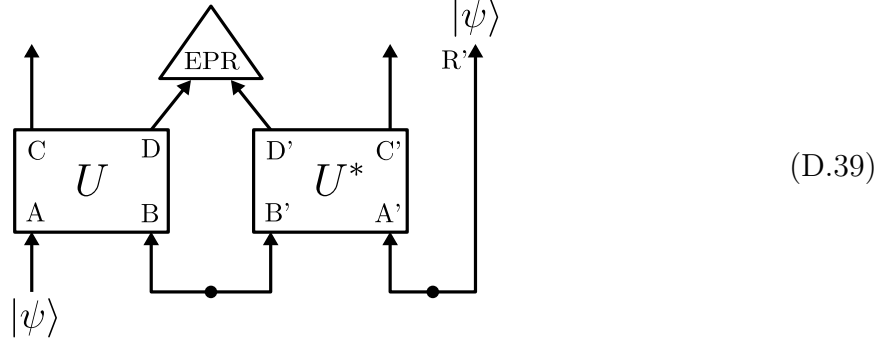
Perfect teleportation requires $d_D \geq d_A$, and succeeds with probability $1/d_A^2$ when $U$ is maximal scrambling. The non-unity success probability signifies that the decoding protocol becomes exponentially more complex with the number of qubits to be teleported.

To provide intuition for the protocol's success, we analyze the action of EPR projection on the initial states $Q_{A,l}(t) |\text{EPR}\rangle$. We restrict to infinite temperature, i.e. EPR pairs in place of the TFD state, in keeping with the original introduction of the HPR protocol in Ref. [517]. We write $Q_A(t)$ as a sum of Pauli strings $S$ on the entire system:

$$Q_A(t) = \sum_R c_R(t) S. \tag{D.40}$$

Denoting the EPR projector on subsystems D, D' as $P_{\text{EPR},D}$ and writing each Pauli string as a tensor product $R = R_C \otimes R_D$ of Paulis on subsystems D and C, we have

$$P_{\text{EPR},D} R_l |\text{EPR}\rangle = \delta_{R_D,\mathbb{1}} R_l |\text{EPR}\rangle, \tag{D.41}$$

since $\langle \text{EPR}_{D,D'}| S_{D,l} |\text{EPR}_{D,D'}\rangle = \text{tr}_D(R_D)/d_D = \delta_{R_D,\mathbb{1}}$. Perfect teleportation is achieved when all input Pauli operators on subsystem A have spread to subsystem D, such that every Pauli string $S$ composing $Q_A(t)$ has non-identity support on subsystem D, for all non-identity $Q_A$. In this situation, the EPR projector has eigenvalue 1 on the thermofield double state and eigenvalue 0 in *all* perturbed states:

$$P_{\text{EPR},D} |\text{EPR}\rangle = |\text{EPR}\rangle \quad, \quad P_{\text{EPR},D} Q_{A,l}(t) |\text{EPR}\rangle = 0. \tag{D.42}$$

However, this is no different than projecting onto EPR pairs between subsystems A and A' before time-evolution by $U_l U_r^*$! This projection would, of course, have an action

$$P_{\text{EPR}} |\text{EPR}\rangle = |\text{EPR}\rangle \quad, \quad P_{\text{EPR}} Q_{A,l} |\text{EPR}\rangle = \text{tr}(Q_A) = 0. \tag{D.43}$$

Expressed diagrammatically, this equivalence is:



$$(D.44)$$

for all initial states $\psi$. However, performing EPR projection between subsystems A, A' before time-evolution is precisely the standard quantum teleportation protocol, applied to subsystems A, A', and R'. The scrambling dynamics of $U$ allow one to perform this teleportation via coupling *any* subsystem D of the system's qubits.

### Deterministic decoding

After scrambling, the probability of successful EPR projection on subsystem D, $\mathcal{O}(1/d_A^2)$, is exponentially small in the size of subsystem A, the state to be teleported. In contrast to standard teleportation, non-successful EPR projection (i.e. projection onto a different maximally entangled state, not $|\text{EPR}_{D,D'}\rangle$) *cannot* be corrected via an additional decoding operation. This exponential decrease in success probability is overcome in the deterministic HPR protocol, which uses an analog of Grover's search algorithm to search for an EPR pair between subsystems D, D'. The protocol requires $\mathcal{O}(d_A)$ steps for completion, again exponential in the number of qubits to be teleported (albeit with half the exponent of the probabilistic decoding).

Grover's search algorithm involves two operations: the first applies a minus sign to the state one is searching for, and the second applies a minus sign to the system's initial state. We will search for an EPR pair on subsystem $D$, so for the first step we apply $W_D \equiv 1 - 2P_{\text{EPR},D} = e^{i\pi P_{\text{EPR},D}}$:

$$(D.45)$$

In the second step, we flip the sign of the initial state (the time-evolved EPR pair between

A' and the reference qubit R') by applying $\widetilde{W}_A \equiv U^* W_A U^T$:

$$
\begin{array}{c}
\text{D'} \quad \text{C'} \quad \text{R'} \\
\boxed{\widetilde{W}_A}
\end{array}
=
\begin{array}{c}
\text{D'} \quad \text{C'} \quad \text{R'} \\
\boxed{U^*} \\
\boxed{W_A} \\
\boxed{U^T}
\end{array}
\tag{D.46}
$$

where $W_A = 1 - 2P_{\text{EPR},A}$ acts on A', R' to apply a minus sign if the two are in an EPR pair.

The entire Grover protocol is identical to the probabilistic protocol, but with EPR measurement replaced by repeated applications of the two above steps until the EPR pair is found. Displaying, for instance, only the first two iterations:

$$
\tag{D.47}
$$



After $\mathcal{O}(d_A)$ iterations, the state $|\psi\rangle$ is found on subsystem R'.

**Single qubit deterministic decoding**
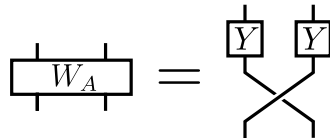
Two important simplifications occur to the deterministic HPR protocol in the case of single qubit teleportation, $d_A = 2$. The first is that the Grover operator $W_A$ is equal to a SWAP operator composed with single-qubit $Y$ operations. To see this, we expand $W_A$ in terms of Pauli operators:

$$
\begin{aligned}
W_A &= 1 - 2P_{\text{EPR},A} \\
&= 1 - \frac{2}{d_A^2} \sum_{P_A} P_{A,l} P_{A,r}^* \\
&= \frac{1}{2} - \frac{1}{2}X_l X_r + \frac{1}{2}Y_l Y_r - \frac{1}{2}Z_l Z_r \\
&= \frac{1}{2}Y_l \left[1 + X_l X_r + Y_l Y_r + Z_l Z_r\right] Y_l \\
&= Y_l \left(\text{SWAP}\right) Y_l, \\
&= Y_l Y_r \left(\text{SWAP}\right)
\end{aligned}
\tag{D.48}
$$

where in the final equality we used $Y_r \text{SWAP} = \text{SWAP}Y_l$, and in the second equality we used the Pauli decomposition for the swap operator between two $d_A$-dimensional boson systems:

$$
\text{SWAP} = \frac{1}{d_A} \sum_{P_A} P_{A,l} P_{A,r}^\dagger.
\tag{D.49}
$$

Expressed graphically, we have



$$\tag{D.50}$$

The second simplification is that Grover's search for an EPR pair D, D' succeeds after only one step; this is a general result for Grover's search in a $d_D^2 = 4$-dimensional database [353].

It implies that the Grover protocol can teleport one qubit through the circuit:

$$
\begin{array}{c}
\text{(D.51)}
\end{array}
$$
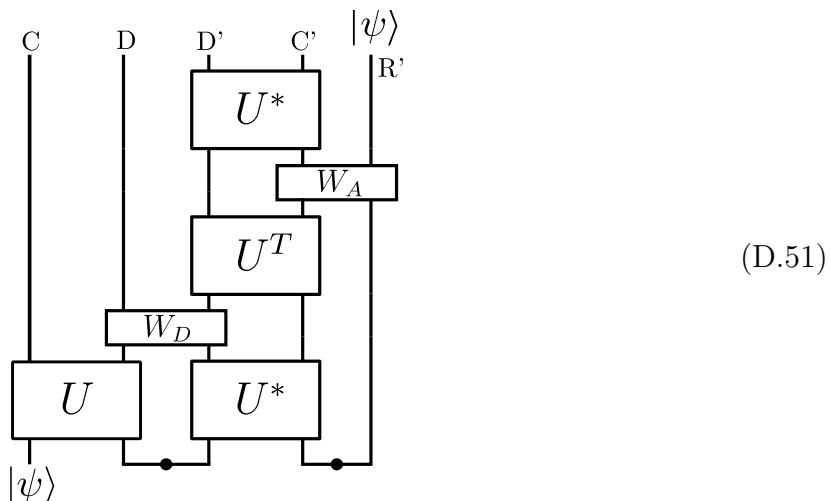
If we only care about the fidelity of the teleported state, we can neglect the final application of $U^*$. Performing the SWAP gate explicitly, and neglecting the action of the final $Y$ operator on R', we have:

$$
\begin{array}{c}
\text{(D.52)}
\end{array}
$$
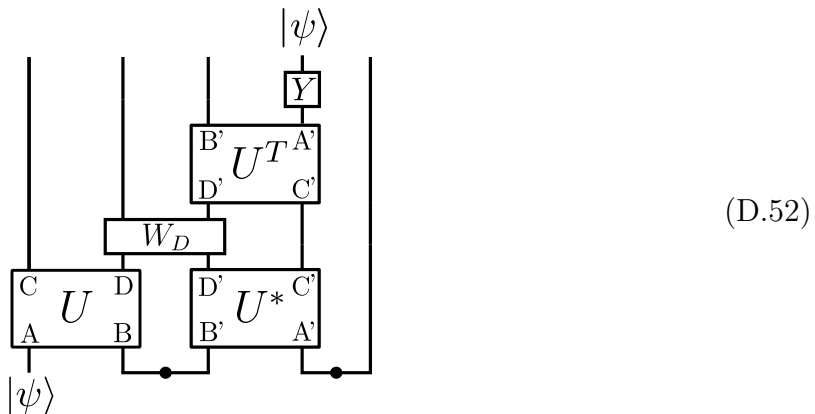
This exact circuit has been performed in trapped ion experiment [274].  We now make a

small cosmetic adjustment, and move the reference qubit R' from the far right to the far left,

$$
\begin{array}{c}
\text{(D.53)}
\end{array}
$$
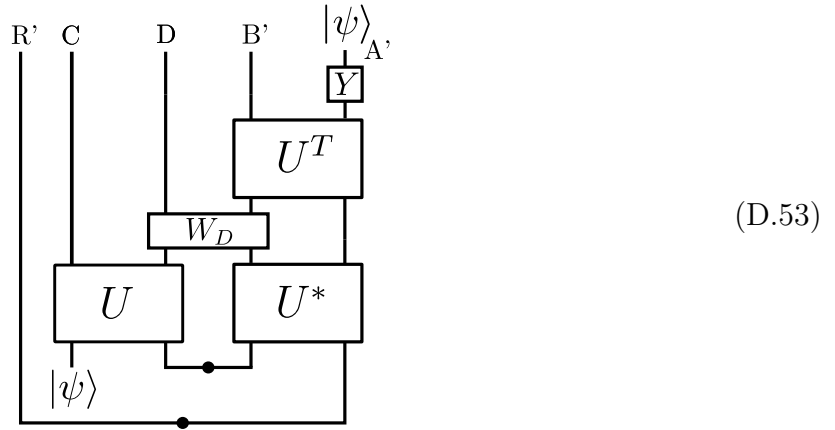
Sliding $U^*$ to the left side using Eq. (9.9), we have:

$$
\begin{array}{c}
\text{(D.54)}
\end{array}
$$
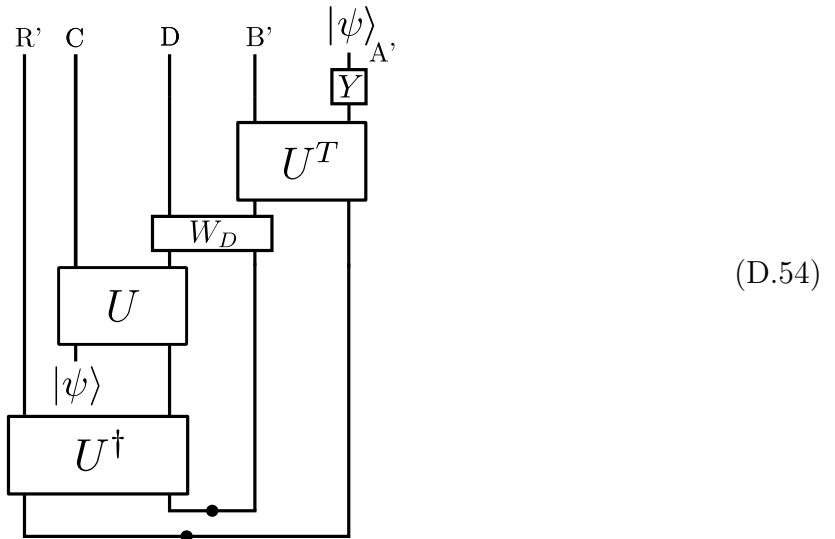
This is the same circuit appearing the teleportation protocol of Ref. [71, 350], modulo the

precise form of the coupling. In the case of EPR teleportation, we would instead have



$$\text{(D.55)}$$
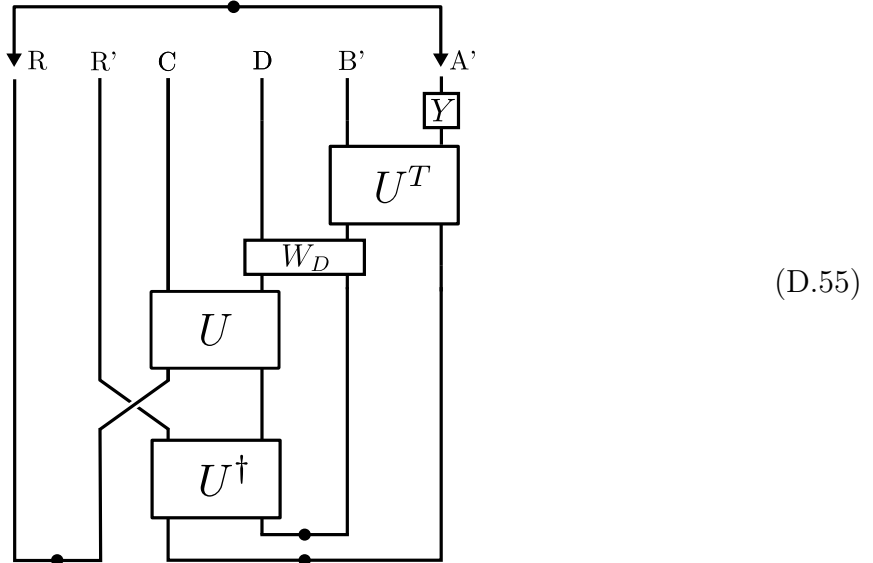
where subsystems R' and A' are in an EPR pair when teleportation is successful. This is the circuit appearing in Ref. [159], modulo the form of the coupling as well as the $Y$ decoding operation. The lack of a $Y$ decoding operation for fermionic teleportation is discussed in Section D.7.

## D.3   State teleportation fidelity

In Section 9.5, we provided a detailed derivation of the teleportation fidelity's relation to the teleportation correlators in the case where one teleports one half of an EPR pair. This allowed us to lower bound the fidelity in Eq. (9.43) and calculate the peaked-size fidelity in Eq. (9.45). Here we do the same for teleportation of a quantum state, as shown in Fig. 12.1(a) and outlined in Section 9.2. Our results provide a rigorous foundation for the arguments of Section 9.3, in particular the insertion of the state $\langle\phi|$ and the subsequent replacement of $|\psi\rangle\langle\phi|$ with a Pauli operator $Q_A$.

We begin with the insertion of $\langle\phi|$ into the protocol Eq. (9.17). We do so by inserting the resolution of the identity $\frac{1}{d_A}\sum_{|\phi\rangle}|\phi\rangle\langle\phi| = \mathbb{1}$ into the ancillary qubit leg of the diagram

for the state teleportation fidelity. We find:

$$F_\psi = \quad = \quad \frac{1}{d_A} \sum_{|\phi\rangle} \qquad (D.56)$$

Plugging Eq. (9.19) into this diagram provides unit teleportation fidelity, as described in Ch. 9. When teleportation is successful each of the $d_A$ terms of the sum must succeed individually, so the right input state $|\phi\rangle$ will not affect the success of the teleportation.

As with EPR distillation [Eq. (9.41)], we can relate the state teleportation fidelity to

correlators of Pauli operators by decomposing the SWAP operator. Diagramatically,



$$F_\psi = \hspace{5cm} = \frac{1}{d_A^2} \sum_{P_1, P_2} \hspace{3cm}, \tag{D.57}$$

and in equation form,

$$F_\psi = \frac{1}{d_A^2} \sum_{P_1, P_2} \langle \psi | P_2 P_1^\dagger | \psi \rangle \cdot \langle \text{TFD} | P_{2,l}^\dagger(t) e^{-igV} | \psi \rangle \langle \psi |_r (-t) \, e^{igV} P_{1,l}(t) | \text{TFD} \rangle . \tag{D.58}$$

When the correlators are maximal with phases $e^{i\theta_P}$, i.e. when

$$e^{igV} P_{1,l}(t) | \text{TFD} \rangle = e^{i\theta_P} P_{1,r}(-t) | \text{TFD} \rangle , \tag{D.59}$$

we can simplify this expression as

$$
\begin{aligned}
F_\psi &\approx \frac{1}{d_A^2} \sum_{P_1, P_2} \langle \psi | P_2 P_1^\dagger | \psi \rangle \cdot \langle \text{TFD} | P_{2,r}^\dagger(-t) | \psi \rangle \langle \psi |_r (-t) \, P_{1,r}(-t) | \text{TFD} \rangle \\
&= \frac{1}{d_A^2} \sum_{P_1, P_2} e^{i(\theta_{P_1} - \theta_{P_2})} \cdot \langle \psi | P_2 P_1^\dagger | \psi \rangle \cdot \text{tr}\left( \rho \, P_2^\dagger | \psi \rangle \langle \psi | P_1 \right) \\
&= \frac{1}{d_A^2} \sum_{P_1, P_2} e^{i(\theta_{P_1} - \theta_{P_2})} \cdot \text{tr}\left( P_1^\dagger | \psi \rangle \langle \psi | P_2 \right) \cdot \text{tr}\left( \rho \, P_2^\dagger | \psi \rangle \langle \psi | P_1 \right) .
\end{aligned}
\tag{D.60}
$$

As expected, when the phases $e^{i\theta_P}$ are the same for all operators, this gives unit fidelity:

$$
\begin{aligned}
F_\psi &= \frac{1}{d_A^2} \sum_{P_1, P_2} \mathrm{tr}(P_1^\dagger \,|\psi\rangle\langle\psi|\, P_2) \cdot \mathrm{tr}(\rho\, P_2^\dagger \,|\psi\rangle\langle\psi|\, P_1) \\
&= \frac{1}{d_A} \sum_{P_1} \mathrm{tr}(P_1^\dagger \,|\psi\rangle\langle\psi|\, |\psi\rangle\langle\psi|\, P_1 \rho) \\
&= \mathrm{tr}(|\psi\rangle\langle\psi|)\mathrm{tr}(\rho) \\
&= 1,
\end{aligned}
\tag{D.61}
$$

using properties of Pauli operators as a 1-design [540]. Differing phases $e^{i\theta_P}$ cause the terms in the sum to interfere with each other, giving lower fidelity. At finite temperature, the fidelity of peaked-size teleportation is again limited. For instance, if $|\psi\rangle$ is a single-qubit eigenstate of the Pauli $Z$ operator, we have:

$$
\begin{aligned}
F_{\mathrm{EPR}} &= \frac{1}{2^2} \sum_{P_1, P_2} \langle\psi|\, P_2 P_1^\dagger \,|\psi\rangle \cdot \langle\mathrm{TFD}|\, P_{2,l}^\dagger(t) e^{-igV} [Y\, |\psi\rangle\langle\psi|Y]_r(-t)\; e^{igV} P_{1,l}(t)\, |\mathrm{TFD}\rangle \\
&= \frac{1}{2^2} \sum_{P_1, P_2} \langle\psi|\, P_2 P_1^\dagger \,|\psi\rangle \cdot \mathrm{tr}(|\psi\rangle\langle\psi|\; \rho^{1/2} P_2^\dagger P_1 \rho^{1/2}) \\
&= \sum_{P} \langle\psi|\, P^\dagger \,|\psi\rangle \cdot \mathrm{tr}(|\psi\rangle\langle\psi|\; \rho^{1/2} P \rho^{1/2}) \\
&= 2\mathrm{tr}(|\psi\rangle\langle\psi|\; \rho^{1/2} \,|\psi\rangle\langle\psi|\, \rho^{1/2}) \\
&\approx \frac{1}{2}\mathrm{tr}((\mathbb{1} + Z)\, \rho^{1/2}(\mathbb{1} + Z)\rho^{1/2}) \\
&\approx \frac{1}{2} + \frac{1}{2}G(t' - t + i\beta/2) + \langle Z\rangle_\beta,
\end{aligned}
\tag{D.62}
$$

where $\langle Z\rangle_\beta = \mathrm{tr}(Z\rho)$, which averages to zero for different initial states $|\psi\rangle$.

## D.4 Rydberg numerical simulations

For the numerical results shown in Fig. 9.7 and 9.8, we simulate the full TW protocol with time evolution generated by the analog Rydberg Hamiltonian [Eq. (9.105)]. In particular, we implement the *one-sided* version of state teleportation, which is obtained by replacing the EPR measurement in Fig. 9.6(b) with a measurement of a two-qubit state $|\psi\rangle \otimes |\psi^*\rangle$. The many-body unitary corresponds to $U = e^{-iHt}$, where $H$ is given in Eq. (9.105) with $\Omega_i = .9$, $\Delta_i = -1.5$, $J_0 = 1$ and open boundary conditions. The teleported state $|\psi\rangle$ is inserted in the middle qubit, and the remaining $K = N - 1$ qubits serve as 'measured' qubits, with $\hat{O}_i = \hat{Z}_i$ (see Section 9.9).

More explicitly, the numerical procedure is given as follows: (*i*) begin in a random initial state, $|o_1 \cdots o_N\rangle$; (*ii*) evolve forward for time $t$ under the Rydberg Hamiltonian; (*iii*) apply

the operator $|\phi\rangle\langle\psi|$ onto the middle qubit; $(iv)$ evolve backward in time, apply $\hat{V}_i = e^{igo_i\hat{Z}_i/K}$ to each of the $K = N - 1$ 'measured' qubits (where $o_i$ is determined by the initial state), and evolve forward again; $(v)$ measure the projector $|\psi\rangle\langle\psi|$ on the middle qubit. We repeat this procedure for $|\phi\rangle \in \{|0\rangle, |1\rangle\}$ [see Eq. (D.57)] and starting from $\sim 100$ random initial states. Moreover, to compute the *average* state fidelity, we average $|\psi\rangle$ over all single-qubit states in a 2-design [274], i.e. $|\psi\rangle \in \{|0\rangle, |1\rangle, \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)\}$). Lastly, we note that the time evolution is implemented with Krylov subspace methods, an iterative technique that is amenable to parallelization and is more computationally efficient than exact diagonalization [1, 257].

# D.5 Random unitary circuit numerics

In this section, we provide additional details and numerical data from our random unitary circuit simulations (Section 9.7 B, C).

## Algorithm

Our goal for the RUC simulations is to compute the Haar-averaged EPR fidelity and operator size distribution for the circuit layouts shown in Fig. 9.2. Crucially, the relevant diagrams for computing these quantities—Eq. (9.41) for the EPR fidelity, and Eq. (9.32) for the operator size distribution—contain at most three copies of $U$ and $U^\dagger$. Together with the fact that Clifford unitaries form a 3-designs for qubits, this implies that can compute the averaged quantities by replacing each Haar-random gate with a random *Clifford* gate [271, 484, 534]. This dramatic simplification has been exploited in prior studies of operator growth in random unitary circuits [340, 476]; here, we adapt these same techniques for computing the full size distribution and the teleportation fidelity.

In more detail, our algorithm consists of the following three ingredients. First, following a standard approach [3, 340], we represent an initial $n$-qubit Pauli operator, $Q$, as a binary string $v = x_1 x_2 \cdots x_n z_1 z_2 \cdots z_n$ of length $2n$:

$$Q = \prod_{i=1}^{n} Q_i(x_i, z_i) \tag{D.63}$$

where $Q_i(0,0) = I_i$, $Q_i(1,0) = X_i$, $Q_i(0,1) = Z_i$, and $Q_i(1,1) = Y_i$ denote individual Pauli operators within the Pauli string. For example, the operator $\mathbb{1} \otimes \mathbb{1} \otimes \mathbb{Z} \otimes \mathbb{1} \otimes \mathbb{1}$ is represented as $x = 00000$ and $z = 00100$. Normally, one would also keep track of the overall phase of $Q$, but for our purposes the phase will be irrelevant and is dropped in the above notation.

Second, we evolve $Q$ under a random Clifford unitary $U$ to obtain $Q(t) = UQU^\dagger$. We consider the circuits shown in Fig. 9.2, which are composed of random 2-qubit Clifford unitaries laid out in a "brick-layer" fashion. Each of the 2-qubit unitaries is sampled uniformly from the set of 2-qubit Clifford unitaries. While an algorithm exists to perform this sampling

directly [291], in practice we find it more convenient to pre-compute and enumerate the entire 2-qubit Clifford set (which consists of 11520 distinct unitaries)[1]. In the binary notation, each 2-qubit Clifford unitary corresponds to a map which acts on the relevant components $v$, i.e. a unitary with support on the $j$th and $k$th qubits updates the values of $(x_j, z_j, x_k, z_k)$. The time complexity of applying the full circuit thus scales linearly with the number of 2-qubit gates and does not otherwise depend on the number of qubits $n$. As a reference point, simulating a 0D circuit until the scrambling time with $10^8$ qubits for a single realization takes approximately one day on a standard single-core processor.

Third, we compute the average operator size distribution and EPR fidelity of the time-evolved operators. For the former, we simply count the size, i.e. number of non-identity terms, of a time-evolved operator $Q(t)$ for a single circuit realization and determine the distribution of sizes with respect to $\sim 10^3$ circuit realizations. Depending on the simulation, we either initialize $Q$ with support on a single site (i.e. $p=1$) or as a $p$-body operator. In either case, the specific terms in $Q$ (e.g. whether each site is initialized as $X$, $Y$, or $Z$) is arbitrary since the averaged quantities are basis independent.

Computing the averaged EPR fidelity requires an additional average over the initial operators. In particular, for a single circuit realization $U$, we compute the EPR fidelity using [Eq. (9.45)]:

$$F_{\text{EPR}} = \left| \frac{1}{d_A^2} \sum_{Q_A} e^{i\theta_{Q_A}} \right|^2 \tag{D.64}$$

where

$$\theta_{Q_A} = \eta_d g \mathcal{S}_K [U Q_A U^\dagger]/K + \pi \mathcal{S}[Q_A]. \tag{D.65}$$

and $\eta_d \equiv 1/(1-1/d^2)$, as defined in Section 9.4. Note that the first term in $\theta_{Q_A}$ corresponds to the phase applied by the coupling, while the second term accounts for minus signs associated with transposition and decoding (see Section 9.7). The sum in Eq. (D.64) is over the complete basis of Pauli operators in subsystem A. For single-qubit teleportation, this consists of three non-trivial Pauli operators and the identity (for which $\theta = 0$), and the sum can be performed explicitly. However, for teleporting many qubits, the number of terms quickly becomes intractable, and we instead approximate the sum by sampling $Q_A$ (e.g. $\sim 100$ randomly selected operators). To compute the average EPR fidelity, we repeat this procedure for $\sim 100$ realizations of $U$. Finally, we note that the coupling strength $g$ enters the fidelity calculation in a computationally efficient manner; in particular, upon determining the distribution of sizes for a particular circuit realization, we can compute the teleportation fidelity for arbitrary values of $g$ "offline" with no additional simulation cost.

## Extended data for 1D and 2D RUCs

*Size distribution*— The average size and size width for time-evolved operators in 1D and 2D for various system sizes are shown in Fig. D.1. In each case, we apply periodic boundary

---

[1]We are grateful to Maxwell Block for sharing code to generate the full set of 2-qubit Clifford unitaries.
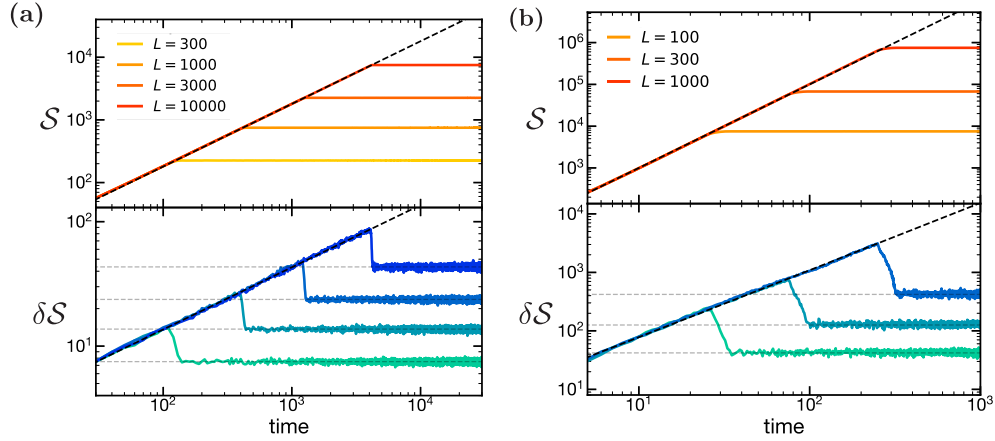
Figure D.1:   Extended data for average operator size and size width in 1D **(a)** and 2D **(b)** RUCs. The average size grows ballistically $\sim t^d$ (dashed line) and saturates at $t_{\mathrm{scr}} \sim L \sim N^{1/d}$. The size width matches the predictions from the KPZ universality class (dashed lines) and allows us to extract the prefactors in Eq. (9.57) and (9.58). In particular, we determine $\alpha_{\mathrm{bulk}}$ and $\beta_{\mathrm{bulk}}$ from the saturation values (light gray), and $\alpha_{\mathrm{boundary}}$ and $\beta_{\mathrm{boundary}}$ from the initial growth rate (dark gray).

conditions and begin with a single-qubit operator. These results match the functional forms predicted by the KPZ universality class [Eq. (9.57) and (9.58)] and allow us to extract the growth parameters $\{\alpha_{\mathrm{bulk}}, \alpha_{\mathrm{boundary}}, \beta_{\mathrm{bulk}}, \beta_{\mathrm{boundary}}\} = \{0.66, 0.70, 1.2, 4.5\}$.

*Multiple qubits*—In Fig. D.2, we present numerical evidence to support our claim that multiple qubits can be teleported in $\geq$ 1D short-range models if their operator light cones are non-overlapping (Section 9.7). In particular, we simulate the teleportation of $n = 5$ qubits that are initially evenly spaced in a 1D RUC with periodic boundary conditions. At early times ($t < 1300$, Region I), we confirm that high-fidelity teleportation is possible for a wide range of coupling strengths, and by measuring the average operator size we infer that during this time the operator light cones have not overlapped. In contrast, after the light cones have overlapped, we generally observe a large suppression in the teleportation fidelity.

Interestingly, there is one noticeable exception to this reasoning: When only adjacent light cones have overlapped (i.e. $1300 < t < 2600$, Region II), high-fidelity teleportation can still occur for specific values of $g$. This situation corresponds to when the multi-qubit size is a multiple of $2\pi K/\eta_d g$ off from the size addition value, e.g. $\mathcal{S}[Q_1(t)Q_2(t)] = \mathcal{S}[Q_1(t)] + \mathcal{S}[Q_1(t)] - 2\pi m(K/\eta_d g)$, where $m$ is an integer value. Therefore, strictly speaking, it is possible to satisfy the conditions for many-body teleportation *without* size addition; nevertheless, it is a non-generic effect that requires finely tuned values of $g$ and evenly spaced input qubits.
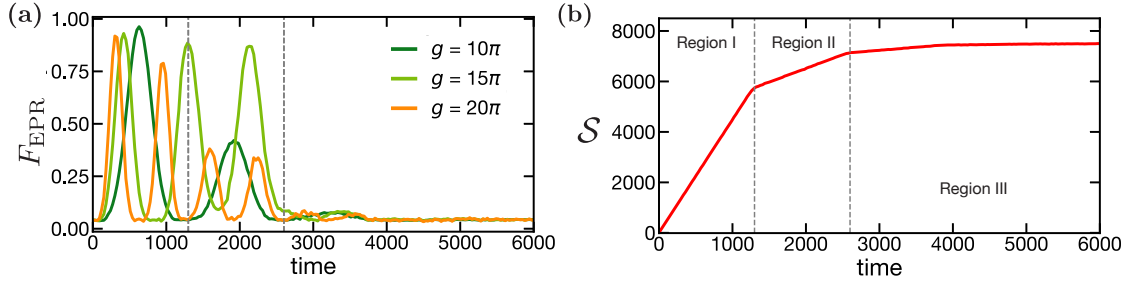
Figure D.2:   Teleporting multiple qubits ($n = 5$) in 1D, where the input qubits are evenly spaced in the system ($N = 10000$). **(a)** Teleportation is achieved with high fidelity for $t \leq 1300$ (Region I). This corresponds to the regime in which the light cones of the operators are non-overlapping. Interestingly, order-one fidelity can also occur for $1300 < t < 2600$ (Region II), when adjacent light cones have overlapped, but only for certain values of $g$. No multi qubit teleportation is possible for $t \geq 2600$ (Region III), as expected from the lack of size addition. **(b)** The three Regions can be detected by changes in the slope of the operator size as a function of time. In particular, the growth rate decreases when nearest neighbor light cones, then next nearest neighbor light cones, etc. begin to overlap.

## Channel capacity for 0D RUCs

An important result of our numerical simulations is substantiating the claim that 0D RUCs exhibit a channel capacity that scales linearly with the number of coupled qubits $K$. To this end, we first recall that our working definition for the channel capacity is based on setting a threshold for the *per qubit fidelity*. The most direct way to compute this fidelity would be to take the $n$-th root of the many-body EPR fidelity; in practice, however, this approach is numerically unstable for large $n$. Thus, we instead consider a modified protocol for estimating the per qubit fidelity where one attempts to send $n$ qubits but only measures the fidelity of one of the $n$ qubits. At infinite temperature and generalizing from one to $m$ qubits, this fidelity is given by:

$$
\begin{aligned}
F_{\text{EPR}}^{(m)} &= \frac{1}{d_A^4} \sum_{Q_1, Q_2} \langle \text{TFD}| \, Q_{2,l}^\dagger(t) \, e^{-igV} \, \tilde{Q}_{2,r}^m(-t) \tilde{Q}_{1,r}^{m\dagger}(-t) \, e^{igV} \, Q_{1,l}(t) \, |\text{TFD}\rangle \, \cdot \text{tr}(Q_1^{u\dagger} Q_2^u) \\
&= \frac{1}{d_m^4 d_u^2} \sum_{Q_1, Q_2} e^{i(\theta_{Q_1} - \theta_{Q_2})} \delta_{Q_1^u, Q_2^u}
\end{aligned} \tag{D.66}
$$
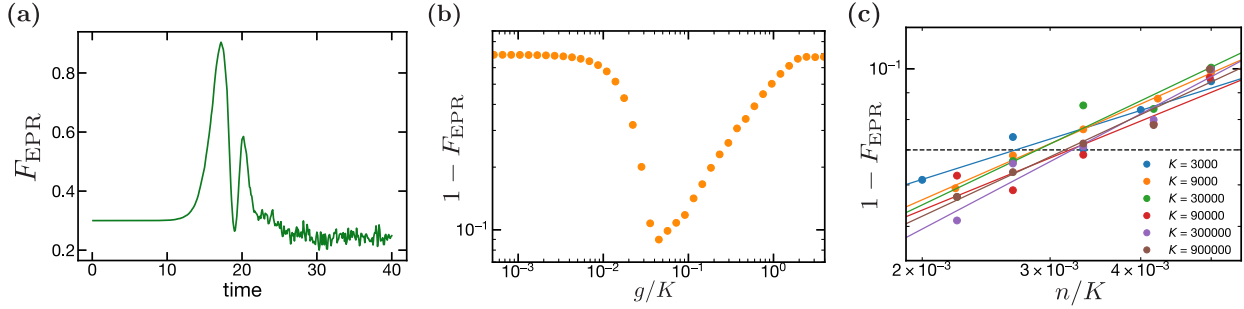
Figure D.3:   Procedure for determining the channel capacity in 0D RUCs. **(a-b)** For fixed $n$ and $K$, we compute the per qubit fidelity while sweeping both the evolution time and coupling strength $g$. (a) The fidelity as a function of evolution time with coupling strength fixed is optimized at the first local maximum, which corresponds to $\eta_d g \mathcal{S}/N = \pi$. (b) After optimizing the evolution time, the fidelity as a function of the coupling strength $g$ is maximal when $g$ (and correspondingly the average operator size $\mathcal{S}$) is tuned to balance errors due to size addition and the finite number of couplings (see Section 9.7 for details). The data shown correspond to $n = 38$ and $K = 9000$. **(c)** The channel capacity is defined as the maximum number of qubits that can be teleported while maintaining the fidelity per qubit above a fixed threshold, i.e. $1 - F_{\text{EPR}}^{(1)} \leq 0.07$ (dashed line). To determine this number, we fit the optimal fidelity as a function of $n$ (for each $K$) with a linear fit in log space and compute the intercept of the fit with the threshold fidelity. The fits approximately collapse with respect to $n/K$, indicating that the channel capacity is linear in $K$.

where $Q = Q^m \otimes Q^u$ and $d_A = d_m d_u$, such that $Q^m$ acts on the measured qubit(s), and $Q^u$ acts on the unmeasured qubits. This can be derived diagrammatically via



$$(\text{D.67})$$

Hence, computing the per qubit fidelity, $F_{\text{EPR}}^{(1)}$, is nearly identical to computing the full many-body fidelity, except we sample only over pairs of Pauli operators $(Q_1, Q_2)$ which are identical on every qubit except for one.

We next discuss how to determine the channel capacity from the teleportation fidelity. Specifically, we compute the maximum number of qubits $n_{\text{max}}$ that can be teleported above a certain teleportation fidelity, where we fix the number of coupled qubits $K$ and optimize over the evolution time $t$ and the coupling strength $g$. We consider each of these parameters in turn. First, when sweeping the evolution time and holding all other parameters fixed, the maximum fidelity occurs during the first peak in the time profile; this corresponds to a size $\eta_d g \mathcal{S} = \pi/N$. After optimizing the evolution time (but holding $n$ and $K$ fixed), the fidelity is non-monotonic with respect to $g$, owing to the competition among errors due the size addition and finite $K$. Finally, after optimizing evolution time and $g$, we determine the maximum number of qubits that can be teleported while maintaining a per qubit fidelity above a fixed threshold value, i.e. $1 - F_{\text{EPR}}^1 \geq 0.07$. Our results from this procedure are shown in Fig. D.3 and demonstrate that the channel capacity follows a linear trend in $K$ across two orders of magnitude, in agreement with our analytical predictions.

## D.6 Random circuit calculations

Here we provide more detailed technical calculations of the size overlap and $K$-size distribution of random Pauli operators of a fixed size. The former is relevant to 0D RUCs (Section 9.7), as the vanishingly small overlap of random Pauli strings with size much less than the system size underlies the circuit's ability to teleport multiple qubits at intermediate times. The latter is applicable to all systems when the $K$ coupled qubits are chosen randomly, and quantifies the width introduced to the $K$-size by this random sampling (Section 9.4). In the appropriate limits, these calculations reproduce the intuitive binomial scalings we argued for in Ch. 9.

### Distribution of the overlap of two random Pauli strings

We are interested in the probability distribution of the size of the overlap, $p$ (not to be confused with the large-$p$ encoding, which we do not reference in this Section) of two randomly samples Pauli strings of fixed size $R_1, R_2$, in a system of $N$ qubits. We expect this to quantify errors to the size addition formula, Eq. (9.55) in Section 9.7, for 0D RUCs with large-$p$ encoding (Section 9.7), where the assumption of random Pauli strings of a fixed size is appropriate. Our precise derivation is necessarily quite technical, however our final result matches that obtained by intuitive arguments in Section 9.7.

This probability distribution is computed exactly as a product of various factorials:

$$P[p] = \frac{C_p^N C_{R_1-p}^{N-p} C_{R_2-p}^{N-R_1}}{C_{R_1}^N C_{R_2}^N} = \frac{1}{p!} \frac{R_1!}{(R_1-p)!} \frac{R_2!}{(R_2-p)!} \frac{(N-R_1)!(N-R_2)!}{N!(N-R_1-R_2+p)!} \qquad \text{(D.68)}$$

The numerator computes the number of distinct configurations with Pauli strings of size $R_1$, $R_2$ and overlap $p$, while the denominator computes the number of distinct Pauli strings of size $R_1$, $R_2$ regardless of the overlap. We are interested in the case where all variables are extensive (scale with $N$), but $N \gg R_1, R_2 \gg p$. We will proceed by applying Stirling's approximation to each term above, which holds as long as all quantities are large compared to 1. For instance, for dummy variables $n, k$, we have:

$$\frac{n!}{(n-k)!} \approx \sqrt{\frac{n}{n-k}} \frac{n^n}{(n-k)^{n-k}} e^{-k} = n^k \left(1 - \frac{k}{n}\right)^{-n+k-1/2} e^{-k} \qquad (\text{D.69})$$

or, taking the logarithm,

$$\log \frac{n!}{(n-k)!} \approx k \log(n) - \left(n - k + \frac{1}{2}\right) \log\left(1 - \frac{k}{n}\right) - k. \qquad (\text{D.70})$$

We will apply this to a few pairs of factorials in our original expression for $P[p]$. For convenience, we only keep track of the $p$-dependence of the probability, and neglect overall constants which serve to normalize the distribution. Anticipating that the average $p$ will be $R_1 R_2/N$, we expand $p = R_1 R_2/N + \delta$ and work to second order in $\delta$. At the end we will show that this is justified. We have:

$$\log \frac{R_1!}{(R_1 - p)!} \approx p \log(R_1) - \left(R_1 - \frac{R_1 R_2}{N} + \frac{1}{2}\right) \log\left(1 - \frac{R_2}{N} - \frac{\delta}{R_1}\right) - \frac{R_1 R_2}{N} - \delta \quad (\text{D.71})$$

Expanding the logarithm using

$$\log(1 - y - x) \approx \log(1 - y) - \frac{x}{1-y} - \frac{1}{2}\frac{x^2}{(1-y)^2} + \mathcal{O}(x^3) \qquad (\text{D.72})$$

we have

$$\begin{aligned}
\log \frac{R_1!}{(R_1 - p)!} \approx{}& p \log(R_1) - \\
& \left(R_1 - \frac{R_1 R_2}{N} - \delta + \frac{1}{2}\right) \left[\log\left(1 - \frac{R_2}{N}\right) - \frac{\delta/R_1}{1 - R_2/N} - \frac{(\delta/R_1)^2}{(1 - R_2/N)^2}\right] \\
& - \delta + \ldots \\
\approx{}& p \log(R_1) + \delta \log\left(1 - \frac{R_2}{N}\right) - \frac{1}{2}\delta^2 \left[\frac{1}{R_1}\frac{1}{1 - R_2/N}\right] \\
& + \mathcal{O}(\delta/R) + \mathcal{O}(\delta^3/R^2) + \ldots.
\end{aligned}$$
$$(\text{D.73})$$

This gives

$$\begin{aligned}
\log \frac{R_1!}{(R_1 - p)!} \frac{R_2!}{(R_2 - p)!} \approx{}& p \log(R_1 R_2) + \delta \log\left((1 - \frac{R_2}{N})(1 - \frac{R_1}{N})\right) \\
& - \frac{1}{2}\delta^2 \left[\frac{1}{R_1}\frac{1}{1 - R_2/N} + \frac{1}{R_2}\frac{1}{1 - R_1/N}\right] \\
& + \mathcal{O}(\delta/R) + \mathcal{O}(\delta^3/R^2) + \ldots.
\end{aligned}$$
$$(\text{D.74})$$

The last piece is

$$
\begin{aligned}
\log &\frac{N!}{(N - R_1 - R_2 + p)!} \\
&\approx -p \log(N) \\
&\quad - \left( N - R_1 - R_2 + \frac{R_1 R_2}{N} + \delta + \frac{1}{2} \right) \log\left( 1 - \frac{R_1}{N} - \frac{R_2}{N} + \frac{R_1 R_2}{N^2} + \frac{\delta}{N} \right) \\
&\quad + \delta + \ldots \\
&\approx -p \log(N) - \left( N - R_1 - R_2 + \frac{R_1 R_2}{N} + \delta + \frac{1}{2} \right) \times \\
&\quad \left[ \log\left( (1 - \frac{R_1}{N})(1 - \frac{R_2}{N}) \right) + \frac{\delta/N}{(1 - \frac{R_1}{N})(1 - \frac{R_2}{N})} - \frac{\delta^2/N^2}{(1 - \frac{R_1}{N})^2 (1 - \frac{R_2}{N})^2} \right] \\
&\quad + \delta + \ldots \\
&\approx -p \log(N) - \delta \left[ \log\left( (1 - \frac{R_1}{N})(1 - \frac{R_2}{N}) \right) \right] \\
&\quad - \frac{1}{2} \delta^2 \left[ \frac{1}{N} \frac{1}{(1 - R_1/N)(1 - R_2/N)} \right] + \mathcal{O}(\delta/N) + \mathcal{O}(\delta^3/N^2)
\end{aligned}
\tag{D.75}
$$

Combining these together, we have

$$
\begin{aligned}
\log P[p] \approx &- \log(p!) + p \log\left( \frac{R_1 R_2}{N} \right) - \\
&\frac{1}{2} \delta^2 \left[ \frac{1}{R_1} \frac{1}{1 - R_2/N} + \frac{1}{R_2} \frac{1}{1 - R_1/N} + \frac{1}{N} \frac{1}{(1 - R_1/N)(1 - R_2/N)} \right] \\
&+ \mathcal{O}(\delta/R) + \mathcal{O}(\delta^3/R^2).
\end{aligned}
\tag{D.76}
$$

Exponentiating,

$$
\begin{aligned}
P[p] \approx &\frac{1}{p!} \left( \frac{R_1 R_2}{N} \right)^p \times \\
&\exp\left( -\frac{1}{2} \left( p - \frac{R_1 R_2}{N} \right)^2 \left[ \frac{R_1 R_2}{R_1 + R_2} + \mathcal{O}(1/N) \right]^{-1} + \mathcal{O}(\delta/R) + \mathcal{O}(\delta^3/R^2) \right).
\end{aligned}
\tag{D.77}
$$

The first two terms are precisely a Poisson distribution, which has mean $R_1 R_2/N$ and width $\sqrt{R_1 R_2/N}$. The exponential is a Gaussian with the same mean $R_1 R_2/N$, and a larger width $\sqrt{R_1 R_2/(R_1 + R_2)}$. The smaller width determines the width of the product of the two functions, so we conclude:

$$
\langle p \rangle = \frac{R_1 R_2}{N}, \qquad \langle p^2 \rangle - \langle p \rangle^2 \approx \frac{R_1 R_2}{N}.
\tag{D.78}
$$

This is what we would expect for drawing $p$ random sites out of $N$, where each site has independent probability $R_i/N$ of being in either Pauli string (Section 9.7). The width is subextensive, $\delta \sim \varepsilon \sqrt{N}$, justifying the higher order terms we neglected along the way.

## Distribution of the $K$-size

Here we are interested in the probability distribution of the $K$-size of a Pauli string of fixed total size $\mathcal{S}$, with $K$ randomly chosen couplings. Our results substantiate those obtained by intuitive arguments beneath Eq. (9.35) in Section 9.4 in Ch. 9.

This objective is in fact an identical problem to calculating the overlap: the $K$-size is the overlap of the $K$ coupled qubits with the $\mathcal{S}$ qubits acted on by the operator of interest. We should just replace $R_1 \to K$, $R_2 \to \mathcal{S}$, $p \to n$ above, where $n$ is the $K$-size. This is confirmed by comparing the factorial expressions:

$$P[n] = \frac{C_n^{\mathcal{S}} C_{K-n}^{N-\mathcal{S}}}{C_K^N} = \frac{1}{n!} \frac{\mathcal{S}!}{(\mathcal{S}-n)!} \frac{K!}{(K-n)!} \frac{(N-\mathcal{S})!(N-K)!}{N!(N-\mathcal{S}-K+n)!} \qquad (D.79)$$

where the numerator computes the number of distinct configurations with $n$ qubits overlapping the Pauli operator support of size $\mathcal{S}$ and $K-n$ qubits not overlapping, and the denominator computes the number of distinct configurations of the $K$ coupled qubits. There are two regimes of interest: when $K$ and $\mathcal{S}$ are both extensive, and when $\mathcal{S}$ is extensive but $K$ is not. The former provides a more accurate measure of the full operator size $(K \to N)$, while the latter is relevant for probing the channel capacity. Both regimes share the same mean $K$-size $\mathcal{S}_K$ and $K$-size width $\delta \mathcal{S}_K$:

$$\mathcal{S}_K \equiv \langle n \rangle = \frac{\mathcal{S}K}{N}, \qquad \delta \mathcal{S}_K^2 \equiv \langle n^2 \rangle - \langle n \rangle^2 \approx \frac{\mathcal{S}K}{N} = \mathcal{S}_K. \qquad (D.80)$$

This matches our prediction in Section 9.4, which was based on a simple scenario of picking $K$ sites, each with a $\mathcal{S}/N$ chance of being in the support of the Pauli operator.

# D.7 Teleportation of fermions

Here we generalize the teleportation protocol to Majorana fermion systems, as discussed in Ch. 9 for the SYK model. This involves a few small modifications, stemming from ($i$) a different definition of fermionic EPR (FEPR) pairs, and ($ii$) a different relation between FEPR pair projector and the SWAP gate. These modifications explain the results of Ref. [159], which find that late time teleportation in the SYK model occurs with less than unit fidelity even at infinite temperature (where we would generally expect perfect fidelity, from late time peaked-size teleportation [Section 9.5, VI]). In particular, we find that the encoding procedure for the late time fermionic protocol must be modified for teleportation to succeed with perfect fidelity, due to modification ($ii$) above.

Consider two complex fermions $\chi_l$ and $\chi_r$, decomposed into pairs of Majorana fermions via $\chi_l = \psi_l^1 + i\psi_l^2$, $\chi_r = \psi_r^1 + i\psi_r^2$. The number operators of the original fermions are Majorana bilinears, e.g. $i\psi_l^1\psi_l^2 = 2\hat{N}_l - 1 = (-1)^{\hat{N}_l}$. We define a single FEPR pair as the positive eigenstate of $i\psi_l^1\psi_r^1$ and $i\psi_l^2\psi_r^2$. In the number operator basis of the original complex fermions, this is the maximally entangled state $(|10\rangle - i|01\rangle)/\sqrt{2}$. Multiple fermion EPR pairs are formed as a tensor product of single FEPR pairs.

This definition leads to some simple relations when 'sliding' fermion operators around FEPR bras and kets in diagrammatic calculations. We have:

$$\psi_l^j |\text{FEPR}\rangle = i\psi_r^j |\text{FEPR}\rangle$$
$$\langle\text{FEPR}| \psi_l^j = -i \langle\text{FEPR}| \psi_r^j,$$

(D.81)

diagrammatically,



(D.82)

As in bosonic systems, the thermofield double state is obtained by applying $\rho^{1/2}$ to one side, $|\text{TFD}\rangle = \rho_l^{1/2} |\text{FEPR}\rangle$. Since the SYK Hamiltonian is composed of 4-fermion terms, we have

$$H_l |\text{TFD}\rangle = (i)^4 H_r |\text{TFD}\rangle = H_r |\text{TFD}\rangle.$$

(D.83)

As in bosonic systems, the coupling for Majorana systems [Eq. (9.70)] measures the size of Majorana strings.

There are two options teleportation in fermionic system. First, we could teleport an ordinary bosonic qubit by encoding it into Majorana fermion operators, for instance:

$$X \equiv i\psi_1\psi_2$$
$$Y \equiv i\psi_2\psi_3$$
$$Z \equiv i\psi_1\psi_3.$$

(D.84)

At infinite temperature before coupling, each of the above operators has a correlator equal to $-1$, which is exactly the result for bosonic systems, but without a need for the decoding operation $Y$. At late times, the coupling $e^{igV}$ applies a relative phase between the identity and non-identity Paulis, giving correlator phases:

| $\mathbb{1}$ | $e^{ig\langle V\rangle}$ |
|---|---|
| $i\psi_1\psi_2$ | $-1$ |
| $i\psi_2\psi_3$ | $-1$ |
| $i\psi_1\psi_3$ | $-1$ |

When $g \langle V \rangle = \pi$ all correlators have the same phase, and peaked-size teleportation succeeds with perfect fidelity at infinite temperature. At intermediate times, peaked-size teleportation of multiple bosonic qubits will succeed just as in bosonic systems.

The second option is to send a fermionic qubit, for instance by inserting half of an ancillary FEPR pair. Here we begin with intermediate times, and discuss a modification necessary for late time teleportation afterwards. We represent a single complex fermion with two Majorana operators $\psi_1, \psi_2$, and suppose that the operators' size distributions are tightly peaked, and the size of $i\psi_1\psi_2$ is twice that of the individual Majorana sizes, denoted $\mathcal{S}$ (this assumption of size addition is appropriate in all-to-all coupled systems, e.g. SYK, but would not necessarily hold for e.g. a 1D Majorana chain). The relevant operator correlators after coupling with a bilinear fermionic interaction, as in Eq. (9.70), are:

| $\mathbb{1}$ | $1$ |
|---|---|
| $\psi_1$ | $-i \cdot e^{ig\mathcal{S}/qN}$ |
| $\psi_2$ | $-i \cdot e^{ig\mathcal{S}/qN}$ |
| $i\psi_1\psi_2$ | $-1 \cdot e^{ig2\mathcal{S}/qN}$ |

At $g\mathcal{S}/qN = \pi/2$ we have perfect teleportation. This generalizes straightforwardly to multiple fermionic qubits: a $p$-fermion operator will gain a phase $i^p$ from sliding across the FEPR pair, and a phase $e^{igp\mathcal{S}/qN}$ from coupling.

At late times, the sizes of initial single-body and two-body Majorana operators are equal, since they have saturated the size of the system, and the above operator correlators do not have the same phase. We now show that an alteration of the encoding procedure can rectify this and lead to perfect late time teleportation. This alteration is explained by the HPR protocol, and we derive it by reexamining the equivalence between the HPR and TW protocols in the case of fermionic qubits. Here, the relevant difference between bosons and fermions is that the fermionic SWAP gate is *not* related to the Grover search operation $1 - 2P_{\text{FEPR}}$ by single-qubit rotations. Since fermions gain a minus sign upon exchange, the fermionic SWAP gate takes the form

$$\text{SWAP}_F = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \frac{i\psi_{1,l}\psi_{2,l} + i\psi_{1,r}\psi_{2,r} + i\psi_{1,l}\psi_{2,r} - i\psi_{2,l}\psi_{1,r}}{2}. \tag{D.85}$$

This is a *two-qubit* controlled-phase (CZ) gate away from $1 - 2P_{\text{FEPR}}$:

$$1 - 2P_{\text{FEPR}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1 - i\psi_{1,l}\psi_{1,r} - i\psi_{2,l}\psi_{2,r} - (i\psi_{1,l}\psi_{1,r})(i\psi_{2,l}\psi_{2,r})}{2} = \text{SWAP}_F \cdot \text{CZ},$$

$$\tag{D.86}$$

where the CZ gate is defined as

$$
\text{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = (1+i)\frac{\psi_{1,l}\psi_{2,l} + i\psi_{1,r}\psi_{2,r}}{2}
$$
$$
= \exp\left(i\frac{\pi}{4}\right) \cdot \exp\left(-i\frac{\pi}{2}[i\psi_{1,l}\psi_{2,l}]\right) \cdot \exp\left(i\frac{\pi}{4}[i\psi_{1,l}\psi_{2,l}][i\psi_{1,r}\psi_{2,r}]\right).
$$

(D.87)

The single-fermion $\exp\left(-i\frac{\pi}{2}[i\psi_{1,l}\psi_{2,l}]\right)$ gate occurs on the swapped-out fermion and may be neglected. Inserting this in place of the second Grover search operation gives the appropriate teleportation protocol:



$$
F_{\text{EPR}} = \qquad = \qquad\qquad (D.88)
$$

In the second diagram we have slid the action of each side of the CZ gate such that the gate acts at the same time and on the same qubits as the initial SWAP gate.

We can relate the fidelity of teleportation to operator correlators by decomposing the encoding gate as

$$
\text{CZ} \cdot \text{SWAP}_F = \frac{1}{2}\sum_{j=1}^{4} S_{j,l}^{L} S_{j,r}^{R}
$$

(D.89)

where we define the operators:

| $j$ | $S_j^L$ | $S_j^R$ | $S_j^R S_j^L$ |
|---|---|---|---|
| 1 | $\mathbb{1}$ | $\mathbb{1}$ | $\mathbb{1}$ |
| 2 | $i\psi_1\psi_2$ | $i\psi_1\psi_2$ | $\mathbb{1}$ |
| 3 | $i\psi_1$ | $\psi_1$ | $i\mathbb{1}$ |
| 4 | $i\psi_2$ | $\psi_2$ | $i\mathbb{1}$ |

according to Eq. (D.86). The final column displays the product $S_j^L S_j^R$, where both act on the same qubit, which will be useful shortly. We find a fidelity:



$$(\text{D.90})$$

In the peaked-size regime with correlator phases $\theta_{R,j}$, we have

$$F_{\text{EPR}} = \frac{1}{2^4} \sum_{j,k} \langle \text{TFD}| \, S_{R,j,l}(t) \, e^{-igV} \, [S_{L,j,r} S_{L,k,r}^\dagger](-t') \, e^{igV} \, S_{R,k,l}^\dagger(t) \, |\text{TFD}\rangle$$

$$= \frac{1}{2^4} \sum_{j,k} \exp(-i[\theta_{R,j} - \theta_{R,k}]) \text{tr}(S_{R,j}(t-t') \, \rho^{1/2} \, S_{L,j}(0) \, S_{L,k}^\dagger(0) \, \rho^{1/2} \, S_{R,k}^\dagger(t-t'))$$

$$(\text{D.91})$$

At infinite temperature, late times, and $g\langle V \rangle = \pi$, we have correlator phases $\theta_{R,j} = 0$ for the identity and two-bosonic operator and $\theta_{R,j} = \pi/2$ for single-body fermionic operators,

and find perfect teleportation fidelity:

$$
\begin{aligned}
F_{\text{EPR}} &= \frac{1}{2^4} \sum_{j,k} \exp(-i[\theta_{R,j} - \theta_{R,k}]) \text{tr}(S_{R,j} \, S_{L,j} \, S_{L,k}^\dagger \, S_{R,k}^\dagger) \\
&= \frac{1}{2^4} \sum_{j,k} \exp(-i[\theta_{R,j} - \theta_{R,k}]) \cdot i^{F_j} \cdot (-i)^{F_k} \cdot \text{tr}(i\psi_1\psi_2 \, i\psi_1\psi_2) \\
&= \frac{1}{2^4} \sum_{j,k} \exp(-i[\theta_{R,j} - \theta_{R,k}]) \cdot i^{F_j} \cdot (-i)^{F_k} \\
&= \frac{1}{2^4} \sum_{j,k} (-i)^{F_j} \cdot i^{F_k} \cdot i^{F_j} \cdot (-i)^{F_k} \\
&= 1,
\end{aligned}
\tag{D.92}
$$

where we define $F_j = 1$ if $S_{L/R,j}$ is fermionic, and 0 if bosonic.

We note that for state, as opposed to EPR, teleportation, the above CZ gate turns out not to be necessary. Since coherent superpositions of different fermion parity cannot be created by physical Hamiltonians, which contain only even combinations of fermionic operators, we should only consider teleporting states of definite fermion parity. The CZ gate applies only an overall phase on these states, and so does not affect the success of teleportation.

We can also briefly analyze the low temperature results of Ref. [159] through the lens of operator correlator phases. Here, state teleportation is found to succeed perfectly at low temperatures only when the initial operators are encoded in $p$-body Majoranas, with $p = q/2 + 1$, despite the operator correlators having maximal magnitude for any value of $p$. At the semiclassical gravity teleportation time, the correlators have phases:

| | |
|---|---|
| $\mathbb{1}$ | 1 |
| $\psi_1$ | $i^p (i)^{2p/q}$ |
| $\psi_2$ | $i^p (i)^{2p/q}$ |
| $i\psi_1\psi_2$ | $(-1)^p (i)^{4p/q}$ |

For single-body Majoranas, $p = 1$, the correlators clearly do not have the same phase—in fact, their phases are nearly identical to their phases at infinite temperature with no coupling—so state teleportation is not possible. When $p = q/2 + 1$, in the large-$q$ limit, these phases are $1, \pm 1, \pm 1, 1$, respectively, where the sign is determined by whether $p = 1, 3 \bmod 4$. When the sign is odd, it can be corrected via the decoding operation $i\psi_1\psi_2 = (-1)^N$, which applies a minus sign when conjugating fermionic operators. Either case can therefore achieve perfect teleportation.

# D.8  Teleportation and inelastic scattering at infinite temperature

In Section 9.8, we found that strong stringy corrections to a bulk theory of gravity led to peaked-size teleportation as well as a deeply inelastic scattering amplitude. We will now demonstrate that these two phenomena—peaked-size teleportation and inelastic scattering—also coincide at infinite temperature, for arbitrary functional forms of the wavefunctions and scattering amplitudes. As we argued before, for a UV complete theory of quantum gravity, strong stringy (and in general deep inelastic) effects are expected to dominate only at high temperatures, $\beta \to 0$.

At infinite temperature, the form of the correlator is constrained by the equality

$$C_\psi(t; g)^* = -C_\psi(t; -g). \tag{D.93}$$

This implies that $C_\psi(t)$ can be written as a real function of $ig$ multiplied by the two-point function:

$$C_\psi(t) = \langle \psi_l \psi_r \rangle e^{-F(ig,t)}. \tag{D.94}$$

When $g = 0$, $C_\psi(t)$ is equal to $\langle \psi_l \psi_r \rangle$, implying

$$F(ig) = ig f_1(t) + \mathcal{O}(g^2), \tag{D.95}$$

where $f_1(t)$ is a real function. Therefore, at this order in $g$, the infinite temperature correlator is simply the two-point function multiplied by a pure phase, matching peaked-size teleportation [Eq. (9.38)].

To justify that higher order terms in $g$ are subleading, we need an additional assumption: that the wavefunction of $\psi(t)$ has a saddle point at some momentum $k$. This is analogous to the boundary assumption that operator sizes are tightly peaked. At early times, this saddle will not be significantly changed by the coupling, since the derivative of the scattering matrix with respect to $k$ will be suppressed by $G_N$, and at early times the time-dependence of the wavefunction will not be strong enough to compensate for this suppression (for example, in semiclassical AdS$_2$, we observed competition between $e^{2\pi t/\beta}$ and $1/G_N$). In such cases, it is easy to see that Eq. (9.96) becomes $\langle \psi_l \psi_r \rangle$ times a pure phase linear in $g$, with higher powers of $g$ suppressed by $G_N$.

Infinite temperature also implies purely inelastic scattering, i.e. the scattering amplitude, $e^{i\delta} = 1 - S(k, s)$, is automatically real. To see this, we first rewrite the correlator in terms of the in-falling momentum operators, $\hat{P}$ and $\hat{K}$, for $\psi_1$ and $\psi(t)$ respectively. For instance, for the former we have:

$$\begin{aligned}
\Psi_{1,r}(s, 0)\Psi_{1,l}^*(s, 0) &= \langle \psi_{1,l}(0)|s\rangle\langle s|\psi_{1,r}(0)\rangle \\
&= \int \frac{da}{2\pi} \langle \psi_{1,l}(0) e^{-ia\hat{P}} \psi_{1,r}(0)\rangle e^{ias}.
\end{aligned} \tag{D.96}$$

As $\psi(t)$ and $\psi_1$ are in principle independent operators, we have $[\hat{K}, \hat{P}] = 0$. Using this, we can rewrite Eq. (9.96) as

$$C_\psi(t) = \langle \psi_r(-t) \exp\left(-igS(\hat{K}, \hat{P})i\psi_{1,l}\psi_{1,r}\right) \psi_l(t)\rangle. \tag{D.97}$$

Taking the complex conjugate gives

$$\begin{aligned} C_\psi(t)^* &= \langle \psi_l(t) \exp\left(igS(\hat{K}, \hat{P})^*(-i)\psi_{1,r}\psi_{1,l}\right) \psi_r(-t)\rangle \\ &= -\langle \psi_r(-t) \exp\left(igS(\hat{K}, \hat{P})^*i\psi_{1,l}\psi_{1,r}\right) \psi_l(t)\rangle \end{aligned} \tag{D.98}$$

where we used the fact that $\hat{K}$, $\hat{P}$ are Hermitian and that at infinite temperature $\psi_l(t) |\text{TFD}\rangle = \psi_r(-t) |\text{TFD}\rangle$. Combining this with Eq. (D.93) then enforces $S(\hat{K}, \hat{P})^* = S(\hat{K}, \hat{P})$, i.e. purely inelastic scattering.

# Appendix E

# Details and additional derivations on operator growth in open quantum systems

In this Appendix we provide additional theoretical and numerical details relating to the results of Ch. 10.

## E.1   Additional analytical details and discussion

In this section, we provide substantially more details and discussion on the approximations leading to Eq. 10.4. We begin by clarifying our notion of operator growth under open-system dynamics, and outlining how this notion connects to different protocols that measure information scrambling. The remainder of the section serves to build intuition and outline the explicit steps that lead to Eq. 10.4. While these steps involve a substantial number of equations, we emphasize that the physical intuition behind Eq. 10.4 is fully contained in Ch. 10.

### Preliminaries

Defining operator growth under open-system dynamics involves a few subtleties not present in unitary dynamics. In Ch. 10, we focused on the time-evolution of operators in the Heisenberg picture [Eq. (10.1)]:

$$\partial_t \hat{M}_H(t) = i[\hat{H}, \hat{M}_H(t)] + \varepsilon \sum_\alpha \left( \hat{L}_\alpha^\dagger \hat{M}_H(t) \hat{L}_\alpha - \frac{1}{2}\{\hat{L}_\alpha^\dagger \hat{L}_\alpha, \hat{M}_H(t)\} \right). \qquad \text{(E.1)}$$

Heisenberg time-evolution is relevant when $\hat{M}$ is a Hermitian observable being measured in an experiment. The growth of $\hat{M}_H(t)$ signifies that the measurement outcome depends on

increasingly non-local correlations in the initial state of the system. Note that Heisenberg time-evolution is unital, i.e. it maps the identity to the identity.

One can also, in principle, consider the growth of operators in the Schrodinger picture. For instance, consider preparing a system in an initial density matrix that is almost entirely mixed, $\rho = \hat{M} \otimes \mathbb{1}^{N-1}/2^N$, where $\hat{M}$ is an initial single-qubit density matrix. The state now evolves via:

$$\partial_t \hat{M}_S(t) = -i[\hat{H}, \hat{M}_S(t)] + \varepsilon \sum_\alpha \left( \hat{L}_\alpha \hat{M}_S(t) \hat{L}_\alpha^\dagger - \frac{1}{2}\{\hat{L}_\alpha^\dagger \hat{L}_\alpha, \hat{M}_S(t)\} \right), \qquad \text{(E.2)}$$

which differs from Heisenberg time-evolution in the sign of the Hamiltonian and the ordering of the Lindblad operators. Note that Schrodinger and Heisenberg time-evolution coincide precisely when the Lindblad operators are Hermitian, $\hat{L}_\alpha = \hat{L}_\alpha^\dagger$, in which case we have $\hat{M}_H(t; \hat{H}) = \hat{M}_S(t; -\hat{H})$. This is trivially the case for unitary dynamics, and also holds for a variety of error models of practical interest (see the subsequent subsection).

Interpreting operator growth in the Schrodinger picture outside of the case $\hat{L}_\alpha = \hat{L}_\alpha^\dagger$ is more difficult. Fundamentally, this difficulty arises because Schrodinger time-evolution is not unital, i.e. the identity operator is time-evolved non-trivially. We illustrate this point with two examples. First, take the Lindblad operators to be zero and the Hamiltonian to implement unitary scrambling dynamics such that the operator $\hat{M}_S(t)$ grows to have support across the entire system. In this case, operator growth corresponds to the formation of non-local correlations under time-evolution. Second, take the Hamiltonian to be zero and the Lindbladian operators to implement spontaneous emission on each qubit (i.e. $\hat{L}_\alpha$ are lowering operators on each qubit). At late enough times, the density matrix will approach the all zero state, $\hat{M}_S(t) \to |0 \ldots 0\rangle\langle 0 \ldots 0|$. This operator is quite large (it has non-trivial support on every qubit of the system), yet its growth clearly does not represent any non-local correlations, since the system remains in a product state throughout time-evolution. For these reasons, in this work we direct our attention to operator growth under Heisenberg time-evolution.

The "correct" time-evolution is dictated by the experimental protocol under study. Let us specify to the protocol depicted in Fig. 10.4, and consider the behavior of the protocol when the unitaries, $U$ and $U^\dagger$, are instead replaced by open-system dynamics under Hamiltonians, $\hat{H}$ and $-\hat{H}$, and Lindblad operators, $\hat{L}_\alpha$. Here, for simplicity we assume the Lindblad operators are the same for forwards and backwards time-evolution, although this assumption is not essential in the subsequent sections. The operator $\hat{M}$ appears in two places in the protocol: First, describing the initial single-qubit state, $|\psi_\pm\rangle\langle\psi_\pm| = (1 \pm \hat{M})/2$; and second, describing the final observable to be measured, again $(1 \pm \hat{M})/2$. Here, we take $\hat{M}$ to be a single-qubit operator, and generalize from preparing only the positive eigenstate of $\hat{M}$ (as considered in Ch. 10) to either the positive or negative eigenstate. The expectation value of the final measurement represents the fidelity to recover the initial state. The fidelity consists

of one copy each of the Schrodinger and Heisenberg time-evolved operators,

$$
\begin{aligned}
\mathcal{F}_{\pm} &= \left\langle \frac{\hat{\mathbb{1}}_S(t; -\hat{H}) \pm \hat{M}_S(t; -\hat{H})}{2^N} \cdot e^{-\mu\mathcal{S}}\left\{ \frac{\hat{\mathbb{1}} \pm \hat{M}_H(t; \hat{H})}{2} \right\} \right\rangle \\
&= \frac{1}{2} \pm \left\langle \hat{\mathbb{1}}_S(t; -\hat{H}) \cdot e^{-\mu\mathcal{S}}\left\{ \hat{M}_H(t; \hat{H}) \right\} \right\rangle + \frac{1}{2} \left\langle \hat{M}_S(t; -\hat{H}) \cdot e^{-\mu\mathcal{S}}\left\{ \hat{M}_H(t; \hat{H}) \right\} \right\rangle,
\end{aligned}
\tag{E.3}
$$

where in the final expression we used the property, $\mathrm{tr}\big(\hat{M}_S(t)\big) = 0$, to neglect one cross term. We can eliminate the middle term on the RHS by averaging the fidelity over the positive and negative eigenstates, $\mathcal{F} = (\mathcal{F}_+ + \mathcal{F}_-)/2$. We have:

$$
\mathcal{F} = \frac{1}{2} + \frac{1}{2}\left\langle \hat{M}_S(t; -\hat{H}) \cdot e^{-\mu\mathcal{S}}\left\{ \hat{M}_H(t; \hat{H}) \right\} \right\rangle,
\tag{E.4}
$$

which is the result quoted in Ch. 10, upon the substitution $\hat{M}(t) \equiv \hat{M}_S(t; -\hat{H}) = \hat{M}_H(t; \hat{H})$ for Hermitian Lindblad operators.

We briefly comment on how open-system dynamics manifest in other protocols for measuring information scrambling. The above analysis applies to a large class of protocols that involve evolving once via forward time-evolution and once via backwards time-evolution [139, 163, 289, 326, 418, 529]. In contrast, protocols that measure scrambling via exclusively forward time-evolution [381, 473] will depend instead on inner products of two copies of the Heisenberg time-evolved operator (although we note that such protocols must in some cases be exponentially inefficient [111]). Finally, in teleportation-based protocols [**blok2020quantum**, 274, 429, 517, 518], the operator of interest enters neither a measured observable nor as an initial state, but rather as a quantum operation applied to one half of EPR pairs. In this case, one should properly consider the time-evolution of the *quantum channel*, $\mathcal{M}\{\cdot\} \equiv \hat{M}(\cdot)\hat{M}^\dagger$, which displays qualitatively different behavior from the Heisenberg time-evolution analyzed in this work [429].

## Examples of Lindblad operators

We can build intuition for the effect of Lindbladian time-evolution on operator growth in the Heisenberg and Schrodinger pictures by analyzing a few simple error models. Let us begin with the simplest case, of isotropic decoherence on each qubit. In this case, we found in Ch. 10 that the action of the Lindbladian was precisely given by the size superoperator:

$$
-\frac{1}{4}\sum_i \sum_P (P_i \hat{M}(t) P_i^\dagger - \hat{M}(t)) = \mathcal{S}\{\hat{M}(t)\} = \# \text{ of } X, Y, Z \text{ in } \hat{M}(t),
\tag{E.5}
$$

where the RHS is intended to indicate that the size superoperator multiplies each Pauli string, $R$, of the operator, $\hat{M}(t) = \sum_R c_R \hat{M}(t)$, by its number of non-identity components (i.e. the number of $X$-, $Y$-, or $Z$-components). Here the Pauli matrices run over $P \in \{\mathbb{1}, X, Y, Z\}$ on each qubit.

Different error models do not exactly measure the operator size. For instance, dephasing noise gives rise to the following superoperator:

$$-\frac{1}{2}\sum_i (Z_i \hat{M}(t) Z_i^\dagger - \hat{M}(t)) = \mathcal{S}_Z\{\hat{M}(t)\} = \# \text{ of } X, Y \text{ in } \hat{M}(t), \qquad \text{(E.6)}$$

while depolarizing noise gives rise to:

$$-\frac{1}{4}\sum_i (X_i \hat{M}(t) X_i^\dagger + Y_i \hat{M}(t) Y_i^\dagger - 2\hat{M}(t)) = \left[\mathcal{S} - \frac{1}{2}\mathcal{S}_Z\right]\{\hat{M}(t)\}$$

$$= \left(\# \text{ of } Z \text{ in } \hat{M}(t)\right) + \frac{1}{2}\left(\# \text{ of } X, Y \text{ in } \hat{M}(t)\right). \qquad \text{(E.7)}$$

In Ch. 10, we argue that these superoperators display similar phenomenology as the operator size when applied to chaotic many-body time-evolved operators. From the definitions above, we see that this is a reasonable assumption if the high-size components of time-evolved operators fluctuate randomly between the local choice of basis (i.e. in these examples, between single-qubit $X$-, $Y$-, and $Z$-components).

Non-Hermitian errors require more careful consideration than Hermitian errors, since their time-evolution differs between the Heisenberg and Schrodinger pictures. For instance, consider spontaneous emission on each qubit, i.e. Lindblad operators equal to the lowering operator, $\hat{\ell}_i = (\hat{X}_i - i\hat{Y}_i)/2$. We denote the raising operator, $\hat{r}_i = \hat{\ell}_i^\dagger$, the projector to the up eigenstate, $\hat{U}_i = \hat{r}_i\hat{\ell}_i = (\mathbb{1}_i + \hat{Z}_i)/2$, and the projector to the down eigenstate, $\hat{D}_i = \hat{\ell}_i\hat{r}_i = \mathbb{1}_i - \hat{U}_i$. In the Heisenberg picture, operators time-evolve via:

$$\partial_t \hat{M}_H(t) = i[\hat{H}, \hat{M}_H(t)] + \varepsilon\sum_i \left(\hat{r}_i \hat{M}_H(t)\hat{\ell}_i - \frac{1}{2}\{\hat{U}_i, \hat{M}_H(t)\}\right). \qquad \text{(E.8)}$$

Let us focus on the effect of spontaneous emission on a single component, $R_i$, of a single Pauli string, $\hat{R} = \bigotimes_i R_i$. We have:

$$-\hat{r}_i \hat{R}_i \hat{\ell}_i + \frac{1}{2}\{\hat{U}_i, \hat{R}_i\} = \begin{cases} 0, & \hat{R}_i = \mathbb{1}_i \\ \hat{X}_i/2, & \hat{R}_i = \hat{X}_i \\ \hat{Y}_i/2, & \hat{R}_i = \hat{Y}_i \\ 2\hat{U}_i, & \hat{R}_i = \hat{Z}_i \end{cases}. \qquad \text{(E.9)}$$

We see that the error cannot increase the size of a Pauli string, since each identity component is mapped to itself under Heisenberg time-evolution. Summing over all qubits, we have:

$$-\sum_i \left(\hat{r}_i \hat{R}\hat{\ell}_i - \frac{1}{2}\{\hat{U}_i, \hat{R}\}\right) = \left[\mathcal{S} - \frac{1}{2}\mathcal{S}_Z\right]\{\hat{R}\} - \sum_i \left(\delta_{R_i = Z_i}\right) \hat{\mathbb{1}}_i \otimes \hat{R}_{\neq i}. \qquad \text{(E.10)}$$

The effect of spontaneous emission in the Heisenberg picture is thus a combination of depolarizing noise (left term) and a superoperator that generates Pauli strings that differ from $\hat{R}$ by replacement of a single $Z$-component with the identity (right term). We do not expect the latter effect to substantially alter operator spreading, since the newly generated strings contain similar support to the original string. Indeed, in the following subsection we find that their effect is suppressed by a factor of $\sim \varepsilon/J$, where $J$ is the local Hamiltonian interaction strength.

As noted in the previous subsection, the time-evolution of operators in the Schrodinger picture can be drastically different. Again considering spontaneous emission, we have:

$$\partial_t \hat{M}_S(t) = -i[\hat{H}, \hat{M}_S(t)] + \varepsilon \sum_i \left( \hat{\ell}_i \hat{M}_S(t) \hat{r}_i - \frac{1}{2} \{ \hat{U}_i, \hat{M}_S(t) \} \right), \qquad \text{(E.11)}$$

where the Lindblad operators now act as,

$$- \hat{\ell}_i \hat{R}_i \hat{r}_i + \frac{1}{2} \{ \hat{U}_i, \hat{R}_i \} = \begin{cases} \hat{Z}_i, & \hat{R}_i = \mathbb{1}_i \\ \hat{X}_i/2, & \hat{R}_i = \hat{X}_i \\ \hat{Y}_i/2, & \hat{R}_i = \hat{Y}_i \\ \hat{Z}_i, & \hat{R}_i = \hat{Z}_i \end{cases}. \qquad \text{(E.12)}$$

Summing over all qubits, we now have:

$$- \sum_i \left( \hat{r}_i \hat{R} \hat{\ell}_i - \frac{1}{2} \{ \hat{U}_i, \hat{R} \} \right) = \left[ \mathcal{S} - \frac{1}{2} \mathcal{S}_Z \right] \{ \hat{R} \} - \sum_i \left( \delta_{R_i = \mathbb{1}_i} \right) \hat{Z}_i \otimes \hat{R}_{\neq i}. \qquad \text{(E.13)}$$

The effect of spontaneous emission in the Schrodinger picture is thus a combination of depolarizing noise (left term) and a superoperator that generates Pauli strings that differ from $\hat{R}$ by a single replacement of an *identity* component with a Pauli-$Z$ operator (right term). As in the previous subsection, we see that non-Hermitian Lindblad operators cause Schrodinger time-evolved operators to immediately gain support on all qubits, irrespective of the locality of the dynamics.

## Approximations leading to Eq. (10.4)

In Ch. 10, we provided intuitive arguments that open-system operator growth is governed by an effective Lindblad equation involving the size superoperator [Eq. (10.4)], in the case of chaotic dynamics and small error rates. In this subsection, we specify the precise approximations which lead to Eq. (10.4).

We begin by formally solving the Heisenberg and Schrodinger time-evolution equations as a Taylor series in the error rate. We have:

$$\hat{M}_H(t) = \sum_{n=0}^{\infty} \varepsilon^n \sum_{\{\alpha_i\}} \int_{t_1 \le \dots \le t_n} dt_1 \dots dt_n e^{-iH(t-t_n)} \mathcal{L}_{\alpha_n} \{ \dots \mathcal{L}_{\alpha_1} \{ e^{-iHt_1} M e^{iHt_1} \} \dots \} e^{iH(t-t_n)},$$

$$\text{(E.14)}$$

and

$$\hat{M}_S(t) = \sum_{n=0}^{\infty} \varepsilon^n \sum_{\{\alpha_i\}} \int_{t_1 \leq \ldots \leq t_n} dt_1 \ldots dt_n e^{-iH(t-t_n)} \mathcal{L}_{\alpha_n}^{\dagger} \{\ldots \mathcal{L}_{\alpha_1}^{\dagger} \{e^{-iHt_1} M e^{iHt_1}\} \ldots\} e^{iH(t-t_n)},$$

(E.15)

where for each Lindblad operator, $\hat{L}_\alpha$, we associate a superoperator, $\mathcal{L}_\alpha\{\cdot\} = \hat{L}_\alpha^{\dagger}(\cdot)\hat{L}_\alpha - \frac{1}{2}\{\hat{L}_\alpha^{\dagger}\hat{L}_\alpha, \cdot\}$, and its adjoint, $\mathcal{L}_\alpha^{\dagger}\{\cdot\} = \hat{L}_\alpha(\cdot)\hat{L}_\alpha^{\dagger} - \frac{1}{2}\{\hat{L}_\alpha^{\dagger}\hat{L}_\alpha, \cdot\}$. Inserting these solutions into Eq. (E.4), we find that the protocol in Fig. 10.4 measures the following quantity:

$$\left\langle \hat{M}_S(t; -\hat{H}) e^{-\mu \mathcal{S}} \left\{\hat{M}_H(t; \hat{H})\right\} \right\rangle$$

$$= \sum_{m,n=0}^{\infty} \varepsilon^{n+m} \sum_{\{\alpha_i\}} \sum_{\{\beta_i\}} \int_{s_1 \leq \ldots \leq s_m} ds_1 \ldots ds_m \int_{t_1 \leq \ldots \leq t_n} dt_1 \ldots dt_n \times$$

$$\left\langle e^{-iH(t-s_m)} \mathcal{L}_{\beta_m}^{\dagger} \{\ldots \mathcal{L}_{\beta_1}^{\dagger} \{e^{-iHs_1} M e^{iHs_1}\} \ldots\} e^{iH(t-s_n)} \right.$$

$$\left. \times e^{-\mu \mathcal{S}} \{e^{-iH(t-t_n)} \mathcal{L}_{\alpha_n} \{\ldots \mathcal{L}_{\alpha_1} \{e^{-iHt_1} M e^{iHt_1}\} \ldots\} e^{iH(t-t_n)}\} \right\rangle.$$

(E.16)

Each expectation value in the summation is a multi-point OTOC involving $\hat{M}$, the Lindblad operators, and the Pauli operators used to implement the quantum channel, $e^{-\mu \mathcal{S}}$.

We apply two approximations to this expression: (i) First, we assume that the precise form of operators entering the OTOC is only relevant for operators that occur at approximately the same "time slice", i.e. their times are separated by less than the inverse local interaction strength, $\sim 1/J$. (ii) Second, we assume that the only OTOCs that are non-zero are those for which all operators in every time-slice product to the identity, i.e. "off-diagonal" OTOCs such as $\langle \hat{Z}_i(t)\hat{X}_j(0)\hat{Z}_i(t)\hat{Y}_j(0) \rangle$ are zero. Properties (i) and (ii) are exact within large-$N$ systems and random matrix theory [399], and in random unitary circuits for single-qubit errors at leading order in the error rate. In the latter, condition (i) follows because the local basis is randomized with each time step, while condition (ii) follows because the sign of off-diagonal OTOCs can be flipped via single-qubit rotations (e.g. conjugation by the Pauli operator $X_j$ at time zero in the aforementioned example), which thus average the OTOC to zero. Corrections arise when the same Lindblad superoperator is applied in both forwards and backwards time-evolution, which occurs with a probability suppressed by $\varepsilon^2$. The application of these approximations to chaotic many-body systems is a conjecture, motivated by the intuition presented in Ch. 10.

We now apply approximations (i) and (ii) to Eq. (E.16). In particular, we show that approximation (ii) implies that operator growth under non-Hermitian Lindblad operators is equal to that under a particular set of Hermitian Lindblad operators, to leading order in $\varepsilon/J$. Since Schrodinger and Heisenberg time-evolution are equivalent for Hermitian Lindblad operators, this verifies that the operator size distribution measured via Eq. (E.4) is well-defined (i.e. it contains no negative elements). From this result, approximation (i) leads to Eq. (10.4).

To establish the first simplification, we define Hermitian operators, $\hat{L}_\alpha^+ = (\hat{L}_\alpha + \hat{L}_\alpha^\dagger)/2$ and $\hat{L}_\alpha^- = i(\hat{L}_\alpha - \hat{L}_\alpha^\dagger)/2$. Note that the inner product between the two Hermitian operators, $\langle \hat{L}_+\hat{L}_- \rangle = \frac{1}{2}[\langle \hat{L}^\dagger\hat{L}^\dagger \rangle - \langle \hat{L}\hat{L} \rangle]$, can be set to zero by multiplying $\hat{L}$ by an overall phase. The Lindblad superoperator can be expressed as:

$$\mathcal{L}_\alpha\{\cdot\} = \hat{L}_\alpha(\cdot)\hat{L}_\alpha^\dagger - \frac{1}{2}\{\hat{L}_\alpha^\dagger\hat{L}_\alpha, (\cdot)\}$$

$$= \mathcal{L}_\alpha^+\{\cdot\} + \mathcal{L}_\alpha^-\{\cdot\} - i\hat{L}_\alpha^-(\cdot)\hat{L}_\alpha^+ + i\hat{L}_\alpha^+(\cdot)\hat{L}_\alpha^- - \frac{i}{2}\{[\hat{L}_\alpha^-, \hat{L}_\alpha^+], (\cdot)\}$$

$$\mathcal{L}_\alpha^\dagger\{\cdot\} = \hat{L}_\alpha^\dagger(\cdot)\hat{L}_\alpha - \frac{1}{2}\{\hat{L}_\alpha^\dagger\hat{L}_\alpha, (\cdot)\}$$

$$= \mathcal{L}_\alpha^+\{\cdot\} + \mathcal{L}_\alpha^-\{\cdot\} + i\hat{L}_\alpha^-(\cdot)\hat{L}_\alpha^+ - i\hat{L}_\alpha^+(\cdot)\hat{L}_\alpha^- - \frac{i}{2}\{[\hat{L}_\alpha^-, \hat{L}_\alpha^+], (\cdot)\}$$

(E.17)

where $\mathcal{L}_\alpha^\pm\{\cdot\}$ are the analogous Lindblad superoperators for $\hat{L}_\alpha^\pm$.

Now, note that the typical separation in time of spatially and temporally nearby Lindblad operators in Eq. (E.16) is given by the inverse local error rate, $\sim 1/\varepsilon$. For small error rates, this is much longer than the thermalization time, $\sim 1/J$; thus we can assume that each Lindblad superoperator in Eq. (E.16) occurs at a different "time-slice" for the purposes of approximations (i), (ii). Applying Eq. (E.17), we see that each time-slice contains at most two of a given Lindblad operator: either two copies of $\hat{L}_\alpha^+$ (from $\mathcal{L}_\alpha^+\{\cdot\}$), two copies of $\hat{L}_\alpha^-$ (from $\mathcal{L}_\alpha^-\{\cdot\}$), or one copy each of $\hat{L}_\alpha^+$ and $\hat{L}_\alpha^-$ (from the latter four terms). However, the latter terms are precisely those that vanish under approximation (ii). Therefore, to leading order in $\varepsilon/J$, we find that time-evolution is effectively governed by the Hermitian Lindblad operators, $\hat{L}_\alpha^\pm$. Concretely, we can replace $\hat{M}_H(t; \hat{H}) \to \hat{M}(t)$ and $\hat{M}_S(t; -\hat{H}) \to \hat{M}(t)$ in Eq. (E.16), where the operator, $\hat{M}(t)$, evolves as:

$$\partial_t\hat{M}(t) = i[\hat{H}, \hat{M}(t)] + \varepsilon\sum_\alpha \left(\mathcal{L}_\alpha^+\{\hat{M}(t)\} + \mathcal{L}_\alpha^-\{\hat{M}(t)\}\right). \tag{E.18}$$

For example, in the case of spontaneous emission, with non-Hermitian Lindblad operator $\hat{L}_i = \hat{\ell}_i$, we have Hermitian Lindblad operators, $\hat{L}_i^+ = X$ and $\hat{L}_i^- = Y$, which represent an effective depolarizing channel. Intuitively, this occurs because the Pauli strings in the final right term of Eqs. (E.10,E.13) are effectively uncorrelated with those of the original operator after time $\sim 1/J$, and thus only contribute at higher order in $\varepsilon/J$.

Applying approximation (i) to Eq. (E.18) leads us to Eq. (10.4). Specifically, we replace each Lindblad superoperator, $\mathcal{L}_\alpha$, with a sum of Pauli Lindblad superoperators, $\mathcal{L}_{P_i}$, acting on the same qubit:

$$\partial_t\hat{M}(t) \approx i[\hat{H}, \hat{M}(t)] + \tilde{\varepsilon}\sum_i\sum_{P_i}\mathcal{L}_{P_i}\{\hat{M}(t)\} = i[\hat{H}, \hat{M}(t)] + 4\tilde{\varepsilon}\mathcal{S}\{\hat{M}(t)\}. \tag{E.19}$$

To be precise, we can adjust the local error rates so that the normalization of the Lindblad operators is unchanged, $\varepsilon\langle\hat{L}_\alpha^\dagger\hat{L}_\alpha\rangle = \tilde{\varepsilon}\sum_{P_i\neq\mathbb{1}_i}\langle\hat{P}_i\hat{P}_i\rangle = 3\tilde{\varepsilon}$. In the remainder of the work,

we replace the notation $\tilde{\varepsilon}$ with $\varepsilon$ for simplicity. If $\mathcal{L}_\alpha$ has support on multiple qubits but the qubits are nearby each other in space, we expect this approximation to apply when $i$ is chosen as any one of the qubits, because the support of $\hat{M}(t)$ will be highly correlated between nearby qubits in a locally interacting system. If $\mathcal{L}_\alpha$ has support on multiple qubits in an all-to-all coupled system, we expect the approximation to apply when $i$ is summed over all qubits, since the support of $\hat{M}(t)$ will be nearly entirely uncorrelated between qubits.

Before concluding, we mention a few notable violations of the approximations (i) and (ii), and show that they nonetheless have a subleading effect. First, note that systems with local conserved quantities violate both approximations. Regarding approximation (i), Lindblad operators that overlap vs. do not overlap the conserved quantity will spread differently; for approximation (ii), certain terms with zero trace in each time-slice are non-zero, for example the two-point function, $\langle \hat{Z}_i(t)\hat{Z}_j(0)\rangle$. Nonetheless, we expect both of these violations to give only subleading contribution (suppressed by $1/\sqrt{t}$), since the majority of a time-evolved operator's support is contained in chaotic large-size Pauli strings.

Second, a distinct argument shows that the approximation (ii) is violated in Hamiltonian systems. Consider the time-derivative of the OTOC under unitary dynamics:

$$\partial_t \langle \hat{M}(t)\hat{Q}\hat{M}(t)\hat{Q}\rangle = i\langle \hat{M}(t)[\hat{Q}, \hat{H}]\hat{M}(t)\hat{Q}\rangle. \tag{E.20}$$

The right side is an "off-diagonal" OTOC since the trace of all operators at time zero vanishes, yet it is non-zero whenever the OTOC on the left side changes in time. We now show that this violation gives only a subleading contribution to Eq. (E.4) in two examples. In all-to-all interacting systems, the right side of Eq. (E.20) contains a sum of at least $\sim N$ Hamiltonian terms, which leads to a $\sim 1/N$ suppression in the value of each off-diagonal OTOC. Meanwhile, in a 1D system, the left side OTOC (before differentiation) is zero when the operator $Q$ is within the light cone of $\hat{M}(t)$, and unity when it is outside. The time-derivative of the OTOC is therefore non-zero only over a small region of width $\sim c\sqrt{t}$, with a magnitude $\sim v_B/c\sqrt{t}$. The total contribution of "off-diagonal" OTOCs to Eq. (E.4) therefore scales as $\varepsilon(c\sqrt{t})(v_B/c\sqrt{t}) \sim \varepsilon \partial_t \mathcal{S}/J \sim \varepsilon$, which is much less than the contribution of "diagonal" OTOCs, $\sim \varepsilon \mathcal{S}$.

## Derivation of Eqs. (10.5), (10.6)

In this section, we derive the Loschmidt echo decay equation [Eq. (10.5)] and the size growth equation [Eq. (10.6)]. Both equations are obtained directly by applying Eq. (10.4) to the quantities of interest. We begin with the Loschmidt echo fidelity, $\mathcal{N} = \langle \hat{M}(t)\hat{M}(t)\rangle$. The time-derivative is:

$$\begin{aligned}
\partial_t \mathcal{N}(t) &= 2\langle \partial_t \hat{M}(t)\hat{M}(t)\rangle \\
&= -2\varepsilon \langle \mathcal{S}\{\hat{M}(t)\}\hat{M}(t)\rangle \\
&= -2\varepsilon \overline{\mathcal{S}}(t)\mathcal{N}(t).
\end{aligned} \tag{E.21}$$

using the definition of the average operator size, $\overline{\mathcal{S}}(t) = \langle \mathcal{S}\{\hat{M}(t)\}\hat{M}(t)\rangle / \langle \hat{M}(t)\hat{M}(t)\rangle$. Taking the time-derivative of the average operator size gives:

$$
\begin{aligned}
\partial_t \overline{\mathcal{S}}(t) &= \frac{\partial_t \langle \hat{M}(t)\mathcal{S}\{\hat{M}(t)\}\rangle}{\langle \hat{M}(t)\hat{M}(t)\rangle} - \frac{\langle \hat{M}(t)\mathcal{S}\{\hat{M}(t)\}\rangle \cdot \partial_t \langle \hat{M}(t)\hat{M}(t)\rangle}{\langle \hat{M}(t)\hat{M}(t)\rangle^2} \\
&= 2i \left( \frac{\langle [\hat{H}, \hat{M}(t)]\mathcal{S}\{\hat{M}(t)\}\rangle}{\langle \hat{M}(t)\hat{M}(t)\rangle} \right) - 2\varepsilon \left( \frac{\langle \hat{M}(t)\mathcal{S}^2\{\hat{M}(t)\}\rangle}{\langle \hat{M}(t)\hat{M}(t)\rangle} \right) + 2\varepsilon \left( \frac{\langle \hat{M}(t)\mathcal{S}\{\hat{M}(t)\}\rangle^2}{\langle \hat{M}(t)\hat{M}(t)\rangle^2} \right) \\
&= 2i \left( \frac{\langle [\hat{H}, \hat{M}(t)]\mathcal{S}\{\hat{M}(t)\}\rangle}{\langle \hat{M}(t)\hat{M}(t)\rangle} \right) - 2\varepsilon \delta \mathcal{S}(t)^2,
\end{aligned}
$$
(E.22)

where the size width is defined as $\delta \mathcal{S}(t)^2 = \overline{\mathcal{S}^2}(t) - \overline{\mathcal{S}}(t)^2$. The first term corresponds to size growth via the unitary dynamics generated by $\hat{H}$.

## Unequal Hamiltonian formulation of the Loschmidt echo

In this section, we address the alternate Loschmidt echo scenario discussed in Ch. 10, in which time-evolution is unitary but is governed by unequal Hamiltonians for forwards and backwards time-evolution. Similar to the previous subsections, we outline the precise approximations which lead to a Loschmidt echo fidelity described by Eq. (10.4). In what follows, we begin by Taylor expanding the fidelity in the difference between the two Hamiltonians, in which case each term is again a sum of various OTOCs. We then apply a single assumption: that these correlation functions relax quickly outside of some correlation length, $\xi_{\text{th}}$, and correlation time, $\tau_{\text{th}}$. We expect these assumptions to hold for chaotic many-body systems when the perturbation $\delta H$ does not overlap any conserved quantities (e.g. the Hamiltonian, $H$). In principle, we expect modifications to our predictions when the perturbation does overlap a conserved quantity. In Fig. 10.3 we explore this scenario numerically for small system sizes, and observe qualitatively similar behavior to that expected from Lindbladian dynamics.

The fidelity of the many-body Loschmidt echo is given by the correlation function:

$$
\mathcal{F}_\eta(t) = \langle e^{-iHt} M e^{iHt} e^{-i(H+\eta\,\delta H)t} M e^{i(H+\eta\,\delta H)t} \rangle.
$$
(E.23)

We begin our derivation by Taylor expanding in the perturbation strength, $\eta$:

$$
\mathcal{F}_\eta(t) = \sum_n (i\eta)^n \mathcal{C}_n(t)
$$
(E.24)

where the $n^{\text{th}}$ coefficient is given by:

$$
\begin{aligned}
\mathcal{C}_n(t) = \sum_{\text{bipart}} (-1)^{n_L} \int dt_1^L \dots dt_{n_L}^L \int dt_1^R \dots dt_{n_R}^R \times \\
\text{tr}(e^{-iHt} M e^{iHt} e^{-iH(t-t_{n_L}^L)} \delta H \dots e^{-iH(t_2^L - t_1^L)} \delta H e^{-iHt_1^L} M \\
\times e^{iHt_1^R} \delta H e^{iH(t_2^R - t_1^R)} \dots \delta H e^{iH(t-t_{n_R}^R)})
\end{aligned}
$$
(E.25)

Here the summation occurs over all bipartitions, $(n_L, n_R)$, with $n_L + n_R = n$ and $n_L, n_R \geq 0$ (since each $\eta$-derivative can hit either $e^{-i(H+\eta\,\delta H)t}$ on the left, or $e^{i(H+\eta\,\delta H)t}$ on the right). The time-ordering is $t_1^L \leq t_2^L \leq \ldots \leq t_{n_L}^L$ and $t_1^R \leq t_2^R \leq \ldots \leq t_{n_R}^R$. Each term in the above equation is an OTOC,

$$
\langle e^{-iHt} M e^{iHt} e^{-iH(t-t_{n_L}^L)} \delta H \ldots e^{-iH(t_2^L - t_1^L)} \delta H e^{-iHt_1^L} M
$$
$$
\times\, e^{iHt_1^R} \delta H e^{iH(t_2^R - t_1^R)} \ldots \delta H e^{iH(t-t_{n_R}^R)} \rangle \tag{E.26}
$$
$$
= \langle \hat{M}(t) \delta H(t - t_{n_L}^L) \ldots \delta H(t - t_1^L) \hat{M}(t) \delta H(t - t_1^R) \ldots \delta H(t - t_n^R) \rangle
$$

To proceed farther, we approximate that each OTOC is zero unless its set of times, $\{t_1^L, \ldots t_{n_L}^L, t_1^R, \ldots, t_{n_R}^R\}$, can be decomposed into pairs of equal times (note that this is equivalent to approximation (ii) in the previous subsection). Note that this immediately eliminates all $\mathcal{C}_n(t)$ with $n$ odd. These pairs can occur either between times on the same side (e.g. $t_1^L = t_2^L$ or $t_1^R = t_2^R$), or on opposite sides (e.g. $t_1^L = t_1^R$). Summing these four possibilities for a fixed set of times, we find that each pair enters in the form:

$$
\mathcal{L}_{\delta H}\{(\ldots)\hat{M}(\ldots)\} = \delta H \delta H(\ldots)\hat{M}(\ldots) + (\ldots)\hat{M}(\ldots)\delta H \delta H - 2\delta H(\ldots)\hat{M}(\ldots)\delta H, \tag{E.27}
$$

where $\mathcal{L}_{\delta H}$ is a superoperator and the ellipses represent intermediate time-evolution. The sum over bipartitions in Eq. (E.25) can thus be re-expressed as:

$$
\mathcal{C}_{2n}(t) \approx \tau_{\text{th}}^n \int dt_1 \ldots dt_n \langle e^{-iHt} M
$$
$$
\times\, e^{iHt} e^{-iH(t-t_n)} \mathcal{L}_{\delta H}\{\ldots e^{-iH(t_2-t_1)} \mathcal{L}_{\delta H}\{e^{-iHt_1} M e^{iHt_1}\} e^{iH(t_2-t_1)} \ldots\} e^{iH(t-t_n)} \rangle. \tag{E.28}
$$

This can be interpreted as the connected part of the OTOC between $M$ and $n$ copies of $\delta H$.

We have not yet assumed anything about the spatial correlations of $\delta H$. However, we can already connect our expression for the Loschmidt echo fidelity to an effective Lindblad equation. Specifically, applying the above results, we can write:

$$
\mathcal{F}_p(t) \approx \sum_n (-p^2)^n \mathcal{C}_{2n}(t) = \langle M_t \tilde{M}_t \rangle \tag{E.29}
$$

with the operator $\tilde{M}_t$ defined as:

$$
\tilde{M}_t \equiv \sum_n (-\tau_{\text{th}} p^2)^n \int dt_1 \ldots dt_n e^{-iH(t-t_n)} \mathcal{L}_{\delta H}\{\ldots e^{-iH(t_2-t_1)} \mathcal{L}_{\delta H}\{e^{-iHt_1} M
$$
$$
\times\, e^{iHt_1}\} e^{iH(t_2-t_1)} \ldots\} e^{iH(t-t_n)}. \tag{E.30}
$$

Taking a time-derivative (note that the time-derivative hits three locations: the first and final applications of $e^{\pm iH(t-t_n)}$, and the upper limit of the integral over $0 \leq t_n \leq t$), we have:

$$
\begin{aligned}
\partial_t \tilde{M}_t &= -i[H, \tilde{M}_t] + \sum_n (-p^2)^n \int dt_1 \ldots dt_{n-1} \mathcal{L}_{\delta H} \{ \ldots e^{-iH(t_2-t_1)} \mathcal{L}_{\delta H} \{ e^{-iHt_1} M \\
&\quad \times e^{iHt_1} \} e^{iH(t_2-t_1)} \ldots \} \\
&= -i[H, \tilde{M}_t] - p^2 \mathcal{L}_{\delta H} \{ \tilde{M}_t \}. \\
&= -i[H, \tilde{M}_t] - 2\tau_{\mathrm{th}} p^2 \left[ \frac{1}{2} \delta H^2 \tilde{M}_t + \frac{1}{2} \tilde{M}_t \delta H^2 - \delta H \tilde{M}_t \delta H \right].
\end{aligned}
\tag{E.31}
$$

The operator is thus the solution to an effective Lindblad equation with a single Lindblad operator, $\hat{L} = \delta H$, and error rate, $\varepsilon = 2\tau_{\mathrm{th}} p^2$. Note that we could have symmetrized the above equation by re-defining the original Hamiltonians so that $H_1 = H - \eta \, \delta H/2$ and $H_2 = H + \eta \, \delta H/2$.

To proceed further, we assume that the perturbation is a sum of local operators, $\delta H = \sum_i \delta H_i$, and that OTOCs involving a pair of local operators at the same fixed time are zero unless the operators are spatially nearby one another. Our notion of 'local' and 'nearby' can be clarified. For locally interacting systems, we have in mind that each $\delta H_i$ is contained within a finite spatial region, e.g. a ball of finite radius. We then assume that the relevant OTOCs are zero unless the pair of local terms within each fixed-time application of $\mathcal{L}_{\delta H}$ are within distance $\xi_{\mathrm{th}}$ of each other—there are $\sim \xi_{\mathrm{th}}^d$ such pairs in dimension $d$. In all-to-all coupled systems, we have in mind that $\delta H_i$ acts on only one of the all-to-all coupled degrees of freedom (e.g. a single qubit in an all-to-all coupled spin system, or a single Majorana fermion in SYK). The OTOCs between pairs of different $\delta H_i$ and $\delta H_j$ are then zero whenever $i \neq j$. For concreteness, we focus on local systems in what follows.

Applying this assumption to the effective Lindblad equation, Eq. (E.31), immediately yields a new effective Lindblad equation:

$$
\partial_t \tilde{M}_t = -i[H, \tilde{M}_t] - 2\tau_{\mathrm{th}} \xi_{\mathrm{th}} p^2 \sum_i \left[ \frac{1}{2} \delta H_i^2 \tilde{M}_t + \frac{1}{2} \tilde{M}_t \delta H_i^2 - \delta H_i \tilde{M}_t \delta H_i \right],
\tag{E.32}
$$

which involves only local Lindblad operators, $\hat{L}_i = \delta H_i$, each with error rate, $2\tau_{\mathrm{th}} \xi_{\mathrm{th}} p^2$. The approximation (i) in the previous subsection then leads to Eq. (10.4).
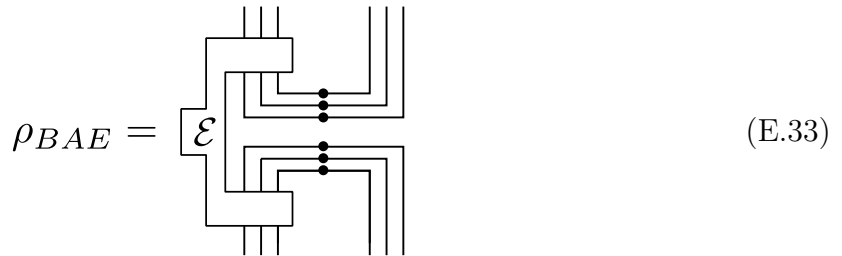
## E.2  Information-theoretic interpretation of the Loschmidt echo

In this section, we show that the Loschmidt echo with respect to a local operator $\hat{M}$ is related to a Renyi-2 mutual information, between the many-body input state of Lindbladian time-evolution and the subsystem of the output state that $\hat{M}$ act upon. Specifically, this

relation holds for the average Loschmidt echo fidelity over a complete basis of operators $\hat{M}$ on the subsystem of interest.

To define the relevant subsystems, we consider a setup where our many-body system $S$ is initially maximally entangled with a reference system $E$. The state then undergoes Lindbladian time-evolution, i.e. we apply the quantum channel, $\mathcal{E} = e^{\mathcal{L}t}$, where $\mathcal{L}$ is the Lindbladian superoperator on system $S$. We are interested in the Renyi-2 mutual information between $E$ and a subsystem $A$ of $S$.

To compute the mutual information, we utilize a tensor network-like diagrammatic notation introduced in Ref. [517] (see Ref. [429] for a more extensive introduction). We represent the density matrix of the entire system, $SE$, after Lindbladian time-evolution via the following diagram:

$$\rho_{BAE} = \quad \boxed{\mathcal{E}} \quad \hspace{3cm} \text{(E.33)}$$

Here system $S$ is represented via three pairs of tensor indices on the left side (decomposed into subsystem $A$ and its complement, $B$), and system $E$ similarly on the right. We denote the number of qubits in $S$ as $N$ and its Hilbert space dimension as $d = 2^N$, and similar for subsystem $A$, with $n_A$ qubits and dimension $d_A = 2^{n_A}$. For each subsystem, the tensor legs terminating downward represent the input to the density matrix (i.e. the ket indices), while the legs terminating upward represent the output (i.e. the bra indices). The middle horizontal lines represent initial EPR pairs between subsystems A, B and the reference system E. Each dot represents a factor of the inverse square root dimension of the tensor leg it is located on, to normalize the EPR pairs. The quantum channel $\mathcal{E}$ acts on system $S$, i.e. the left half of the EPR pairs.

The Renyi-2 mutual information of interest is defined as:

$$I^{(2)}(A, E) = S^{(2)}(E) + S^{(2)}(A) - S^{(2)}(AE) = N + n_A - S^{(2)}(AE), \quad \text{(E.34)}$$

where we use the fact $E$ and $A$ are in maximally mixed states. Here, the Renyi-2 entropy is defined as:

$$S^{(2)}(AE) = -\log_2\left[\text{tr}(\rho_{AE}^2)\right], \quad \text{(E.35)}$$

with $\rho_{AE} = \text{tr}_B(\rho_{BAE})$. In the diagrammatic notation, the quantity $\text{tr}(\rho_{AE}^2)$ can be expressed

as:



$$\text{tr}(\rho_{AE}^2) = \quad [\text{diagram}] \quad = \frac{1}{d_A} \sum_{P_A} [\text{diagram}] \quad = \frac{1}{d_A} \sum_{P_A} [\text{diagram}]$$

(E.36)

The leftmost diagram follows straightforwardly from Eq. (E.33) by tracing over subsystem $B$, squaring, and then tracing over $A$ and $E$. In the first equality, we use the identity, $\text{tr}_A(\cdot) = \frac{1}{d_A} \sum_{P_A} \hat{P}_A(\cdot)\hat{P}_A^\dagger$, where $\hat{P}_A$ form a complete basis for operators on subsystem $A$ (e.g. the Pauli operators), to re-arrange the tensor indices. In the second equality, we complex conjugate the entire diagram and apply the definition of the adjoint in order to "slide" the channel $\mathcal{E}$ across the EPR pairs (note that $\text{tr}(\rho_{AE}^2)$ is real and so is invariant under complex conjugation; we also re-index the summation $\hat{P}_A \to \hat{P}_A^*$ for convenience). This can be derived explicitly by decomposing $\mathcal{E}$ in terms of Kraus operators, $\mathcal{E}(\cdot) = \sum_i \hat{E}_i(\cdot)\hat{E}_i^\dagger$, and utilizing the identity [429], $\hat{E}_{i,\text{left}}^T |\text{EPR}\rangle = \hat{E}_{i,\text{right}} |\text{EPR}\rangle$, for each Kraus operator. Now, we note that this diagram contains precisely the Heisenberg time-evolution of the operators, $\hat{P}_A(t) = \mathcal{E}^\dagger\{\hat{P}_A\}$. Contracting the tensor indices and keeping track of the various Hilbert space dimensions, we thus arrive at our final result:

$$I^{(2)}(A, E) = N + n_A + \log_2\left[\frac{1}{dd_A}\sum_{P_A}\left\langle \hat{P}_A(t)\hat{P}_A^\dagger(t)\right\rangle\right]$$

$$= 2n_A + \log_2\left[\frac{1}{d_A^2}\sum_{P_A}\left\langle \hat{P}_A(t)\hat{P}_A^\dagger(t)\right\rangle\right].$$

(E.37)

The argument of the logarithm is the average operator normalization over a complete basis of $d_A^2$ operators on subsystem $A$. We can easily check the limits on this equation. Under unitary dynamics, the average normalization is unity and so the mutual information is maximal, $I^{(2)}(A, E) = 2n_A$. Under a maximally decohering channel, the normalization approaches zero for all operators except the identity. This gives an average normalization of $1/d_A^2$, and thus a mutual information of zero, $I^{(2)}(A, E) = 2n_A - 2n_A = 0$.

# E.3 Open-system operator growth in long-range interacting systems

In this section, we turn to operator growth in long-range interacting open-system dynamics. We begin by summarizing previous results on operator growth under long-range interacting unitary dynamics. We then provide a brief intuitive picture to explain the effect of open-system dynamics, and explore these predictions numerically in random unitary circuits (Fig. E.1).

Operator growth in long-range interacting dynamics has been studied in the context of random unitary circuits [97, 531] and biological physics [93, 192], and is known to display qualitatively distinct behavior compared to short-range dynamics. In particular, these works establish several regimes of operator growth, dependent on the power-law, $\alpha$, of the interaction and the spatial dimension, $d$. Here, long-range hops are designated to occur with a probability, $\sim 1/r^{2\alpha}$. For $d + 1/2 < \alpha$, operator growth proceeds similarly to in short-range dynamics: the average operator size grows ballistically, $\mathcal{S} \sim t^d$, and the size width grows at a slower rate dominated by fluctuations at the light cone boundary. For $d < \alpha < d + 1/2$, the average operator size exhibits a *super-ballistic* power law in time, $\mathcal{S} \sim t^{\frac{1}{2\alpha - 2d}}$. Moreover, fluctuations in the light cone boundary become the same order as the light cone width [192], implying that the size distribution becomes broad ($\delta \mathcal{S} \sim \overline{S}$). For $d/2 < \alpha < d$, the average operator size exhibits a stretched exponential, $\mathcal{S} \sim e^{t^{\log_2(d/\alpha)}}$, again with large fluctuations at the boundary. Finally, for $\alpha < d/2$, the system is effectively all-to-all coupled and operators grow according a simple exponential in time.

In Ref. [192], all of the above patterns of growth are explained by a single remarkably simple model for operator growth. The central observation of this model is that the value of the light cone radius at time $t$, $r(t)$, is determined primarily by contributions from long-range hops that were "seeded" from the bulk of the light cone at an earlier time, $t_{\text{seed}} \approx r^{-1}(r(t)/2)$. These seeds grow separately from the bulk of the light cone until the time of interest, $t$, at which the light cones of the seed and core combine. The radius $r(t)$ can then be solved for in a self-consistent approach, by demanding that the expected number of such hops up to the time $t$ [expressed in terms of $d$, $\alpha$, and $r(t)$] is equal to one. For more details, we refer to Ref. [192]. We note that this simple picture explains why the size distributions are broad (for power laws $\alpha < d + 1/2$). The long-range hops that contribute to the light cone boundary at time $t$ occur via a Poisson process with expected value 1. This leads to comparable fluctuations in the number of such hops; in particular, there will be some $\sim \mathcal{O}(1)$ probability that no hop of distance $r(t)$ has occurred at all. This is in sharp contrast to operator growth in short-range systems, where fluctuations in the light cone boundary receive equal contributions from fluctuations at all previous times [340].

A full analysis of long-range operator spreading in the presence of error would represent a substantial undertaking. We note that the manifestation of error that is relevant for time-evolution of quantum mechanical operators has no physical analog in the biological context where the previous studies of operator spreading originated [93, 97, 192, 531]. Here, we
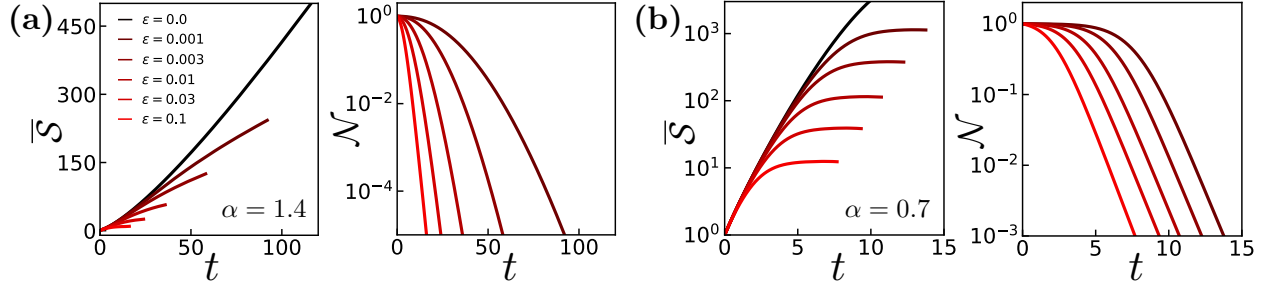
Figure E.1: Operator growth in long-range interacting random unitary circuits under unitary
and open-system dynamics. **(a)** Power-law exponent $\alpha = 0.14$. The average operator size
grows super-linearly under unitary dynamics (black), according to the power-law [531], $\overline{\mathcal{S}} \sim
t^{1/(2\alpha-1)}$. Under open-system dynamics (red), the average operator size eventually reverts
to sub-linear growth. **(b)** Power-law exponent $\alpha = 0.7$. The average operator size grows
super-linearly under unitary dynamics (black), now according to a shrunk exponential [531],
$\overline{\mathcal{S}} \sim e^{t^{\log_2(1/\alpha)}}$. Under open-system dynamics (red), the average operator size eventually
plateaus, similar to our observation in all-to-all interacting systems [Fig. 10.2(b)]. The
arguments in the text predict that this apparent plateau is in fact a slow linear growth,
however, it is difficult to precisely distinguish these the two scenarios here.

instead propose a simple modification of the intuitive picture above. Consider two possi-
ble outcomes of the aforementioned Poisson process: one in which no hop of distance $r(t)$
occurs, and one in which $\geq 1$ such hops occur. Both outcomes occur with $\sim \mathcal{O}(1)$ proba-
bility under unitary dynamics. However, under open-system dynamics, the latter scenario
will have be suppressed *exponentially* in the operator size of the seed integrated over time,
i.e. proportional to $\sim \exp\left(-\varepsilon \int_{t_{\text{seed}}}^{t} dt' \, \mathcal{S}_{\text{seed}}(t')\right)$. We therefore expect that sufficiently long-
range hops will cease to contribute to open-system operator growth, when this suppression
becomes small. This implies that at sufficiently late times, operator growth will asymptote
to ballistic behavior, $\mathcal{S} \sim t^d$, dominated by finite-range hops.

We explore these predictions by numerically simulating open-system long-range interact-
ing RUCs for $\alpha = 0.7, 1.4$ (Fig. E.1). For $\alpha = 1.4$, consistent with previous results, we observe
that the average operator size grows super-ballistically under unitary dynamics (black). In
contrast, under open-system dynamics (red), we indeed find that the average operator size
eventually crosses over to sub-ballistic growth at late times. We observe qualitatively similar
results for $\alpha = 0.7$. Namely, the average operator size grows consistently with a shrunk expo-
nential at early times, and appears to plateau to a constant at late times. While we cannot
firmly establish the functional form of the late time behavior in either case, the crossover
from super-linear to sub-linear growth is consistent with the simple model outlined in the
previous paragraph.

## E.4 Open-system operator growth in free-fermion integrable spin chains

In this section, we address operator growth in free-fermion integrable spin chains (FF-ISCs) with open-system dynamics (Fig. E.2). Under unitary dynamics, FFISCs have been found to exhibit qualitatively different behavior of operator growth compared to chaotic spin chains [381, 504]. We will begin by showing that under unitary dynamics the size distributions of FFISCs are in fact *broad* ($\delta\mathcal{S} \sim \mathcal{S}$), unlike those of chaotic short-range 1D systems ($\delta\mathcal{S} \sim \sqrt{\mathcal{S}}$). We will then show that open-system dynamics eventually cause the average operator size of FFISCs to decrease in time, $\overline{\mathcal{S}} \sim 1/t$. In contrast to our other examples, we do not expect this behavior to be universal for all local error models. For instance, errors such as spontaneous emission will destroy the free fermion integrability of the model and are thus beyond the analytic framework we introduce below. Other errors, such as dephasing, will have a substantially weaker effect owing to the particular structure of time-evolved operators in FFISCs.

For concreteness, we focus on the XX model:

$$H = J \sum_{i=-\infty}^{\infty} (X_i X_{i+1} + Y_i Y_{i+1}) = J \sum_{i=-\infty}^{\infty} (r_i \ell_{i+1} + \ell_i r_{i+1}). \tag{E.38}$$

We expect the physics of other free-fermion integrable models to be similar. On the right side, we re-write the Pauli matrices, $X$ and $Y$, in terms of raising and lowering operators, $X_i = r_i + \ell_i$, $Y_i = i(r_i - \ell_i)$. We can solve this Hamiltonian by defining fermion operators via the Jordan-Wigner transform:

$$r_i = \left( \prod_{0 \le j < i} [2c_j^\dagger c_j - 1] \right) c_i^\dagger$$

$$\ell_i = \left( \prod_{0 \le j < i} [2c_j^\dagger c_j - 1] \right) c_i \tag{E.39}$$

$$Z_i = 2c_i^\dagger c_i - 1.$$

The Hamiltonian becomes:

$$H = J \sum_{i=-\infty}^{\infty} \left( c_i^\dagger c_{i+1} + c_i c_{i+1}^\dagger \right), \tag{E.40}$$

corresponding to hopping on a one-dimensional chain of free fermions.

*Operator growth under unitary dynamics.*—The Jordan-Wigner transformation allows us to exactly solve for the time-evolution of operators that map to local fermion bilinears. We focus on one such example: the fermion number operator, $Z_0 = 2c_0^\dagger c_0 - 1$ (taken to lie at site $i = 0$). Under unitary dynamics, the Heisenberg equation is:

$$\partial_t Z_0(t) = i[Z_0(t), H]. \tag{E.41}$$

Note that repeated commutations $Z_0$ with the Hamiltonian give rise to operators exclusively of the form $r_x Z_{x+1} \ldots Z_{x'-1} \ell_{x'}$, for $x' \geq x$, and $\ell_{x'} Z_{x'+1} \ldots Z_{x-1} r_x$, for $x \geq x'$. In the fermion picture, these correspond to bilinears, $c_x^\dagger c_{x'}$. Making this explicit, we define the Pauli strings,

$$S_{x,x'} \equiv c_x^\dagger c_{x'} = \begin{cases} -P_x Z_{x+1} \ldots Z_{x'-1} M_{x'} & x < x' \\ +P_x M_x & x = x' \\ -M_{x'} Z_{x'+1} \ldots Z_{x-1} P_x & x > x' \end{cases} . \tag{E.42}$$

We can decompose the time-evolved operator as a sum of such strings,

$$Z_0(t) = 2 \sum_{x,x'} a_{x,x'}(t) \left( S_{x,x'} - \delta_{x,x} \mathbb{1}/2 \right), \tag{E.43}$$

with coefficients $a_{x,x'}(t)$. Note that the identity factors guarantee that $\text{tr}(Z_0(t)) = 0$. The Heisenberg equation becomes:

$$\begin{aligned} \sum_{x,x'} \partial_t a_{x,x'}(t) \left( S_{x,x'} - \delta_{x,x} \mathbb{1}/2 \right) &= i \sum_{x,x'} a_{x,x'}(t) [S_{x,x'}, H] \\ &= iJ \sum_{x,x'} a_{x,x'}(t) [S_{x-1,x'} + S_{x+1,x'} - S_{x,x'-1} - S_{x,x'+1}] \end{aligned} \tag{E.44}$$

which gives the following time-evolution for the coefficients, $a_{x,x'}(t)$:

$$\partial_t a_{x,x'}(t) = iJ[a_{x-1,x'}(t) + a_{x+1,x'}(t) - a_{x,x'-1}(t) - a_{x,x'+1}(t)]. \tag{E.45}$$

This is the equation of motion for two (decoupled) free fermions, indexed by $x$ and $x'$ respectively. It can be solved exactly via Fourier transform.

We now present the solution to this equation in the continuum limit, $a_{x,x'}(t) \to a(x, x'; t)$. This limit neglects lattice-scale fluctuations in the coefficients $a_{x,x'}(t)$, but we expect it to broadly capture the operators size distribution in the bulk of the light cone (see Ref. [504] for a full treatment of unitary operator growth without the continuum approximation). The Heisenberg equation becomes:

$$-i\partial_t a(x, x'; t) = J\partial_x^2 a(x, x'; t) - J\partial_{x'}^2 a(x, x'; t). \tag{E.46}$$

We can solve this via a Fourier transform with the initial condition, $a(x, x'; 0) = \delta(x)\delta(x')$:

$$\begin{aligned} a(x, x'; t) &= \frac{1}{(2\pi)^2} \int dk \int dk' \, e^{-ikx} e^{-ik'x'} e^{-iJk^2 t} e^{iJk'^2 t} \\ &= \frac{1}{(2\pi)2Jt} e^{-ix^2/4Jt} e^{ix'^2/4Jt}. \end{aligned} \tag{E.47}$$

The solution has non-zero support at every site at any non-zero time, which is clearly unphysical. We can rectify this by smoothing the wavefunction over an initial Gaussian wavepacket

of width $\sim 1$ (i.e. the lattice spacing):

$$
\begin{aligned}
a_s(x, x'; t) &= \frac{1}{\sqrt{2\pi}} \int dy \int dy' \exp\left(-\frac{(x-y)^2}{2} - \frac{(x'-y')^2}{2}\right) a(y, y'; t) \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{(2\pi)2Jt} \int dy \int dy' \exp\left(-\frac{(x-y)^2}{2} - i\frac{y^2}{4Jt} - \frac{(x'-y')^2}{2} + i\frac{y'^2}{4Jt}\right) \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{(2\pi)2Jt} \int dy \int dy' \\
&\quad \times \exp\left(-\frac{y^2(1 + i/2Jt)}{2} - yx - \frac{x^2}{2} - \frac{y'^2(1 - i/2Jt)}{2} - y'x' - \frac{x'^2}{2}\right) \\
&= \frac{1}{\sqrt{2\pi}2Jt\sqrt{(1 + i/2Jt)(1 - i/2Jt)}} \\
&\quad \times \exp\left(-\frac{x^2}{2}\left[1 - \frac{1}{1 + i/2Jt}\right] - \frac{x'^2}{2}\left[1 - \frac{1}{1 - i/2Jt}\right]\right) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{(2Jt)^2 + 1}} \exp\left(-\frac{x^2}{2(1 - i2Jt)}\right) \exp\left(-\frac{x'^2}{2(1 + i2Jt)}\right).
\end{aligned}
$$

(E.48)

This gives a normalized probability distribution for $x, x'$,

$$
\begin{aligned}
|a_s(x, x'; t)|^2 &= \frac{1}{2\pi[(2Jt)^2 + 1]} \exp\left(-\frac{x^2}{2(1 - i2Jt)} - \frac{x^2}{2(1 + i2Jt)}\right) \\
&\quad \times \exp\left(-\frac{x'^2}{2(1 + i2Jt)} - \frac{x'^2}{2(1 - i2Jt)}\right) \\
&= \frac{1}{2\pi[(2Jt)^2 + 1]} \exp\left(-\frac{x^2/2}{1 + (2Jt)^2}\right) \exp\left(-\frac{x'^2/2}{1 + (2Jt)^2}\right) \\
&\approx \frac{1}{2\pi(2Jt)^2} \exp\left(-\frac{1}{2}(x/2Jt)^2\right) \exp\left(-\frac{1}{2}(x'/2Jt)^2\right),
\end{aligned}
$$

(E.49)

corresponding to a Gaussian centered about $x = x' = 0$, with a ballistically growing width, $2Jt$.

To compute the operator size distribution, note that the size superoperator acts on each string, $S_{x,x'}$, as:

$$
\mathcal{S}\{S_{x,x'}\} = (|x - x'| + 1)S_{x,x'} - \frac{1}{2}\mathbb{1}\delta_{x,x'}.
$$

(E.50)

This implies that $S_{x,x'} - \frac{1}{2}\mathbb{1}\delta_{x,x'}$ is a size eigenstate with size $(|x - x'| + 1)$. The size distribution of $Z_0(t)$ is therefore:

$$
P(\mathcal{S}) = \sum_{\{x,x':|x-x'|+1=\mathcal{S}\}} |a_{x,x'}(t)|^2 \approx \int dx \left(|a_s(x, x + \mathcal{S} - 1; t)|^2 + |a_s(x, x - \mathcal{S} + 1; t)|^2\right)
$$

(E.51)

Changing coordinates, $\delta H = x + x'$, $\delta = x - x'$, the operator wavefunction becomes:

$$|a_s(\delta H, \delta; t)|^2 \approx \frac{1}{2(2\pi)(2Jt)^2} \exp\left(-\frac{1}{4}(\delta H/2Jt)^2\right) \exp\left(-\frac{1}{4}(\delta/2Jt)^2\right), \quad (E.52)$$

which gives a size distribution:

$$
\begin{aligned}
P(\mathcal{S}) &\approx \int d\delta H \, |a_s(\delta H, \mathcal{S} - 1; t)|^2 + |a_s(\delta H, -\mathcal{S} + 1; t)|^2 \\
&= \frac{1}{2\sqrt{\pi}Jt} \exp\left(-\frac{1}{4}\left(\frac{\mathcal{S} - 1}{2Jt}\right)^2\right). \\
&\approx \frac{1}{2\sqrt{\pi}Jt} \exp\left(-\frac{1}{4}\left(\frac{\mathcal{S}}{2Jt}\right)^2\right).
\end{aligned}
\quad (E.53)
$$

The size distribution is the positive half of a Gaussian with mean zero and width $2\sqrt{2}Jt$. The average size thus grows ballistically,

$$\mathcal{S} \approx \frac{4Jt}{\sqrt{\pi}}, \quad (E.54)$$

and the size width likewise,

$$\delta\mathcal{S} \approx 2\sqrt{2}\sqrt{1 - 2/\pi}\, Jt. \quad (E.55)$$

This is in sharp contrast with chaotic spin chains, which exhibit diffusive growth of the size width, $\delta\mathcal{S} \sim \sqrt{\mathcal{S}}$.

*Operator growth under open-system dynamics.*—We now turn to the effect of open-system dynamics on operator growth. For simplicity, we study isotropic decoherence on each qubit, in which case the latter portion of the Lindbladian is directly given by the operator size. Again, in contrast to chaotic systems, we expect that different forms of open-system dynamics will affect FFISCs in different ways, since the large-size Pauli strings composing $\hat{Z}_0(t)$ contain nearly entirely Pauli-$Z$ components [Eq. (E.42)]. For example, this implies that dephasing dynamics will have a nearly negligible effect on time-evolution. On the other hand, open-system dynamics such as spontaneous emission will evolve $\hat{Z}_0(t)$ outside the given set of strings, breaking free fermion integrability.

The Heisenberg equation becomes:

$$\partial_t Z_0(t) = i[Z_0(t), H] - \varepsilon\mathcal{S}\{Z_0(t)\}. \quad (E.56)$$

Using the expansion in Eq. (E.43), we have:

$$
\begin{aligned}
\sum_{x,x'} \partial_t a_{x,x'}(t)\left(S_{x,x'} - \delta_{x,x}\mathbb{1}/2\right) &= \sum_{x,x'} a_{x,x'}(t)\left(i[S_{x,x'}, H] - \varepsilon\mathcal{S}\{S_{x,x'}\}\right) \\
&= \sum_{x,x'} a_{x,x'}(t)\Big[iJ[S_{x-1,x'} + S_{x+1,x'} - S_{x,x'-1} - S_{x,x'+1}] \\
&\quad - \varepsilon(|x - x'| + 1)\left(S_{x,x'} - \delta_{x,x}\mathbb{1}/2\right)\Big],
\end{aligned}
\quad (E.57)
$$
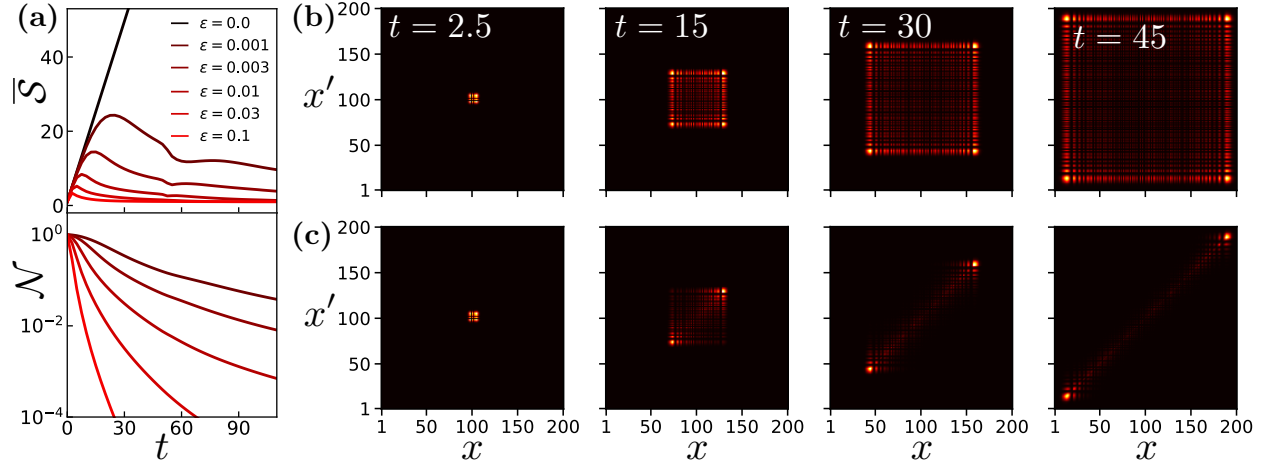
Figure E.2: Operator growth in free-fermion integrable spin chains with isotropic single-qubit decoherence [Eq. (E.58)]. **(a)** The average operator size increases linearly at early times, and before inverting to decay towards one at a time-scale set by the error rate, $\varepsilon$. The Loschmidt echo fidelity (i.e. the operator normalization) displays non-concave behavior as a result of this reverse in operator growth. **(b)** The probability distribution, $a_{x,x'}(t)$, for the location of the endpoints, $x, x'$, of a time-evolved operators' Pauli strings under unitary dynamics. Each endpoint spreads ballistically in either direction, and the interior of the distribution is approximately uniform. **(c)** The same probability distribution under open-system dynamics with $\varepsilon = 0.003$. The small-size components (which lie along the diagonal, $x \approx x'$) are nearly unaffected by error, while large-size components (which lie highly off-diagonal) are strongly damped. Note that the time of the final image, $t = 45$, occurs shortly before the first kink appearing in **(a)**.

which gives the following time-evolution for the Pauli string coefficients:

$$\partial_t a_{x,x'}(t) = iJ[a_{x-1,x'}(t) + a_{x+1,x'}(t) - a_{x,x'-1}(t) - a_{x,x'+1}(t)] - \varepsilon(|x - x'| + 1)a_{x,x'}(t). \quad \text{(E.58)}$$

Note that the coordinates, $x$ and $x'$, are now coupled by the Lindbladian, since the error is proportional the absolute value of the separation between the two coordinates (i.e. the size).

   To approximately solve this equation, first recall that in the fermion picture for unitary time-evolution, the growth of $Z_0(t)$ corresponds a fermionic creation and annihilation operator each dispersing independently. At large times, $t \gg 1/J$, each momentum-component of each fermion will propagate independently of all other momenta (formally, this corresponds to a saddle point approximation when solving Eq. (E.44) via a Fourier transform, see Ref. [504]). For a given momentum $k$, the group velocity, $v(k) = \partial_k \epsilon(k)$, is determined by the dispersion relation, $\epsilon(k)$, of the Hamiltonian (for the XX model, we have $\epsilon(k) = 2J \cos(k) \approx 2J + Jk^2$ in the continuum limit). Turning to Lindbladian time-evolution, we now consider a pair of momenta, $k, k'$, for the creation and annihilation operator respectively. The size of the Pauli

strings corresponding to these momenta grows linearly in time, according to:

$$\mathcal{S}_{k,k'}(t) = |v(k) - v(k')|t. \tag{E.59}$$

We therefore expect normalization of the particular pair of momenta to decay by a factor,

$$\sim \exp\left(-2\varepsilon \int_0^t dt' \mathcal{S}_{k,k'}(t)\right) = \exp\left(-\varepsilon t \mathcal{S}_{k,k'}(t)\right). \tag{E.60}$$

Integrating over all momenta pairs gives a size distribution,

$$P(\mathcal{S}) \approx \frac{1}{2\sqrt{\pi}Jt} \exp\left(-\frac{1}{4}\left(\frac{\mathcal{S}}{2Jt}\right)^2 - \varepsilon\mathcal{S}t\right). \tag{E.61}$$

We expect this treatment to hold when $\varepsilon/J \ll 1$, which guarantees that the effect of error, $\sim \varepsilon Jt^2$, will be negligible until sufficiently large times ($t \gg 1/J$) that our approximation holds.

The normalization of the operator size distribution is:

$$\mathcal{N} = \int_0^\infty d\mathcal{S}\, P(\mathcal{S}) = \exp\left(4J^2\varepsilon^2 t^4\right) \cdot \text{erfc}\left(2J\varepsilon t^2\right), \tag{E.62}$$

where the complementary error function is defined as, $\text{erfc}(x) \equiv (2/\sqrt{\pi})\int_x^\infty e^{-x^2}$. The average size can also be computed exactly in terms of the error function, although we do not express it here. At late times ($J\varepsilon t^2 \gg 1$), the size distribution asymptotes to,

$$P(\mathcal{S}) \to \frac{1}{2\sqrt{\pi}Jt} \exp\left(-\varepsilon\mathcal{S}t\right), \tag{E.63}$$

which gives an algebraic decay for the normalization,

$$\mathcal{N} \to \frac{1}{2\sqrt{\pi}J\varepsilon t^2}. \tag{E.64}$$

This can be derived from Eq. (E.62) using the approximation, $\text{erfc}(x) \approx e^{-x^2}/(\sqrt{\pi}x)$, for the complementary error function at large $x$. This normalization decay implies a $\sim 1/t$ decay of the average operator size,

$$\mathcal{S} = \frac{-\partial_t \log(\mathcal{N})}{2\varepsilon} \to \frac{1}{\varepsilon t}, \tag{E.65}$$

as quoted in Table 10.1.

To explore open-system operator growth in FFISCs outside the continuum limit, we numerically solve Eq. (E.58) for $N = 200$ with open boundary conditions. As shown in Fig. E.2, we observe that the average operator size indeed increases linearly at early times and decays towards its minimal value, one, at late times. We also observe "kinks" in the time profile of the average operator size. By analyzing the full Pauli string distribution (Fig. E.2), $a_{x,x'}(t)$, we see that these kinks occur precisely when fermion operators reflect off the boundary of the finite-size system.

# E.5 Numerical simulation details

In this section we provide details on the numerical simulations presented in Figs. 10.2, 10.3, E.1, and E.2.

*Haar-random unitary circuits.*—We begin with the random unitary circuit (RUC) simulations presented in Fig. 10.2 and Fig. E.1. Haar-random unitary circuits are toy models for operator spreading, in which averages of quantities that contain two or fewer copies of a time-evolved operator are efficiently simulable by averaging instead over Clifford-random unitaries [122, 171, 340, 341]. Note that the generating function for the size distribution (Fig. 10.4) contains two copies of $\hat{M}(t)$ can thus be efficiently simulated. The zeroth and first moments of the size distribution provide the Loschmidt echo fidelity and the average operator size. The RUC formalism can be also adapted to incorporate *open-system* operator spreading [326]. Below, we do so for the specific error model in Eq. (10.4). As discussed earlier in the supplemental material, this model exactly captures the effect of single-qubit Pauli errors, and also applies to leading order in the error rate for arbitrary single-qubit errors (assuming in both cases that errors occur at the same rate at each qubit, although the model is easily modified if this is not the case).

We consider three RUC models: (i) 1D short-range interacting [340], (ii) 1D long-range interacting [97, 531], and (iii) 0D all-to-all interacting [429]. In each model, a single time step, $\delta t$, consists of choosing a pair of qubits and acting on this pair with a Haar-random unitary, and applying the global quantum channel, $\exp(-\varepsilon \mathcal{S} \delta t)$. A unit of physical time consists of $N$ iterations of this procedure, so that each qubit is acted on by two unitaries per unit time, on average. The pairs are chosen as follows: (i) in the short-range 1D RUC, a qubit $j$ is sampled randomly and paired with its nearest neighbor, $j + 1$, (ii) in the long-range 1D RUC, a pair of qubits $i$ and $j$ are sampled according to a power law distribution in their distance, $\sim 1/|i - j|^{2\alpha}$ [97, 531], and (iii) in the 0D RUC, pairs are sampled entirely randomly. In cases (i) and (ii) we take periodic boundary conditions.

We simulate the average generating function of the operator size distribution by tracking the Clifford time-evolution of a local Pauli operator of interest, and sampling over realizations of Clifford-random unitaries [340, 341]. In each Clifford realization, $\mathcal{C}$, the time-evolved operator, $\hat{M}_{\mathcal{C}}(t)$, has a well-defined size at every time step (since Pauli operators are eigenstates of the size superoperator). Under open-system dynamics, the individual Pauli operator $\hat{M}_{\mathcal{C}}(t)$ thus evolves exactly as:

$$\hat{M}_{\mathcal{C}}(t) = \exp\left(-\varepsilon \int_0^t dt' \mathcal{S}_{\mathcal{C}}(t')\right) \cdot \hat{M}_{\mathcal{C},U}(t). \tag{E.66}$$

Here, $\hat{M}_{\mathcal{C},U}(t)$ represents the operator time-evolved by Clifford unitaries under *unitary* dynamics (i.e. $\varepsilon = 0$), and $\mathcal{S}_{\mathcal{C}}(t')$ is the size of this operator at time $t'$. The average generating function of the size distribution over Haar-random realizations, $\mathcal{U}$, is therefore equal to:

$$G_{\mathcal{S}}(\mu) = \frac{1}{N_{\mathcal{C}}} \sum_{\mathcal{C}} \left\langle \hat{M}_{\mathcal{C}}(t) e^{-\mu \mathcal{S}} \left\{ \hat{M}_{\mathcal{C}}(t) \right\} \right\rangle = \frac{1}{N_{\mathcal{C}}} \sum_{\mathcal{C}} \mathcal{N}_{\mathcal{C}}(t) e^{-\mu \mathcal{S}_{\mathcal{C}}(t)}, \tag{E.67}$$

where we denote the normalization of a single Clifford realization as:

$$\mathcal{N}_{\mathcal{C}}(t) = \exp\left(-2\varepsilon \int_0^t dt' \mathcal{S}_{\mathcal{C}}(t')\right). \tag{E.68}$$

The average generating function is a sum of the generating functions over each Clifford realization (i.e. the simple exponentials $e^{-\mu \mathcal{S}_{\mathcal{C}}(t)}$), weighted by its normalization, $\mathcal{N}_{\mathcal{C}}(t)$. From Eq. (E.67), we obtain the average Loschmidt echo fidelity,

$$\mathcal{N}_t = \frac{1}{N_{\mathcal{C}}} \sum_{\mathcal{C}} \mathcal{N}_{\mathcal{C}}(t), \tag{E.69}$$

and the moments of the size distribution,

$$\overline{\mathcal{S}^n} = \frac{\frac{1}{N_{\mathcal{C}}} \sum_{\mathcal{C}} \mathcal{N}_{\mathcal{C}}(t) \left(\mathcal{S}^{\mathcal{C}}(t)\right)^n}{\mathcal{N}_t}. \tag{E.70}$$

In Fig. 10.2, we plot our phenomenological predictions as dashed lines in addition to data from numerical simulations. For Fig. 10.2(a) [i.e. model (i)], the dashed curve is given by $\overline{\mathcal{S}} = 2v_B t - \varepsilon c t^2/2$, with the butterfly velocity, $v_B = (3/5)(3/4) = 9/20$, set by theory [340], and the size width parameter, $c = 1.14$, fit to the average operator size data via linear regression. For Fig. 10.2(b) [i.e. model (iii)], the dashed curve is given by $\overline{\mathcal{S}} = \frac{e^{\lambda t}}{1 + e^{\lambda t}/\mathcal{S}_p}$, with the value of the plateau size, $\mathcal{S}_p = \frac{\lambda}{b^2 \varepsilon + 4\lambda/3N}$, and the Lyapunov exponent, $\lambda = 2(3/5) = 6/5$, are set by theory, and the size width parameter, $b = 1.00$, is fit to the average operator size via linear regression. This is the solution to the phenomenological size growth equation, $\partial_t \overline{\mathcal{S}} = \lambda \overline{\mathcal{S}} - (\varepsilon b^2 + \frac{4\lambda}{3N})\overline{\mathcal{S}}^2$, obtained from the equation in Ch. 10 by setting $\delta \mathcal{S} = b\mathcal{S}$ and including a $1/N$ correction to ensure that under unitary dynamics the average operator size saturates to the late time value, $3N/4$.

*Hamiltonian dynamics.*—In Fig. 10.3, we compute the OTOCs of local operators in the scenario in which forwards and backwards time-evolution are both performed unitarily, but under unequal Hamiltonians. Specifically, we consider the following spin chain Hamiltonians:

$$\hat{H}_1 = \hat{H} - \frac{\eta}{2}\delta\hat{H}, \quad \hat{H}_2 = \hat{H} + \frac{\eta}{2}\delta\hat{H}. \tag{E.71}$$

Here, the Hamiltonian $\hat{H}$ is given by:

$$\hat{H} = \sum_{i=1}^{N} h_z \hat{Z}_i + \sum_{i=1}^{N-1} \left( J_x \hat{X}_i \hat{X}_{i+1} + J_y \hat{Y}_i \hat{Y}_{i+1} + J_z \hat{Z}_i \hat{Z}_{i+1} \right)$$

$$\times + \sum_{i=1}^{N-2} \left( J'_x \hat{X}_i \hat{X}_{i+2} + J'_y \hat{Y}_i \hat{Y}_{i+2} + J'_z \hat{Z}_i \hat{Z}_{i+2} \right), \tag{E.72}$$

which consists of single-site fields, $h_z = 0.4$, strong nearest-neighbor interactions, $J_x = 0.6, J_y = -0.8, J_z = 1.1$, and weak next-nearest-neighbor interactions, $J'_x = -0.3, J'_y = 0.3, J'_z = 0.15$. Here, we choose "generic" values for the interaction strengths in order to avoid behavior associated with closeness to integrable points (for example, the Heisenberg point at $J_\alpha = J, J'_\alpha = 0$). We do not otherwise expect our results to depend on the specific values chosen. The perturbing Hamiltonian,

$$\eta \, \delta\hat{H} = \sum_i \left( h_{x,i} \hat{X}_i + h_{y,i} \hat{Y}_i \right), \tag{E.73}$$

consists of disordered single-site fields in the perpendicular directions, with each magnitude drawn randomly from the interval $[-\eta, \eta]$ with $\eta = 0.4$. This choice of Hamiltonians possesses one special feature, namely that the perturbing Hamiltonian has zero overlap with all powers of the average Hamiltonian, $\text{tr}\left(\delta\hat{H}\,\hat{H}^m\right) = 0$ for all $m$. As discussed earlier in the supplemental material, we expect errors with non-zero overlap to contribute subleading corrections (suppressed by powers of $\sim 1/t$) to our picture of operator spreading.

We numerically compute the Loschmidt echo fidelity,

$$\mathcal{N}(t) = \langle e^{-iH_1 t} \hat{M} e^{iH_1 t} e^{-iH_2 t} \hat{M} e^{iH_2 t} \rangle, \tag{E.74}$$

and the OTOCs,

$$C(t; \hat{P}_i) = \langle e^{-iH_1 t} \hat{M} e^{iH_1 t} \hat{P}_i e^{-iH_2 t} \hat{M} e^{iH_2 t} \hat{P}_i \rangle \tag{E.75}$$

of the above Hamiltonians via Krylov-subspace methods for $N = 12$ and $10^4$ disorder realizations. Note that Loschmidt echo fidelity, $\mathcal{N}(t)$, is not guaranteed to be positive in this scenario (although we expect it to be in the many-body limit). Indeed, at late times where $\mathcal{N}(t)$ is expected to be exponentially small, we observe that the value of $\mathcal{N}(t)$ for individual disorder realizations may take small negative values. We find that the average of $\mathcal{N}(t)$ over disorder realizations is significantly more stable and remains positive for all times considered.

In Fig. 10.3, we plot the average of the OTOC over both disorder realizations and the set of single-qubit Pauli operators, $\hat{P}_i \in \{\mathbb{1}_i, X_i, Y_i, Z_i\}$, divided by the average Loschmidt echo fidelity over disorder realizations. We comment that, if the experiments we wish to model contain only a single realization of the perturbing Hamiltonian, it would be more appropriate to divide the OTOC by the Loschmidt echo fidelity *before* averaging over disorder realizations. However, we find that this leads to numerical instabilities at the small system sizes we consider. In Fig. 10.3(a) we take the initial operator $\hat{M} = \hat{X}_0$, which does not overlap the Hamiltonian, while in Fig. 10.3(b) we take the initial operator $\hat{M} = \hat{Z}_0$, which has non-zero overlap.

*Free fermions integrable spin chains.*—Time-evolution for Fig. E.2 is performed by exactly solving Eq. (E.58) via matrix exponentiation, with open boundary condition for $J = 1$ and $N = 200$.

# Appendix F

# Details on scrambling and complexity in phase space

In this Appendix we provide additional theoretical and numerical details relating to the results of Ch. 11.

## F.1 Basic Gaussian unitaries

In this appendix, we give some examples of symplectic matrices corresponding to Gaussian unitries [485].

1. Single-mode phase rotation (2 by 2)

$$\boldsymbol{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{F.1}$$

2. Single-mode squeezing (2 by 2)

$$\boldsymbol{S}(r) = \begin{pmatrix} e^{-r} & 0 \\ 0 & e^{r} \end{pmatrix}. \tag{F.2}$$

3. Beam-splitter transform (4 by 4) written in a block form

$$\boldsymbol{B}(\eta) = \begin{pmatrix} \sqrt{\eta}\mathbf{I} & \sqrt{1-\eta}\mathbf{I} \\ -\sqrt{1-\eta}\mathbf{I} & \sqrt{\eta}\mathbf{I} \end{pmatrix}, \tag{F.3}$$

where $\mathbf{I}$ is the 2 by 2 identity matrix.

4. Two-mode squeezing (4 by 4) written in a block form

$$\boldsymbol{S}_2(r) = \begin{pmatrix} \cosh r\mathbf{I} & \sinh r\mathbf{Z} \\ \sinh r\mathbf{Z} & \cosh r\mathbf{I} \end{pmatrix}, \tag{F.4}$$

where $\mathbf{Z}$ is the 2 by 2 Pauli Z matrix.

In the following, we formalize the analog between displacement operators and Pauli operators, and between Gaussian unitaries and Clifford unitaries [472].

**Lemma 1.** *Completeness of displacement operators, $tr\left(D\left(\boldsymbol{\xi}\right)D\left(\boldsymbol{\xi}'\right)\right) = \pi^N\delta\left(\boldsymbol{\xi}+\boldsymbol{\xi}'\right)$, and $A = 1/\pi^N\int d^{2N}\boldsymbol{\xi}\;\chi\left(\boldsymbol{\xi};A\right)D\left(-\boldsymbol{\xi}\right),$ when $\chi\left(\boldsymbol{\xi};A\right)$ exists. Moreover, we have two useful identities*

$$\frac{1}{\pi^N}\int d^{2N}\boldsymbol{\xi}\;D\left(\boldsymbol{\xi}\right)AD^\dagger\left(\boldsymbol{\xi}\right) = tr\left(A\right)\boldsymbol{I}, \tag{F.5}$$

$$\frac{1}{\pi^N}\int d^{2N}\boldsymbol{\xi}\;tr\left[D\left(\boldsymbol{\xi}\right)A\right]tr\left[D^\dagger\left(\boldsymbol{\xi}\right)B\right] = tr\left(AB\right). \tag{F.6}$$

The proof can be found in Ref. [84].

**Lemma 2.** *For Gaussian unitary,*

$$U_{\boldsymbol{S},\boldsymbol{d}}^\dagger D\left(\boldsymbol{\xi}\right)U_{\boldsymbol{S},\boldsymbol{d}} = \exp\left(i\boldsymbol{d}^T\boldsymbol{\Omega}\boldsymbol{\xi}\right)D\left(\boldsymbol{S}^{-1}\boldsymbol{\xi}\right). \tag{F.7}$$

The proof is straightforward calculation.

**Lemma 3.** *Gaussian unitary corresponds to a linear coordinate transform of the Wigner characteristic function and Wigner function,*

$$\chi\left(\boldsymbol{\xi};U_{\boldsymbol{S},\boldsymbol{d}}AU_{\boldsymbol{S},\boldsymbol{d}}^\dagger\right) = \chi\left(\boldsymbol{S}^{-1}\boldsymbol{\xi};A\right)\exp\left(i\boldsymbol{d}^T\boldsymbol{\Omega}\boldsymbol{\xi}\right) \tag{F.8}$$

*and*

$$W\left(\boldsymbol{x};U_{\boldsymbol{S},\boldsymbol{d}}AU_{\boldsymbol{S},\boldsymbol{d}}^\dagger\right) = W\left(\boldsymbol{S}^{-1}\left(\boldsymbol{x}-\boldsymbol{d}\right);A\right). \tag{F.9}$$

The proof for Lemma 3 can also be found in Ref. [454]

The above analog motivates us to consider displacement operators as a 1-design and Gaussian unitaries as a 2-design analog.

## F.2   Operator distributions and OTOCs in quasi scramblers

### Volume and average OTOCs

Consider some ensemble $\mathcal{E}$ of displacement operators. Under a quasi scrambling dynamics by a Gaussian unitary $U(t)$, the ensemble evolves to $\mathcal{E}(t)$ which is also an ensemble of displacement operators. Our goal is to develop a probe of $\mathcal{E}(t)$ via OTOCs. For this purpose, let us introduce a formal definition of the average OTOCs for quasi scramblers. Given a pair of ensembles of displacement operators $\mathcal{E}_1$ and $\mathcal{E}_2$, define

$$\overline{\mathcal{C}_2}\left(\mathcal{E}_1,\mathcal{E}_2\right)_\rho \equiv \mathbb{E}_{V\sim\mathcal{E}_1,W\sim\mathcal{E}_2}\mathrm{tr}\left[\rho V^\dagger W^\dagger VW\right]. \tag{F.10}$$

For the purpose of characterizing operator distributions, $\mathcal{E}_1$ may be regarded as the distribution of our interest, e.g., $\mathcal{E}_1 = \mathcal{E}(t)$, while $\mathcal{E}_2$ is a probe. The necessity of taking an average over $\mathcal{E}_1$ and $\mathcal{E}_2$ can be understood from the fact that the amplitude of OTOCs for Gaussian unitaries is always unity.

While the average OTOCs of quasi scramblers in Eq. (F.10) bears great similarity with the one of genuine scramblers in Eq. (11.38), there is a subtle difference. In the case of genuine scramblers, $D^\dagger(\boldsymbol{\xi}_1; t)$ and $D(\boldsymbol{\xi}_1; t)$ are decomposed by different displacement operators with weights $\chi^\star[\boldsymbol{\xi}]$ and $\chi[\boldsymbol{\xi}']$. Observe that, due to $\rho$, there are contributions for $\boldsymbol{\xi} \neq \boldsymbol{\xi}'$ to the average OTOCs. This is not the case for quasi scramblers where we sample the same displacement operator for $U^\dagger$ and $U$. Namely, the average OTOCs for quasi scramblers do not depend on $\rho$. We note that the $\rho$ dependence may be recovered by considering thermally regulated OTOCs, but detailed discussions on these generalizations of OTOCs are beyond the scope of this paper.

As in the case with genuine scramblers, the average OTOC is closely related to operator distributions. For instance, given an arbitrary ensemble of dispalcement operators $\mathbb{D}_{P(\cdot)}$ with the distribution $P(\cdot)$, let us consider cases where the probe $\mathcal{E}_2$ is the global Gaussian displacement ensemble $\mathbb{D}_n$ defined in Eq. (11.39):

$$\overline{\mathcal{C}_2}\left(\mathbb{D}_n, \mathbb{D}_{P(\cdot)}\right)_\rho = \mathbb{E}_{V \sim \mathbb{D}_n, W \sim \mathbb{D}_{P(\cdot)}} \mathrm{tr}\left[\rho V^\dagger W^\dagger V W\right] \tag{F.11}$$

$$= \mathbb{E}_{\boldsymbol{\xi} \sim P(\cdot)}\left[\exp\left(-n|\boldsymbol{\xi}|^2\right)\right]. \tag{F.12}$$

This measures the extent of the operator spreading with Gaussian filters.

In fact, an explicit relation between OTOCs and frame potentials can be derived when distributions of our interest are Gaussian distributions with zero mean. As an example, we look at the zero mean ensemble $\mathbb{D}_{\boldsymbol{0},\boldsymbol{V}}$ in Eq. (11.22). One can show that

$$\overline{\mathcal{C}_2}\left(\mathbb{D}_{\boldsymbol{0},\boldsymbol{V}}, \mathbb{D}_n\right)_\rho = \prod_{\ell=1}^{2N} \frac{1}{\sqrt{1 + 2\lambda_\ell n}}, \tag{F.13}$$

where $\lambda_\ell$'s are the eigenvalues of the matrix $\boldsymbol{V}$. Note that for displacement operators, the OTOC is independent of the state $\rho$. In particular, if $n = 2\left(2n_{\mathrm{th}} + 1\right)$, comparing with Eq. (11.23) we have

$$\overline{\mathcal{C}_2}\left(\mathbb{D}_{\boldsymbol{0},\boldsymbol{V}}, \mathbb{D}_n\right)_\rho = \mathcal{F}_{\mathbb{D}_{\boldsymbol{\xi}_0,\boldsymbol{V}}}\left(\tilde{\rho}_{n_{\mathrm{th}}}\right). \tag{F.14}$$

An important merit of OTOCs over the frame potential is that OTOCs can measure not only the ensemble volume, but also the projected volume on certain subspace of the $2N-$dimensional phase space (Fig. 11.6). If one chooses one of the ensembles as displacement operators localized on certain mode $w$, i.e. $\mathbb{D}_n^w = \left\{D^w\left(\boldsymbol{\xi}\right) | \boldsymbol{\xi} \sim P_D^G\left(\boldsymbol{\xi}; n\right) \equiv \exp\left(-|\boldsymbol{\xi}|^2/n\right)/\left(\pi n\right)\right\}$, we have

$$\overline{\mathcal{C}_2}\left(\mathbb{D}_{\boldsymbol{0},\boldsymbol{V}}, \mathbb{D}_n^w\right)_\rho = \frac{1}{\sqrt{\left(1 + 2\lambda_1^w n\right)\left(1 + 2\lambda_2^w n\right)}}, \tag{F.15}$$

where $\lambda_1^w, \lambda_2^w$ are the eigenvalues of the marginal covariance matrix $\boldsymbol{V}_{2w,2w+1}$. This enables us to define the projected volume on mode $w$ as

$$\mathrm{vol}^w(\mathcal{E}) = \lim_{n \to \infty} \left(\frac{1}{2n}\right) \frac{1}{\overline{\mathcal{C}_2}(\mathcal{E}, \mathbb{D}_n^w)_\rho}, \tag{F.16}$$

in analog to Eqs. (11.20) and (11.13).

It is worth looking at a few simple examples. We shall consider the average OTOCs of the form

$$\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho \equiv \overline{\mathcal{C}_2}(\mathbb{D}_n^\omega(t), \mathbb{D}_n^v), \tag{F.17}$$

where displacement operators are chosen to be Gaussian distributed at $t = 0$, and $\mathbb{D}_n^\omega$ evolves to $\mathbb{D}_n^\omega(t)$ under a Gaussian unitary $U$. Note that in Eq. (F.16), the limit is only on the probe ensemble, which is not taking $n$ to infinity in Eq. (F.17). First we look at single-mode cases $w = v$.

i) For $U$ being displacements and phase rotations, we have

$$\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho = \frac{1}{1 + n^2}. \tag{F.18}$$

This includes the case where $U$ is an identity operator.

ii) For single-mode squeezing, we have $\alpha' = e^{-r}\mathrm{Re}\alpha + e^r\mathrm{Im}\alpha$, and

$$\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho = \frac{1}{\sqrt{1 + n^4 + 2n^2\cosh(2r)}} \tag{F.19}$$

where the average OTOC decays as the squeezing parameter $r$ increases. This is an example of the increase of coarse-grained volume measured by finite temperature average OTOCs.

Next we look at multi-mode cases.

iii) Passive linear optics. Let $\eta_{w,v}$ be the transmissivity between the two modes $w, v$. We have

$$\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho = \frac{1}{(1 + \eta_{w,v}^2 n^2)}. \tag{F.20}$$

When $\eta_{w,v} = 1$, it corresponds to swap gate and $\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho = 1/(1 + n^2)$. For generic passive linear optics, a typical transmissivity between the two modes $\eta_{w,v} \sim 1/\sqrt{N}$ decreases with the system size, thus the decay of the average OTOCs becomes less significant.

iv) For a two-mode squeezing between modes $w, v$, we find

$$\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho = 1/(1 + \sinh^2(r)n^2). \tag{F.21}$$

We see that as the amount of squeezing $|r|$ increases, $\overline{\mathcal{C}_2^{w,v}}(n; t)_\rho \sim e^{-2|r|}$ decreases exponentially.

## Quantum Liouville's theorem

Genuine scrambling via non-Gaussian unitaries is of quantum nature whereas quasi scrambling via Gaussian unitaries admits description by a classical chaotic Hamiltonian dynamics since a single phase space point stays localized. This observation motivates us to look for a generalization of the classical Liouville's theorem applicable to the Gaussian quantum dynamics.

In classical mechanics, Liouville's theorem asserts that the phase space volume is preserved in Hamiltonian dynamics. Namely, a phase space volume may stretch and distort to complicated forms, yet its volume is always conserved. In analog to the classical theorem, let us consider an arbitrary displacement ensemble $\mathbb{D}_{P(.)}$ with the distribution $P(\cdot)$. We are interested in the volume of the ensemble $\mathbb{D}_{P(.)}(t)$ at time $t$. To be more rigorous, we shall consider the volume of operator distributions as defined in Eq. (11.20). Noticing that volume is given by an integration in Eq. (11.17) and $\det(\boldsymbol{S}) = 1$, we immediately have the following theorem.

**Theorem 3.** *(Quantum Liouville's theorem) Phase space volume of a general ensemble of displacement operators is preserved if $U(t)$ is a Gaussian unitary.*

Despite its simplicity, the theorem provides interesting lessons. First, since the theorem suggests no volume increase for Gaussian unitaries, one might wonder why Gaussian quasi scramblers can transform a displacement operator into the one with larger displacement. Indeed, given the fact that OTOCs are sensitive to volume growths, the decay of the average OTOCs appear to contradict with the theorem. However, recall that the average OTOCs with respect to local modes measure projected volumes. Namely, even if the whole volume in $2N$-dimensional space is fixed, the projected volumes on local modes can be large in the presence of large squeezing (Fig. 11.6). Second, the theorem holds at the infinite temperature limit whereas coarse-grained volumes at finite temperature may not be preserved, suggesting the emergence of quantum nature at low temperature phases. This coarse-grained volume also has a nice correspondence to the idea of Kolmogorov-Sinai (KS) entropy. In classical Hamiltonian dynamics, the phase space volume is preserved. The KS entropy counts the number of coarse-grained phase space boxes of the phase space volume, which increases for generic chaotic dynamics.

Another interesting implication concerns characterization of quantum chaos in CV systems. The theorem implies that any non-Gaussian effect will lead to some change of the volume. Although the volume may either decrease or increase depending on $U$ and initial conditions, we expect that the volume can only increase for generic choice of chaotic non-Gaussian dynamics. This speculation suggests that the volume of operator distributions is a monotonically increasing quantity, providing an intuition similar to the second-law of entropies.

## F.3 OTOC in presence of loss

In this section, we demonstrate that OTOCs decay due to a loss channel. Consider the original unitary channel $\rho \to U^\dagger(t)\rho U(t)$ combined with a thermal loss channel. Let's suppose the loss happens before the unitary, then

$$D^w(\alpha;t) = U(t)^\dagger \mathcal{N}_\eta^{N_E}(D^w(\alpha))U(t),\qquad\text{(F.22)}$$

the channel maps $x \to \sqrt{\eta}x + \sqrt{1-\eta}x_e$, where $x_e$ are quadrature operators for an ancilla in thermal state $\rho_E$ with mean photon number $N_E$. To obtain how operators evolve under the loss channel we consider the trace

$$
\begin{aligned}
&\mathrm{tr}_A\left(D_A(\alpha)\mathcal{N}_\eta^{N_E}(\rho_A)\right)\\
&= \mathrm{tr}_A\left(D_A(\alpha)\,\mathrm{tr}_E U_{AE,\eta}(\rho_A \otimes \rho_E)U_{AE,\eta}^\dagger\right)\\
&= \mathrm{tr}_E\mathrm{tr}_A\left(U_{AE,\eta}^\dagger D_A(\alpha)U_{AE,\eta}\rho_A \otimes \rho_E\right)\\
&= \mathrm{tr}_E\mathrm{tr}_A\left(D_A(\sqrt{\eta}\alpha)D_E\left(\sqrt{1-\eta}\alpha\right)\rho_A \otimes \rho_E\right)\\
&= \chi\left(\mathrm{Re}\sqrt{1-\eta}\alpha,\mathrm{Im}\sqrt{1-\eta}\alpha;\rho_E\right)\mathrm{tr}_A\left(D_A(\sqrt{\eta}\alpha)\rho_A\right)\\
&= \exp\left(-(1-\eta)|\alpha|^2(2N_E+1)/2\right)\mathrm{tr}_A\left(D_A(\sqrt{\eta}\alpha)\rho_A\right),\qquad\text{(F.23)}
\end{aligned}
$$

where we have used the Stinespring dilation of a pure loss channel. Thus for a single loss channel

$$\mathcal{N}_\eta^{N_E}(D(\alpha)) = \exp\left(-(1-\eta)|\alpha|^2(2N_E+1)/2\right)D(\sqrt{\eta}\alpha).\qquad\text{(F.24)}$$

From this we expect, the OTOC decreases exponentially in presence of loss and noise. The actual decay relies on the measurement scheme of the OTOC, but a rule-of-thumb estimation by considering two lossy noisy channel gives $\mathcal{C}_2^{w,v}(\alpha,\beta;t)_T|_{\eta,N_E} \sim \exp\left(-(1-\eta)|\alpha|^2(2N_E+1)\right)$.

## F.4 Twice-regulated frame potential

In the main text, we have employed the finite-temperature frame potential to characterize operator distributions in an ensemble of CV unitaries. Despite the fact that it provides us with a quantitative understanding of volumes of distributions, there are two drawbacks. First, it is lower bounded by zero even though $\rho$ is a quantum state with finite entropy. This implies that any normalized probability distribution won't saturate the lower bound while our Hilbert space of interest is regulated by $\rho$. This concern is also related to the second drawback. When we seek for appropriate definitions of CV design, it would be natural to consider ensembles of unitaries which approximately preserve the density matrix $\rho$. Hence we hope to construct a modification of frame potentials which is lower bounded by some $\rho$-dependent finite value.

We now outline the derivation of the new frame potential. The guiding principle is the definition (11.67) of DV $k$-designs. For CV systems with $d \to \infty$, it converges to the following:

$$\lim_{d \to \infty} d^k \, \mathbb{E}_{\text{Haar}} \left\{ U^{\otimes k} \otimes (U^\dagger)^{\otimes k} \right\} = \sum_\pi S_\leftrightarrow (W_\pi \otimes W_\pi^{-1}), \tag{F.25}$$

where subleading contributions from the Weingarten functions are neglected. We will use this as a definition for Haar randomness for CV systems. Here $W_\pi$ performs the permutation $\pi$ on the $k$-copied Hilbert space $\mathcal{H}^{\otimes k}$ and $S_\leftrightarrow$ is the swap operator between the two copies of $\mathcal{H}^{\otimes k}$ (one acted on by $U^{\otimes k}$ and one acted on by $(U^\dagger)^{\otimes k}$).

The new frame potential seeks to characterize how closely a given ensemble $\mathcal{E}$ resemble the Haar ensemble. Define the followings:

$$|\mathcal{H}) \equiv \mathbb{E}_{\text{Haar}} \left\{ U^{\otimes k} \otimes (U^\dagger)^{\otimes k} \right\} \tag{F.26}$$

$$|\mathcal{E}) \equiv \mathcal{N} \, \mathbb{E}_{\mathcal{E}} \left\{ U^{\otimes k} \otimes (U^\dagger)^{\otimes k} \right\}, \tag{F.27}$$

with some appropriate normalization constant $\mathcal{N}$ which will be specified later. We define the inner product between operators $\mathcal{O}_1, \mathcal{O}_2$ on the $2k$-copied Hilbert space $\mathcal{H}^{\otimes 2k}$ as

$$(\mathcal{O}_1 | \mathcal{O}_2) = \text{tr} \left( \mathcal{O}_1^\dagger \, P_k \, \mathcal{O}_2 \, P_k \right), \tag{F.28}$$

with positive Hermitian operators $P_k = [\left( \rho^{\frac{1}{2k}} \right)^{\otimes k} \otimes \left( \rho^{\frac{1}{2k}} \right)^{\otimes k}]$. One can show that it satisfies all properties of inner product if $\rho$ is full rank. If $\rho$ is not full rank, it satisfies properties of a semi-definite inner product. It can be viewed as the usual matrix inner product of the 'low-energy weighted' operators $P_k^{1/2} \mathcal{O}_1 P_k^{1/2}$ and $P_k^{1/2} \mathcal{O}_2 P_k^{1/2}$.

The inner product defines a natural distance measure with which to compare $\mathcal{E}$ and $\mathcal{H}$, $\min_{\mathcal{N}} \| |\mathcal{E}) - |\mathcal{H}) \|^2 \equiv \min_{\mathcal{N}} \left[ (\mathcal{E}| - (\mathcal{H}| \right] \left[ |\mathcal{E}) - |\mathcal{H}) \right] \geq 0$. Here we minimize over the normalization constant $\mathcal{N}$ such that the distance is minimum. After some calculations, we arrive at

$$\min_N \| |\mathcal{E}) - |\mathcal{H}) \|^2 = \mathcal{H}^{(k)}(\rho) - \frac{1}{\mathcal{J}_{\mathcal{E}}^{(k)}(\rho)} \geq 0, \tag{F.29}$$

where

$$\mathcal{H}^{(k)}(\rho) = k! \sum_\pi |\text{tr} \left( W_\pi \left( \rho^{\frac{1}{2k}} \right)^{\otimes k} \left( \rho^{\frac{1}{2k}} \right)^{\otimes k} \right)|^2, \tag{F.30}$$

$$\mathcal{J}_{\mathcal{E}}^{(k)}(\rho) = \frac{\mathbb{E}_{U,V \in \mathcal{E}} \left[ |\text{tr} \left( U^\dagger \rho^{\frac{1}{2k}} V \rho^{\frac{1}{2k}} \right)|^{2k} \right]}{(k!)^2 \left[ \mathbb{E}_{U \in \mathcal{E}} \left[ \text{tr} \left( U^\dagger \rho^{1/k} U \rho^{1/k} \right) \right]^k \right]^2}. \tag{F.31}$$

Note that $\mathcal{H}^{(k)}(\rho)$ depends only on $\rho$. Observing

$$\mathcal{J}_{\mathcal{E}}^{(k)}(\rho) \geq \frac{1}{\mathcal{H}^{(k)}(\rho)}, \tag{F.32}$$

we recognize $\mathcal{J}_{\mathcal{E}}^{(k)}(\rho)$ as the frame potential. For $k = 1, 2$, we find the following lower bounds

$$\mathcal{H}^{(1)}(\rho) = 1, \qquad \mathcal{H}^{(2)}(\rho) = 2 + 2\operatorname{tr}(\rho^{1/2})^4. \tag{F.33}$$

We would like to emphasize that the density matrix $\rho$ reflects the Hilbert space of interest. Namely, when the frame potential achieves its lower bound, the ensemble $\mathcal{E}$ is CV $k$-design with respect to $\rho$. For instance, for a pure state $\rho = |\psi\rangle\langle\psi|$, the lower bound is always saturated for any ensembles.

The new frame potential can be used to verify CV $k$-designs. As an illuminating example, we calculate the $k = 1$ CV frame potential for our 1-design candidate $\mathbb{D}_n$ on a thermal state $\tilde{\rho}_{n_{\text{th}}}$ of mean photon number $n_{\text{th}}$. We obtain

$$\mathcal{J}_{\mathbb{D}_n}^{(1)}(\tilde{\rho}_{n_{\text{th}}}) = \left(\frac{(1 + n + 2n_{\text{th}})^2}{(1 + n)^2 + 4nn_{\text{th}}}\right)^N \geq 1 = \frac{1}{\mathcal{H}^{(1)}}, \tag{F.34}$$

where $N$ is the number of modes. It is illuminating to examine two limits: $n \gg n_{\text{th}}$, where the ensemble $\mathbb{D}_n$ contains an approximate basis for all operators acting on the subspace of $\lesssim n_{\text{th}}$ photons, and $n \ll n_{\text{th}}$, where it does not. For the latter, the frame potential scales as $(n_{\text{th}}/n)^N$ and continues decreasing as we increase the size $n$ of the ensemble. Once $n \gtrsim n_{\text{th}}$, the frame potential algebraically saturates to its lower bound of 1, up to terms of order $(1/n)$ and $(n_{\text{th}}/n)^2$, confirming again that $\lim_{n\to\infty}\mathbb{D}_n$ is a proper CV 1-design.

# Appendix G

# Details on learning quantum systems via out-of-time-order correlators

In this Appendix we provide additional theoretical and numerical details relating to the results of Ch. 13.

## G.1 Details of numerical simulations

Here we provide further details on the numerical simulations displayed in Figs. 12.2, 12.3, and 12.4.

### Correlation functions

We begin by explicitly writing down the correlation functions used in Figs. 12.2, 12.3, 12.4. Throughout, we denote time-evolved operators as $V(t) \equiv U(t,0)VU(t,0)^\dagger$, where the time-evolution unitary is $U(t_2, t_1) = \mathcal{T}\left\{e^{-i\int_{t_1}^{t_2} dt\, H(t)}\right\}$ and $H(t)$ is the time-dependent stroboscopic Floquet Hamiltonian specified in Ch. 12 (unless otherwise stated, in Fig. G.2).

In the restricted access scenario considered in Figs. 12.2, G.2, we use the following correlation functions:

$$
\begin{aligned}
C_{\text{TOC}}(t) &= \langle V_p(t)\, V_p(0)\rangle \\
C_{\text{TOC}}(x,t) &= \langle V_p(t)\, W_x(t/2)\, V_p(0)\, W_x(t/2)\rangle \\
C_{\text{OTOC}}(x,t) &= \langle V_p(t)\, W_x(0)\, V_p(t)\, W_x(0)\rangle
\end{aligned}
\tag{G.1}
$$

where $p$ denotes the probe qubit, and $\langle\cdot\rangle \equiv 2^{-L}\text{tr}(\cdot)$ is an infinite temperature average. Each of these correlation functions can be measured using state preparation and read-out on the probe qubit, combined with time-evolution and a single local unitary operation on the larger system. (In the case of the auto-correlation function $C_{\text{TOC}}(t)$ no local unitary operation is needed, $W_x = \mathbb{1}$.)

In principle, we envision allowing $V, W$ to run over all local operators in the system. For instance, they could run over all $4^w \binom{N}{w}$ Pauli operators of weight $\leq w$, where $w \sim O(1)$. This is naturally achieved by randomized measurement strategies such as shadow tomography with local Clifford unitaries and $O(3^w)$ measurements [112, 217]. In practice, we must restrict $V, W$ to a few possible values in numerical simulations. Specifically, we take $V = W \in \{\sigma_x, \sigma_z\}$ for TOCs, and $V = W \in \{\sigma_z\}$ for OTOCs. The OTOC is observed to be relatively insensitive to basis of $V$ and $W$, hence our choice to restrict to a single operator, $\sigma_z$ (further, we note that adding $\sigma_x$ OTOCs could only improve the relative advantage of OTOCs compared to TOCs). More broadly, we do not expect that adding additional pairs of $\{V, W\}$ will change the qualitative behavior of learning via TOCs and OTOCs. Specifically, we have seen that the learning advantage of OTOCs arises from their ability to detect highly non-local correlations in the system (i.e. large-weight components of the time-evolved operator $V_p(t)$, see Appendix G.2 for more detailed phenomenological estimates). These correlations are not detectable by any time-ordered correlator involving only few-body operators; indeed, in ergodic systems we generically expect that they are not efficiently detectable by *any* time-ordered measurement.

For Figs. 12.3, 12.4, we utilize two-point correlation functions between pairs of local operators:

$$
\begin{aligned}
C_{\text{TOC}} &= \langle V_x(t) \, W_{x'}(0) \rangle, \\
C_{\text{OTOC}} &= \langle V_x(t) \, W_{x'}(0) \, V_x(t) \, W_{x'}(0) \rangle.
\end{aligned}
\tag{G.2}
$$

We again take $V = W \in \{\sigma_x, \sigma_z\}$ for TOCs and $V = W \in \{\sigma_z\}$ for OTOCs. We allow $x, x'$ to span all qubits within a distance 2 of the link—this consists of 6 possible values for each of $x, x'$, corresponding to distances 0, 1, and 2 to both the left and right of the link. In principle, we would like $x, x'$ to run over the entire lattice; however, in practice we observe that correlation functions involving qubits distant from the link provide little information, and so can be safely neglected.

We now briefly comment on our numerical methods for computing the above correlation functions and the Fisher information [Figs. 12.2(b), 12.3(b)]. We compute the infinite temperature average in the correlation functions by sampling over Haar-random initial states $|\psi\rangle$. To motivate this, we can insert a resolution of the identity, $\mathbb{1} = \frac{1}{2^L} \sum_\psi |\psi\rangle\langle\psi|$ into the correlation functions Eq. (G.2) to obtain:

$$
C_{\text{TOC}} = \frac{1}{2^L} \sum_\psi \langle\psi| V_i(t) \, W_j(0) |\psi\rangle,
\tag{G.3}
$$

and similarly for the OTOC. In numerics, we approximate this sum by sampling a finite number $N_\psi$ of states $|\psi\rangle$ drawn from the Haar distribution; errors in this approximation will scale as $\sim 1/\sqrt{N_\psi 2^L}$.

In the learning problems considered in Ch. 12 [Fig. 12.2(a), 12.3(a), 12.4], we take $N_\psi = 25, 25, 10, 1, 1$ for system sizes $L = 6, 8, 10, 12, 14$, respectively. In contrast, when
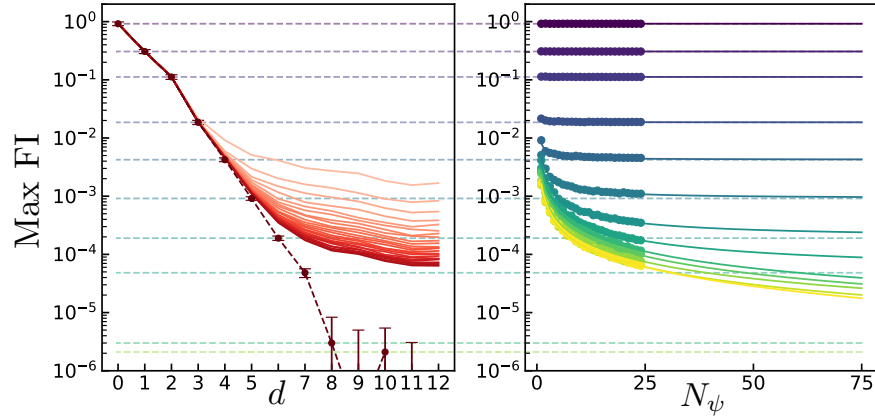
Figure G.1: Depiction of the extrapolation method used to calculate the maximum Fisher information over time-ordered correlators [Fig. 12.2(b)]. Each correlation function is computed for 25 Haar-random values of the state $|\psi\rangle$ [Eq. (G.3)]. For each value of $N_\psi$ between 1 and 25, we choose a random subset of $N_\psi$ values of $|\psi\rangle$ and compute the average correlation function over the subset. (Left) For each value of $N_\psi$, we then compute the maximum Fisher information over all correlation functions, max FI($N_\psi$) (solid red lines, darker lines corresponds to higher $N_\psi$). (Right) Our estimate of the maximum Fisher information at infinite temperature (dotted lines, both plots) is obtained by fitting max FI($N_\psi$) = max FI($\infty$) + $A/N_\psi$ and taking $N_\psi \to \infty$ (points denote data, solid lines denote $1/N_\psi$-fit).

estimating the Fisher information [Fig. 12.2(b), 12.3(b)], we perform a large-$N_\psi$ extrapolation to improve precision. This is required in order to establish the asymptotic scaling of the Fisher information at large $d$ [Fig. 12.2(b)] and small $J_\ell$ [Fig. 12.3(b)]. Specifically, we compute the estimated correlation function $C_{N_\psi}$ averaged over $N_\psi = 1, \ldots, 25$ Haar-random initial states, as well as the resultant Fisher information max FI($N_\psi$) $\equiv |\partial C_{N_\psi}/\partial J|^2$, maximized over all relevant correlation functions. We then perform a linear fit max FI($N_\psi$) = max FI($\infty$) + $\frac{A}{N_\psi}$, where max FI($\infty$) and $A$ are fitting parameters. Finally, the fitting parameter max FI($\infty$) represents our estimation of the Fisher information at $N_\psi \to \infty$, which we plot in Figs. 12.2(b), 12.3(b). We illustrate this procedure in Fig. G.1, using the data for Fig. 12.2(b). On the left of Fig. G.1, we plot the maximum Fisher information, max FI($N_\psi$), for each $N_\psi$, as a function of the distance $d$. We observe that in regions where $C$ is relatively large (i.e. small $d$), the estimates are quite accurate even for $N_\psi = 1$, while in regions where $C$ is small (i.e. large $d$) the Fisher information becomes successively smaller as the number of sampled states $N_\psi$ increases. On the right of Fig. G.1, we re-plot the Fisher information for each $d$ as a function of $N_\psi$. Solid lines represent the results of the linear fit, which we observe to fit the $N_\psi$-dependence of the data quite well. The extrapolated Fisher information [as displayed in Fig. 12.2(b)] is shown in Fig. G.1 as a dashed line.

## Imperfect time-reversal via cavity mode

In Fig. 12.4(a), we benchmark the effects of decoherence on learning by coupling the spin system to a single cavity mode. Our motivation for studying this model is two-fold. First, in ergodic many-body systems the effect of local errors on OTOCs is expected to be independent of the precise microscopic form of the error [427]. We therefore expect the spin-cavity system to display similar OTOC physics to more generic local error models. Second, for $L = 10$ spins the spin-cavity system can be exactly simulated in a Hilbert space of size $2^L \times L$ (we assume the cavity initially has zero occupation number; since the sum of spin magnetization and the cavity occupation is conserved, the cavity occupation is upper bounded by $L$). This is substantially smaller than the requirements to exactly simulate a mixed state quantum system, $2^{2L}$.

More specifically, the spin-cavity Hamiltonian is as follows. We modify the Floquet time-evolution described in Ch. 12 to alternate between the following two Hamiltonians:

$$
\begin{aligned}
H_1 &= \pm H_f + g \sum_i \left( a^\dagger \sigma_i^- + a \sigma_i^+ \right) + \omega a^\dagger a \\
H_2 &= \pm H_c + g \sum_i \left( a^\dagger \sigma_i^- + a \sigma_i^+ \right) + \omega a^\dagger a
\end{aligned}
\tag{G.4}
$$

where $H_f, H_c$ are the field and coupling Hamiltonians written in Ch. 12, $a, a^\dagger$ are lowering/raising operators for a bosonic cavity mode, $g = \{0.0, 0.25, 0.5\}$ is the spin-cavity interaction strength, and $\omega = 1.7$ is the cavity frequency. Here, the $\pm$ denote values during forwards/backwards time-evolution; note that we do *not* reverse the spin-cavity interaction or the cavity frequency during backwards time-evolution.

## Learning model

We now detail the machine learning techniques used in Figs. 12.2(a), 12.3(a), 12.4 and G.3(b). Throughout, read-out error is mimicked by adding a random Gaussian variable with mean zero and standard deviation $\delta$ to the exact correlation functions.

We begin with Fig. 12.2(a). Our goal is to predict the value of $d$ [which specifies the geometry of the spin system, see Fig. 12.2(a)] from the correlation functions of the system, Eq. (G.1). To do so, we train a learning model on 3000 randomly drawn disorder realizations of the Hamiltonian, consisting of 300 realizations each for $d = 0, 1, \ldots, 9$. We test model performance on 2000 additional disorder realizations, again consisting of 200 realizations each for $d = 0, 1, \ldots, 9$. For each disorder realization, the input to our learning model consists of the correlation functions Eq. (G.1), evaluated at $x = 2, \ldots, L$ and 30 evenly spaced times between 0 and 12. We apply Gaussian distributed read-out error $\delta = 3\%$ to each correlation function. We repeat this procedure, as well as the model training and evaluation that follows, first using only TOCs as input to the learning algorithm, and second using both TOCs and OTOCs.

Next, we input these correlation functions into a support vector regression (SVR) model with radial basis functions [78]. The radial SVR contains two hyperparameters: $C$, the regularization parameter, and $\gamma$, which controls the width of the radial basis functions. We choose $C$ and $\gamma$ by performing five-fold cross-validation over the sets $C = \{10, 30, 100\}$, $\gamma = \{0.03, 0.06, 0.1, 0.3, 1, 3, 6, 10\}$. We obtain $C = 10, \gamma = 1$ for learning via TOCs, and $C = 10, \gamma = 0.03$ for learning via both TOCs and OTOCs. In the identical learning task for Hamiltonian evolution [Fig. G.2(a)], we obtain $C = 10, \gamma = 1$ for learning via TOCs, and $C = 100, \gamma = 0.03$ for learning via both TOCs and OTOCs.

We now turn to Fig. 12.3(a) and 12.4. Our goal is to perform binary classification using the correlation functions Eq. (G.2) to distinguish whether the link interaction strength is zero or nonzero. To do so, we simulate the correlation functions of 300 randomly drawn disorder realizations of the Hamiltonian for each link strength,

$$J_\ell = \{0, 0.01, 0.017, 0.03, 0.06, 0.1, 0.17, 0.3, 0.6, 1.0\}. \tag{G.5}$$

For each nonzero $J_\ell$, we train a learning model to perform binary classification between link strength 0 and $J_\ell$. We test model performance on 400 additional disorder realizations, again consisting of 200 realizations each for link strength 0 and $J_\ell$.

The first step of our learning model is to prune the correlation functions used as input. We do so by estimating the mutual information between each individual correlation function and the link interaction strength, and selecting the $K$ correlation functions with the highest mutual information. Here $K$ is a hyperparameter that will ultimately be chosen via cross-validation. To estimate the mutual information, we fit the distribution of correlation functions values over disorder realizations to a Gaussian for each link strength, and compute the Jensen-Shannon divergence between the Gaussian distributions. The Jensen-Shannon divergence is equal to the desired mutual information [296].

As before, we input the selected correlation functions into a support vector machine (SVM) with radial basis functions [78]. We now have three hyperparameters: the SVM hyperparameters $C$ and $\gamma$ and the number of selected correlation functions $K$. We choose $C$, $\gamma$, and $K$ by performing five-fold cross-validation over each value $C = \{0.1, 1, 10, 100, 1000\}$, $\gamma = \{0.1, 0.3, 1, 3, 10, 30\}$. We obtain Figs. 12.4(b), by repeating this procedure for various simulated read-out errors,

$$\delta = \{0.0001, 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1\}. \tag{G.6}$$

At each read-out error, we perform a linear interpolation of the classification accuracy as a function of $J_\ell$ [as shown in Fig. 12.3(a) for $\delta = 0.03$]. The minimum detectable link strength $J_\ell^*$ occurs at the intersection of this interpolation with a horizontal line (not depicted) corresponding to a classification accuracy of 90%.

Finally, we turn to Fig. G.3(b) in Appendix G.3. In the probe qubit scenario, training and testing are performed on 300 and 200 samples respectively for each geometry and each value of $d$. In the global state preparation and read-out scenario, we instead use 60 and 40 samples respectively for each geometry. Our learning model consists of a support vector machine

with hyperparameters chosen via 4-fold cross-validation from the sets $C = \{0.1, 1, 10\}$, $\gamma = \{0.3, 3, 30\}$. As in our previous learning tasks, we apply a read-out error $\delta = 3\%$ to each correlation function before use in learning.

## G.2   Phenomenological estimates

In this section, we provide more detailed reasoning behind the phenomenological estimates of the Fisher information presented in Ch. 12. We begin with the Fisher information under unitary dynamics with restricted access and then turn to the effects of imperfect time-reversal and decoherence.

### Fisher information in restricted access scenario

At sufficiently large times and distances, we expect the profile of correlation functions in ergodic many-body systems to be described by only a few phenomenological parameters. For instance, in one-dimensional systems the out-of-time-order correlator is predicted to take the following functional form [17, 341],

$$C_{\text{OTOC}}(x,t) \approx f\left(\frac{x/v_B - t}{A\sqrt{t}}\right) \tag{G.7}$$

where the phenomenological parameters $v_B$ and $A$ describe the butterfly velocity and the width of the OTOC wavefront, respectively. Here $f$ is a compactly supported bump function which interpolates between zero and one and then zero again within an $O(1)$-sized region about the origin. Meanwhile, in systems with a local conservation law, we expect time-ordered correlators to be dominated by diffusion of the conserved quantity. This leads to the following profile for the auto-correlation function,

$$C_{\text{TOC}}(t) \sim \frac{1}{\sqrt{Dt}}, \tag{G.8}$$

where $D$ is a diffusion constant. In the absence of conserved quantities, one expects time-ordered correlation functions to instead decay exponentially in time,

$$C_{\text{TOC}}(t) \sim \exp(-\gamma t), \tag{G.9}$$

parameterized by a decay rate $\gamma$.

To obtain the Fisher information, $\text{FI}(J|C) = |\partial C/\partial J|^2$, we must compute the derivative of the correlation functions with respect to a local coupling strength, $J_y$. To do so while leveraging the above phenomenological predictions, we must first recognize that the phenomenological parameters are themselves dependent on the local coupling strengths of the system, e.g. $v_B \to v_B(\{J_y\})$. We expand on this in further detail for each case below. The resultant scaling of the Fisher information in various physical regimes is summarized in Table G.1.

Table G.1: Maximum Fisher information in restricted access scenarios

| Correlation function | Probe qubit with local unitary control | Probe qubit with global unitary control |
|---|---|---|
| TOC without conserved quantity | $O(\exp(-d))$, [Fig. 12.2(b)] | $O(\exp(-d))$, |
| TOC with conserved quantity | $O(1/d^4)$ | $O(1/d^4)$ |
| OTOC | $O(1/d)$, [Fig. 12.2(b)] | $O(1/d^2)$, |

Table G.2: Phenomenological estimates of the scaling of the Fisher information in the restricted access scenario, for learning an interaction that lies a distance $d$ from the probe qubit.

## Fisher information of OTOCs

We begin with the Fisher information of OTOCs. Our treatment is broken into two parts, corresponding to the scenarios where the experimenter has either local or global unitary control over the larger system. The former scenario is simulated numerically in Fig. 12.2.

*Local unitary control.*—We consider local OTOCs [Eq. (G.1)] and are interested in the dependence of the OTOC on the local coupling strengths, $\{J_y\}$. To approach this, we will assume that OTOC takes the same functional form as in Eq. (G.7),

$$C_{\text{OTOC}}(x,t) = \langle V_p(t) W_x(0) V_p(t) W_x(0) \rangle \approx f\left(\frac{x/v_B(x) - t}{A\sqrt{t}}\right), \tag{G.10}$$

but now with a *position-dependent* butterfly velocity, $v_B(x)$. Specifically, we assume that the effective butterfly velocity at time $t$ receives contributions from all couplings that have been visited thus far, i.e. all $J_y$ with $y \lesssim x$. Since the time to traverse a single coupling is proportional to the inverse coupling strength $1/J_y$, we expect the time to traverse all couplings up to a distance $x$ to be proportional to the sum $\sum_{y=0}^{x} 1/J_y$. Equating this time to the distance divided by the effective butterfly velocity $x/v_B(x)$, we have:

$$v_B(x) \approx \left[\frac{1}{x}\sum_{x=0}^{x}\frac{1}{J_x}\right]^{-1}. \tag{G.11}$$

If each coupling strength is drawn independently from some disorder realization, then at large times the butterfly velocity will be close to its typical value, $v_B = \overline{1/J}$.

We can now compute derivatives of the correlation function with respect to a given coupling strength via the chain rule. The derivative of the butterfly velocity is

$$\partial_{J_d} v_B(x) \approx \frac{v_B(x)^2}{J_d^2 x} \cdot \delta_{d \leq x}, \tag{G.12}$$

which yields the following for the OTOC:

$$\begin{aligned}
\partial_{J_d} C_{\text{OTOC}}(x,t) &\approx -\frac{x}{A\sqrt{t} v_B(x)^2} \cdot f'\left(\frac{x/v_B(x) - t}{A\sqrt{t}}\right) \cdot \partial_{J_d} v_B(x) \\
&\approx -\frac{1}{A J_d^2 \sqrt{t}} \cdot \delta_{d \leq x} \cdot f'\left(\frac{x/v_B(x) - t}{A\sqrt{t}}\right).
\end{aligned} \tag{G.13}$$

There are two parameters of the local OTOC chosen by a potential experimentalist: the position $x$ of the local perturbation, and the evolution time $t$. We are interested in the maximum Fisher information given an optimal choice of $x$ and $t$. Observing Eq. (G.13), we see that the derivative $f'$ is maximized by the choice $x = v_B t$, while the delta function then sets $v_B t = d$. Plugging these values in, we find the Fisher information

$$\max_C \text{FI}(J | C_{\text{OTOC}}) \approx \left| \frac{f'(0)}{A J_y^2 \sqrt{v_B d}} \right|^2 \sim \frac{1}{d}. \tag{G.14}$$

*Global unitary control.*—We now turn to an alternate experimental scenario, where one has only global unitary control over the larger system. In this scenario, the natural generalization of the correlation functions Eq. (G.1) is the following:

$$\begin{aligned}
C_{\text{TOC}} &= \langle V_p(t) V_p(0) \rangle \\
C_{\text{TOC}} &= \langle V_p(t) e^{i\phi \sum_x W_x(t/2)} V_p(0) e^{-i\phi \sum_x W_x(t/2)} \rangle \\
C_{\text{OTOC}} &= \langle V_p(t) e^{i\phi \sum_x W_x(0)} V_p(t) e^{-i\phi \sum_x W_x(0)} \cdot \rangle
\end{aligned} \tag{G.15}$$

Here we replace the local unitary operations of Eq. (G.1) with global spin rotations, $e^{i\phi \sum_x W_x}$, by an angle $\phi$ (here, $W_x$ is a local Hermitian operator on qubit $x$).

We expect the behavior of the OTOC under global control to be governed by the "size" of time-evolved operators [139, 416, 418]. The size corresponds to the average of local OTOCs over all qubits in the system [379]. In one-dimensional ergodic systems the size grows linearly $\sim v_B t$, which yields the following phenomenological expectation for the global OTOC [139]:

$$C_{\text{OTOC}}^{\text{glob}} = \exp\left(-\phi^2 v_B(t) t\right). \tag{G.16}$$

Here we have made the butterfly velocity time-dependent to capture its dependence on the local coupling strengths,

$$v_B(t) \approx \left[ \frac{1}{v_B t} \sum_{y=0}^{v_B t} \frac{1}{J_y} \right]^{-1}, \tag{G.17}$$

where $v_B = \overline{1/J}$ is the typical butterfly velocity.

Taking the derivative of the OTOC via the chain rule, we have:

$$
\begin{aligned}
\partial_{J_d} C_{\mathrm{OTOC}}^{\mathrm{glob}} &\approx -\phi^2 t \cdot \partial_{J_d} v_B(t) \cdot \exp\left(-\phi^2 v_B(t) t\right) \\
&= -\phi^2 t \cdot \frac{v_B(t)^2}{v_B J_d^2 t} \cdot \delta_{d \le v_B(t) t} \cdot \exp\left(-\phi^2 v_B(t) t\right) \\
&\approx -\phi^2 \cdot \frac{v_B}{J_d^2} \cdot \delta_{d \le v_B t} \cdot \exp\left(-\phi^2 v_B t\right).
\end{aligned}
\tag{G.18}
$$

We would like to maximize the Fisher information over the parameters $(\phi, t)$. This entails taking the time $t$ to be as early as allowed by the delta function, $t \approx d/v_B$, in order to minimize the exponential. The correlator is then maximized by choosing $\phi$ such that $\phi^2 v_B t \sim 1$. This gives a Fisher information:

$$
\max_C \mathrm{FI}(J|C_{\mathrm{OTOC}}) \approx \left| \frac{v_B e^{-1}}{J_y^2 d} \right|^2 \sim \frac{1}{d^2},
\tag{G.19}
$$

which decays algebraically, with an additional factor of $d$ compared to the local unitary control scenario.

Before moving on, we briefly summarize the intuition behind the two above estimates. In both cases, an $O(1)$ perturbation in a local coupling strength produces an $O(1)$ shift in the location of the OTOC wavefront. With local control, this shift produces an $O(1/\sqrt{d})$ change in the OTOC, since the OTOC wavefront is spread across a width $\sim \sqrt{d}$ by the time it reaches the coupling. With global control, this produces an $O(1/d)$ change in the OTOC, since the global OTOC depends on the average of $\sim d$ individual coupling strengths. Since the Fisher information involves the square of the OTOC derivative, these lead to an $O(1/d)$ and $O(1/d^2)$ Fisher information, respectively.

**Fisher information of TOCs in absence of conserved quantities**

We now turn to a simpler case, the Fisher information of time-ordered correlators in the absence of conserved quantities [Fig. 12.2(b)]. Under ergodic dynamics, we expect such correlation functions to decay exponentially in time at sufficiently large times, see Eq. (G.9). Now, consider the derivative of the correlation function with respect to a local coupling strength at a distance $d$ away from the probe qubit. By causality, this derivative can only be non-zero after a time $t \gtrsim d/v_B$. However, at such times the magnitude of the correlation function has already decayed by a factor of $e^{-\gamma t}$. This suggests that the Fisher information will decay exponentially in the distance $d$,

$$
\max_C \mathrm{FI}(J|C_{\mathrm{TOC}}) \lesssim \exp(-2\gamma x/v_B),
\tag{G.20}
$$

as observed numerically in Fig. 12.2(b).

**Fisher information of TOCs in presence of conserved quantities**

The scaling of the Fisher information for TOCs is modified in the presence of a conserved quantities. In this case, one expects the TOC at sufficiently large times to be dominated by slow diffusive dynamics of the conserved quantity. This lies in contrast to the exponential decay expected in the absence of conserved quantities.

To study this, we begin with the auto-correlation function [first line of Eq. (G.1)]. Recall that the auto-correlation function can be measured with access solely to the probe qubit, and is thus accessible in both the local and global control scenarios. Similar to the case of OTOCs, we will assume that the dependence of the correlation function on the local coupling strengths is captured by replacing the diffusion constant, $D$, with a *time-dependent* value,

$$C_{\text{TOC}}(t) = \frac{1}{\sqrt{D(t)t}}. \tag{G.21}$$

Following the logic of the previous section, we assume the effective diffusion constant takes the form [392],

$$D(t) \sim \left[ \frac{1}{\sqrt{Dt}} \sum_{y=0}^{\sqrt{Dt}} \frac{1}{J_y} \right]^{-1}, \tag{G.22}$$

where $D = \overline{1/J}$ is the diffusion constant's typical value. Differentiating with respect to the local coupling strength gives

$$\partial_{J_d} D(t) \sim \frac{D(t)^2}{J_d^2 \sqrt{Dt}} \cdot \delta_{d < \sqrt{Dt}}. \tag{G.23}$$

Computing the derivative of the auto-correlation function, we have:

$$\begin{aligned} \partial_{J_d} C_{\text{TOC}}(t) &= -\frac{1}{2D(t)^{3/2}t^{1/2}} \cdot \partial_{J_d} D(t) \\ &= -\frac{D(t)^{1/2}}{2J_d^2 t \sqrt{D}} \cdot \delta_{d < \sqrt{Dt}} \\ &\approx -\frac{1}{2J_d^2 t} \cdot \delta_{d < \sqrt{Dt}}. \end{aligned} \tag{G.24}$$

The magnitude of the derivative is maximized by taking the minimum possible time, $t \approx d^2/D$, which yields a Fisher information,

$$\max_C \text{FI}(J | C_{\text{TOC}}) \approx \left| \frac{D}{2J_x^2 d^2} \right|^2 \sim \frac{1}{d^4}. \tag{G.25}$$

This can be understood intuitively as follows. For the auto-correlation function to be sensitive to the coupling strength $J_d$, the conserved quantity must have spread to at least distance

$d$. At such a distance the magnitude of the auto-correlation function is $O(1/d)$, since the conserved quantity has spread over $\sim d$ sites. In addition, the derivative with respect to an individual coupling strength is suppressed by an additional factor $O(1/d)$, since the auto-correlator depends only on the average (inverse) coupling strength over $\sim d$ sites. Combining these two factors and squaring leads to an $O(1/d^4)$ Fisher information.

We now turn to the remaining time-ordered correlation functions in Eqs. (G.1, G.15), which require either local or global unitary control over the non-probe qubits. We will find that such correlators provide no scaling advantage beyond the auto-correlator.

We first consider the case of local unitary control [Eq. (G.1)]. Physically, these correlation functions correspond to preparing an amount of the conserved quantity (e.g. a spin polarization) at the probe qubit, letting it diffuse for a time $t/2$, flipping the spin polarization at a qubit $x$, and measuring the polarization at the probe qubit after an additional time $t/2$. We thus expect the TOC behave as follows,

$$C_{\mathrm{TOC}}(x,t) = \langle V_p(t)\, W_x(0)\, V_p(t)\, W_x(0)\rangle \approx q(0,t) - 2q(x,t/2)\cdot q(x,t/2), \qquad \text{(G.26)}$$

where $q(x,t) \approx (2\pi D(t)t)^{-1/2}\exp(-x^2/(2D(t)t))$ is the propagator of the conserved quantity from position 0 to position $x$ (or vice versa). The first term is equal to the auto-correlation function. The second term arises from the spin flip at position $x$ and time $t/2$. The spin flip effectively inserts a negative polarization $-2q(x,t/2)$ on the qubit $x$, which propagates back to the probe qubit with amplitude $q(x,t/2)$.

The derivative of the second term is as follows,

$$
\begin{aligned}
\partial_{J_d}\left[q(x,t/2)^2\right] &= \partial_{J_d}\left[\frac{1}{\pi D(t)t}\exp\left(-\frac{2x^2}{D(t)t}\right)\right] \\
&\approx \partial_{J_d}\left[\frac{-1}{\pi J_y^2 D^{1/2}t^{3/2}}\exp\left(-\frac{2x^2}{D(t)t}\right) + \frac{2x^2}{\pi J_y^2 D^{3/2}t^{5/2}}\exp\left(-\frac{2x^2}{D(t)t}\right)\right]\cdot\delta_{d<\sqrt{Dt}}\,.
\end{aligned}
$$

$$\text{(G.27)}$$

The magnitude of the derivative is maximized at $Dt \sim x^2, x \sim d$, and is of order $O(1/d^3)$. This is subleading compared to the auto-correlation function, of order $O(1/d^2)$, and thus does not affect the asymptotic scaling of the Fisher information with $d$.

The case of a global control [Eq. (G.15)] is even simpler. A global spin rotation about the $x$-axis by an angle $\phi$ multiplies the conserved quantity at each site by a factor of $\cos(\phi)$. Here we assume that the $x$- and $y$-components of spin that are generated by the rotation quickly decay in time if they are not conserved by the ergodic dynamics. The resulting correlation function is then given by the auto-correlation multiplied by $\cos(\phi)$. Again, this provides no scaling advantage in the Fisher information.

## Effect of imperfect time-reversal and decoherence on Fisher information

We now incorporate imperfect time-reversal dynamics into our estimates of the Fisher information of OTOCs. Previous works have been found that a wide range of experimental

errors (e.g. extrinsic decoherence, coherent errors in time-reversal) have a similar effect on OTOC measurements, as long as the relevant errors are local and the dynamics are ergodic [139, 326, 416, 427, 450].

Specifically, in one-dimensional systems, one expects that the OTOC under open-system dynamics, $\tilde{C}_{\mathrm{OTOC}}$, is equal to the same OTOC under unitary dynamics, $C_{\mathrm{OTOC}}$, multiplied by an overall Gaussian decay in time [427]:

$$\tilde{C}_{\mathrm{OTOC}} \approx \exp\left(-a\varepsilon v_B t^2\right) \times C_{\mathrm{OTOC}}. \tag{G.28}$$

Here $\varepsilon$ is an effective local error rate, $v_B$ is the butterfly velocity, and $a$ is an order one constant. The argument of the above exponential is proportional the volume of the time-evolved operator's light cone. Intuitively, Eq. (G.28) states that each error in the causal past of an operator contributes a roughly equal amount to the decay of the OTOC. We note that in finite-size systems we do not expect Eq. (G.28) to precisely hold, however, corrections are expected to be suppressed by $\sim \varepsilon/J$ where $J$ is the local interaction strength [427], so we neglect them here.

Substituting Eq. (G.28) into our estimate for the Fisher information [Eq. (G.19)] and setting $v_B t \approx d$, we find:

$$\max_C \mathrm{FI}(J_d|\tilde{C}_{\mathrm{OTOC}}) \sim \frac{1}{d^2} \exp\left(-a\varepsilon d^2/v_B\right). \tag{G.29}$$

Meanwhile, we assume that the Fisher information with respect to TOCs is comparatively unaffected by error, and thus once again follows a linear exponential decay in $d$:

$$\max_C \mathrm{FI}(J_d|\tilde{C}_{\mathrm{TOC}}) \sim \exp(-\gamma d). \tag{G.30}$$

Setting the two exponentials to be equal, $\max_C \mathrm{FI}(J_d|\tilde{C}_{\mathrm{OTOC}}) \sim \max_C \mathrm{FI}(J_d|\tilde{C}_{\mathrm{TOC}})$, we find that the OTOC continues to provide an advantage over the TOC up to

$$d \lesssim \frac{\gamma v_B}{\varepsilon}, \tag{G.31}$$

as quoted in Ch. 12.

We now apply the same analysis to our second learning regime. Let us set $v_B \sim J$ for consistency with Ch. 12. The Fisher information of an OTOC between operators on either side of the link with respect to the link interaction strength is now modified to

$$\max_C \mathrm{FI}(J_\ell|\tilde{C}_{\mathrm{OTOC}}) \sim \frac{J_\ell^4 t^2}{J^2} \exp\left(-a\varepsilon J t^2\right), \quad Jt \lesssim L. \tag{G.32}$$

The maximum of the Fisher information as a function of time now occurs at

$$t^* \sim \min\left\{\sqrt{\frac{1}{\varepsilon J}}, \frac{L}{J}\right\}, \tag{G.33}$$
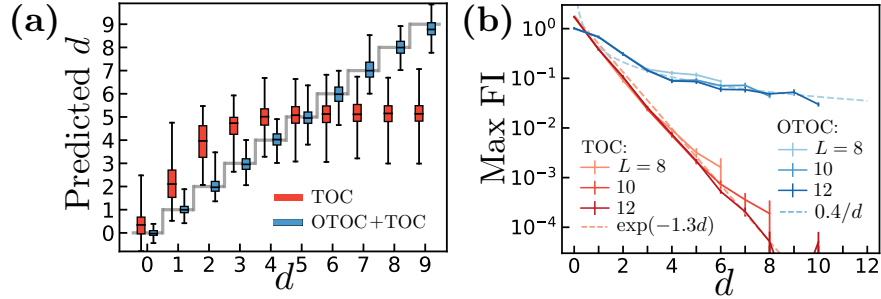
Figure G.2: Learning in the restricted access scenario under Hamiltonian evolution. Numerical simulations are performed identically to Fig. 12.2 but now with Hamiltonian evolution under $(H_c + H_f)/2$ instead of Floquet evolution. In both **(a)** the learning task and **(b)** the Fisher information, the results for learning Hamiltonian dynamics are qualitatively similar to the results for learning Floquet dynamics [Fig. 12.2]. In **(b)**, the maximum Fisher information is averaged over 100 disorder realizations for both TOCs and OTOCs. At large $d$, we expect the Fisher information for Hamiltonian evolution to approach a power law decay $\sim 1/d^4$ (see Appendix G.2), but this cannot be observed in our finite-size numerics.

with value

$$\max_C \mathrm{FI}(J_\ell | \tilde{C}_{\mathrm{OTOC}}) \sim \min\left\{ \frac{J_\ell^4}{\varepsilon J^3}, \frac{J_\ell^4 L^2}{J^4} \right\}, \qquad (\mathrm{G.34})$$

which differs from the unitary OTOC for sufficiently high error rates. Again, we assume that the Fisher information of TOC is not affected by error to leading order. Taking the square root of Eq. (G.34), we thus find the $L$-fold advantage of the OTOC is replaced by a $\sqrt{J/\varepsilon}$-fold advantage at error rates $\sqrt{\varepsilon/J} \gtrsim 1/L$, as quoted in Ch. 12.

# G.3 Additional numerics

In this section we provide numerical results in two additional learning scenarios. We begin by repeating the simulations leading to Fig. 12.2 for time-independent Hamiltonian evolution instead of Floquet evolution. We then discuss learning in the restricted access scenario with global unitary control, in contrast to local unitary control as considered in Ch. 12).

## Learning with time-independent Hamiltonian evolution

In the numerical simulations in Ch. 12, we utilized Floquet time-evolution in which the spin interactions and local fields were applied in a stroboscopic fashion. Our motivations for using Floquet time-evolution instead of Hamiltonian time-evolution were three-fold: First, Floquet dynamics are prevalent in a variety of quantum systems that one might wish to

learn, e.g. in digital quantum simulators, and NMR or solid-state defect setups with optical driving. Second, the Floquet dynamics considered are moderately faster to simulate via Krylov subspace methods than Hamiltonian dynamics, since the Hamiltonian of the former contains fewer terms at a given instant in time. Third, we do not expect the behavior of learning via TOCs or OTOCs under the two dynamics to qualitatively differ at moderate times and distances (although at large distances they may, see Appendix G.2).

Here, we check the latter assumption by repeating the numerical analysis of Fig. 12.2 using time-independent Hamiltonian dynamics. As shown in Fig. G.2(a), we find that the results of the learning task of Fig. 12.2(a) behave quite similarly for Hamiltonian and Floquet dynamics. In particular, access to OTOCs continues to enable substantially more accurate predictions for the crossing distance $d$ for all $d \gtrsim 3$. In Fig. G.2(b), we turn to the behavior of the Fisher information as a function of a coupling's distance from the probe qubit. Unfortunately, we are not able to discern the $\sim 1/d^4$ scaling predicted in Appendix G.2 in our finite-size numerics. Instead, the Fisher information behaves qualitatively similar to that of Floquet dynamics [Fig. 12.2(b)]. We anticipate that at sufficiently large distances the Fisher information of Hamiltonian dynamics will indeed asymptote to the expected power law decay. However, at such distances the Fisher information will likely already be too small to be useful for most practical purposes.

## Learning under restricted access with global unitary control

We now turn to learning when one has only global unitary control over the system of interest. We consider a learning task where one wishes to classify the geometry of an unknown spin system, which we assume is drawn with equal probability from the three geometries shown in Fig. G.3(a). We find that access to OTOCs provides a substantial advantage in this classification task. Notably, we find that OTOCs continue to improve learning even when one has only global state preparation, control, and read-out (i.e. even in the absence of a probe qubit).

The classification problem we consider is a close variant of those introduced in Ch. 12. We suppose that one has access to the correlation functions of an unknown Hamiltonian whose connectivity corresponds to one of the three geometries shown in Fig. G.3(a). The goal is to distinguish which geometry describes the Hamiltonian. We again approach this task by training and testing a support vector machine on samples of disorder realizations, see Section G.1 for details.

We consider learning in two different experimental access scenarios. First, we consider the scenario where one has state preparation and read-out from a single probe qubit, and global control over the remainder of the system. In this case, we take the probe qubit to be a distance $d$ away from any distinguishing features of the geometry (see Fig. G.3), and study the learnability as a function of $d$. Note that we are restricted to relatively small distances, $d \leq 4$, owing to the particular form of the three geometries considered. We find that access to OTOCs increases the classification accuracy between 10% and 35% for all values of $d$
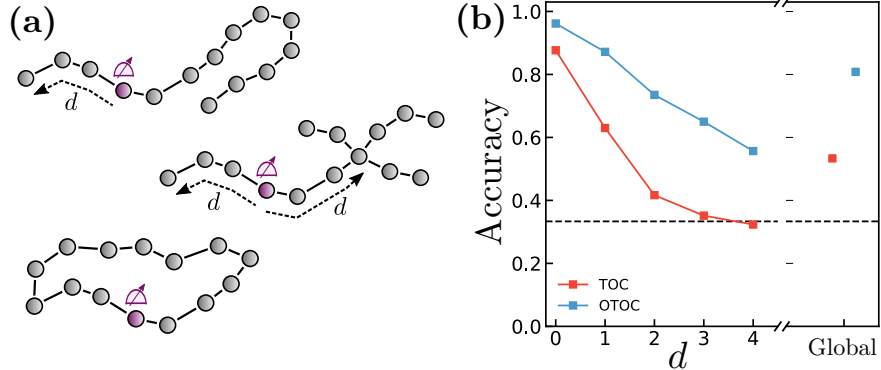
Figure G.3:   **(a)** The three spin geometries considered in the learning task defined in the text. Each geometry consists of $L = 14$ spins. The probe qubit (purple) is located along a subset of the system that is identical between the three geometries up to a distance $d$ away from the probe. **(b)** Accuracy of classification, using correlation functions that can be measured with (left) state preparation and read-out on the probe qubit and global unitary control over the remaining system, and (right) global state preparation, unitary control, and read-out. For the former, the accuracy is plotted as a function of the distance $d$ of the probe qubit from the geometric feature of interest. In both scenarios, access to OTOCs (blue) substantially improves the classification accuracy compare solely accessing TOCs (red).

[Fig. G.3(b)]. For instance, OTOCs allow classification with accuracy $\sim 65\%$ at $d = 3$, at which learning via TOCs has nearly trivial accuracy.

Our second scenario is even more restrictive: we suppose that one has only global state preparation, control and read-out over the entire system. Despite being commonplace in experiments such as NMR spectroscopy [279], learning in this scenario remains quite difficult in strongly-interacting systems, due to the combination of time-ordered correlators decaying quickly and local information being averaged out by global control and measurement. Indeed, in our learning task, we find that learning via TOCs features a classification accuracy of only $\sim 55\%$. Intuitively, we expect access to global OTOCs to improve learning, as operator spreading at late times is dependent on global geometric features of the system. In keeping with this intuition, we find that learning via both TOCs and OTOCs improves the classification accuracy to $\sim 80\%$.

# Appendix H

# Proof of information-theoretic hardness of out-of-time-order correlators

In this Appendix we provide a detailed proof of the main result of Ch. 13.

## H.1 Technical preliminaries

Before delving into the proof of Theorem 1, we require some technical definitions and tools. Included are more precise versions of Definition 1 and Definition 2 from Ch. 13.

### Notation

We will work with an $n$ qubit Hilbert space $\mathcal{H} \simeq (\mathbb{C}^2)^{\otimes n}$, for $n$ even. The dimension of the Hilbert space is $2^n$, which we denote by $d$. Our main proof will involve extensive use of diagrammatic tensor network notation, reviewed in detail in [96]. Our conventions for the diagrams will match those of [96].

### Haar integration and Weingarten functions

Consider the unitary group $U(d)$. It will be convenient to use multi-index notation, wherein $I = (i_1, ..., i_k)$ and analogously for $I', J, J'$. Letting

$$U^{\otimes k}_{IJ} := U_{i_1 j_1} U_{i_2 j_2} \cdots U_{i_k j_k} \,, \tag{H.1}$$

we will be interested in computing expectation values of the form $\mathbb{E}_{U \sim \mathrm{Haar}(d)}[U^{\otimes k}_{IJ} U^{\dagger \otimes k}_{J'I'}]$. To write out the result of this expectation value, we denote by $S_k$ the symmetric group on $k$ elements; for $\sigma \in S_k$, we adopt the notation

$$\delta_{\sigma(I), I'} := \delta_{i_{\sigma(1)}, i'_1} \delta_{i_{\sigma(2)}, i'_2} \cdots \delta_{i_{\sigma(k)}, i'_k} \,. \tag{H.2}$$

Then we have the useful identity (see e.g. [262] for a review)

$$\mathbb{E}_{U \sim \text{Haar}(d)}[U_{IJ}^{\otimes k} U_{J'I'}^{\otimes k}] = \sum_{\sigma,\tau \in S_k} \delta_{\sigma(I),I'} \delta_{\tau(J),J'} \text{Wg}^U(\sigma\tau^{-1}, d)\,, \tag{H.3}$$

where $\text{Wg}^U(\,\cdot\,, d) : S_k \to \mathbb{R}$ is the Weingarten function. This function can be constructed somewhat explicitly in the following way. In a slight abuse of notation, let us also denote by $\sigma, \tau$ their representation on $\mathcal{H}^{\otimes k}$, and define

$$G^U(\sigma\tau^{-1}, d) := \text{tr}(\sigma\tau^{-1}) = d^{\#(\sigma\tau^{-1})} \tag{H.4}$$

where $\#(\sigma\tau^{-1})$ is the number of cycles of $\sigma\tau^{-1}$. Viewing $G^U(\sigma\tau^{-1}, d)$ as a $k! \times k!$ matrix $G_{\sigma^{-1}, \tau}$, we have that $\text{Wg}^U(\sigma^{-1}\tau, d)$ is its matrix inverse. That is,

$$\sum_{\tau \in S^k} \text{Wg}^U(\sigma^{-1}\tau, d)\, G^U(\tau^{-1}\pi, d) = \delta_{\sigma,\pi}\,. \tag{H.5}$$

Having defined the Weingarten function, let us state a few useful results from the literature which we will leverage in our proofs:

**Theorem 3.2 of [108] 1.** *For any $\sigma \in S_k$ and $d > \sqrt{6}\, k^{7/4}$,*

$$\frac{1}{1 - \frac{k-1}{d}} \leq \frac{(-1)^{k-\#(\sigma)}d^{2k-\#(\sigma)}\, Wg^U(\sigma, d)}{\prod_i \frac{(2\ell_i-2)!}{(\ell_i-1)!\ell_i!}} \leq \frac{1}{1 - \frac{6k^{7/2}}{d^2}}\,, \tag{H.6}$$

*where the left-hand side inequality is valid for any $d \geq k$. Here $\sigma \in S_k$ has cycle type $(\ell_1, \ell_2, ...)$.*

We will in fact use the following corollary of this result:

**Corollary 1.** $|Wg^U(\mathbb{1}, d) - d^{-k}| \leq O(k^{7/2}d^{-(k+2)})$.

Finally we state a Lemma from [10]:

**Lemma 6 of [10] 1.** $\sum_{\tau \in S_k} |Wg^U(\tau, d)| = \frac{(d-k)!}{d!}$.

## Learning tree formalism

In Ch. 13, we provided definitions of time-ordered and out-of-time-order experiments based on the learning tree framework in quantum learning theory [10, 96]. It is useful to formalize these more precisely; our definitions below are closely based off of Definition 6.1 of [96].

**Definition 3** (Tree representation for learning a collection of channels without a quantum memory)**.** *Let $S = \{\mathcal{C}_i\}_i$ be a set of quantum channels on states on $\mathcal{H}$. A quantum learning algorithm without memory can be cast as a rooted tree $\mathcal{T}$ of depth $T$ where each vertex encodes all of the classical measurement outcomes that have been obtained by the algorithm up until then. The tree $\mathcal{T}$ satisfies the following properties:*

1. *Each note $u$ has an associated n-qubit unnormalized state $\rho^S(u)$ corresponding to the current state of the system.*

2. *At the root $r$ of the tree, $\rho^S(r)$ is the initial state $\rho_0$.*

3. *At each node $u$ (except the root node) we apply a POVM measurement $\{F^u_s\}_s$ on $\rho(u)$ to obtain a classical outcome $s$. Without loss of generality we take all of the $F^u_s$'s to be rank one; if they are not, we can simply refine $\{F^u_s\}_s$ so that each of its elements is rank one. We also have a function $f^u$ which takes the index set of $\{F^u_s\}_s$ to the index set of $S = \{\mathcal{C}_i\}_i$. Then we apply the channel $\mathcal{C}_{f^u(s)}$ to the present state. Each child node $v$ of $u$ is connected through the edge $e_{u,s}$.*

4. *If $v$ is the child node of $u$ connected through the edge $e_{u,s}$, then*

$$\rho^S(v) := \mathcal{C}_{f^u(s)}[F^u_s \, \rho^S(u) \, (F^u_s)^\dagger]. \tag{H.7}$$

   *Here $F^u_s \, \rho^S(u) \, (F^u_s)^\dagger$ is the unnormalized post-measurement state, to which the channel $\mathcal{C}_{f^u(s)}$ is applied.*

5. *For any node $u$ at depth $t$ in the tree, $p^S(u) := tr(\rho^S(u))$ is the probability that the transcript of measurement outcomes observed by the learning algorithm after $t$ measurements is $u$. Moreover, $\rho^S(u)/p^S(u)$ is the state of the system at the node $u$.*

Using this definition, we can provide the following formalizations of Definitions 1 and 2 in Ch. 13:

**Definition 4** (Time-ordered experiment, formal). *A time-ordered experiment is a tree representation for learning a single unitary channel $\mathcal{U}$ without quantum memory.*

**Definition 5** (Out-of-time-order experiment, formal). *An out-of-time-order experiment is a tree representation for learning the collection of two unitary channels $\{\mathcal{U}, \mathcal{U}^\dagger\}$ without quantum memory, where the channels are inverses of one another.*

Suppose we have a tree representation $\mathcal{T}$ for learning a collection of channels without quantum memory, with depth $T$. Let its associated collection of channels be $S = \{\mathcal{C}_i\}_i$. Then the probability distribution over measurement outcomes is given by $p^S(\ell)$ where $\ell$ runs over the leafs of the tree. If instead we had a collection of channels $S' = \{\mathcal{C}'_i\}_i$ with the same index set as $S = \{\mathcal{C}_i\}_i$, then we could run $S'$ through the same learning tree protocol so that the probability distribution over measurement outcomes is now $p^{S'}(\ell)$. If we did not know if we were handed $S$ or $S'$, then Le Cam's two point method [519] implies that any post-processing algorithm we might use on our measurement data to distinguish between $S$ and $S'$ can succeed with a probability $p \geq 1/2$ only if

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^S(\ell) - p^{S'}(\ell)| \geq 2p - 1. \tag{H.8}$$

Circling back to Theorem 2, it is thus sufficient to show that for any time-ordered experiment corresponding to a learning tree $\mathcal{T}$ of depth $T$,

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |\mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)] - \mathbb{E}_{\mathcal{U}_1,\mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]| \leq \frac{2}{3} \tag{H.9}$$

for $T \leq \Omega(d^{1/4})$. This inequality would show that we cannot distinguish between the two ensembles with success probability $p \geq 5/6$ using fewer than $\Omega(d^{1/4})$ queries to the unknown unitary. Here $\mathbb{E}_{\mathcal{U}}$ denotes the Haar average over $U$ in the unitary channel $\mathcal{U}[\rho] = U\rho U^\dagger$, and similarly for $\mathbb{E}_{\mathcal{U}_1,\mathcal{U}_2}$ and $(\mathcal{U}_1 \otimes \mathcal{U}_2)[\rho] = (U_1 \otimes U_2)\rho(U_1^\dagger \otimes U_2^\dagger)$. We will prove the inequality (H.9) below.

## H.2 Main proofs

As explained above, we can reformulate Theorem 2 in the following manner:

**Theorem 4** (Equivalent to Theorem 2). *For any time-ordered experiment with learning tree $\mathcal{T}$ with depth $T \leq \Omega(d^{1/4})$, we have*

$$\frac{1}{2} \sum_{\ell \in leaf(\mathcal{T})} |\mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)] - \mathbb{E}_{\mathcal{U}_1,\mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]| \leq \frac{2}{3}. \tag{H.10}$$

*Proof.* Let $\mathcal{D}$ be the maximally depolarizing channel so that $p^{\mathcal{D}}(\ell) = 1/d^T$. Then using the triangle inequality,

$$\begin{aligned}
\frac{1}{2} &\sum_{\ell \in \text{leaf}(\mathcal{T})} |\mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)] - \mathbb{E}_{\mathcal{U}_1,\mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]| \\
&\leq \frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)]| + \frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}_1,\mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]|.
\end{aligned} \tag{H.11}$$

By Proposition 1 below, the first term on the right-hand side is less than or equal to $1/3$ for $T \leq \Omega(d^{1/3})$. Similarly, by Proposition 2 below the second term is less than or equal to $1/3$ for $T \leq \Omega(d^{1/4})$. This completes the proof. □

### Unitary channel versus maximally depolarizing channel

We begin by establishing more notation. Given a learning tree $\mathcal{T}$ of depth $T$, let $v_0, v_1, ..., v_T = \ell$ be a root-to-leaf path through the tree. This corresponds to having measured a sequence of POVM elements; let us denote them by $F_{v_1}, F_{v_2}, ..., F_{v_T}$. Without loss of generality these can be assumed to be rank one, as we explained previously. We can treat the last round (i.e. the $T$th round) differently than all of the others, since we do not need

to have a residual state after measuring. This allows us to replace $F_{v_T}$ by a bra $\langle\psi_{v_T}|$. Our $F_{v_i}$'s and $\langle\psi_{v_T}|$'s satisfy completeness relations, namely

$$\sum_{v \in \text{child}(v_{i-1})} F_v^\dagger F_v = \mathbb{1} \qquad \text{for } i = 1, ..., T-1, \tag{H.12}$$

and also

$$\sum_{v \in \text{child}(v_{T-1})} |\psi_v\rangle\langle\psi_v| = \mathbb{1}. \tag{H.13}$$

With this notation at hand, we can write $p^{\mathcal{U}}(\ell)$ as

$$p^{\mathcal{U}}(\ell) = \langle\psi_{v_T}|U F_{v_{T-1}} U \cdots U F_{v_2} \rho_0 F_{v_1}^\dagger U^\dagger \cdots U^\dagger F_{v_{T-1}}^\dagger U^\dagger |\psi_{v_T}\rangle. \tag{H.14}$$

For later, it will also be convenient to define $F_\ell$ by $F_\ell := F_{v_1} \otimes F_{v_2} \otimes \cdots \otimes F_{v_{T-1}}$.

With these preparations in order, we turn to our desired Proposition:

**Proposition 1.** *For $T \leq \Omega(d^{1/3})$, we have*

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)]| \leq \frac{1}{3}. \tag{H.15}$$

*Proof.* Using our Haar integration results from earlier, we have



$$\mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)] = \sum_{\sigma, \tau \in S_T} \qquad \text{Wg}^U(\tau\sigma^{-1}, d).$$

We let $p_{\sigma,\tau}(\ell)$ denote the summand of the above. Now we can upper bound the left-hand side of (H.15) using the triangle inequality and the Cauchy-Schwarz inequality as

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}}[p^{\mathcal{U}}(\ell)]| \leq \frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - p_{\mathbb{1},\mathbb{1}}(\ell)|$$

$$+ \frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} \sum_{\sigma \neq \mathbb{1}} |p_{\sigma,\mathbb{1}}(\ell)| + \frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} \sum_{\tau \neq \mathbb{1}, \sigma} |p_{\sigma,\tau}(\ell)|. \tag{H.16}$$

We will proceed by bounding each of the three terms on the right-hand side of (H.16) in turn.

**First term**

For the first term, we can apply Cauchy-Schwarz to find the upper bound



$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} \left| \; \right| \left| \text{Wg}^U(\mathbb{1}, d) - \frac{1}{d^T} \right|.$$

The absolute value in the first term can be removed since its argument is positive. Now we can explicitly sum over leafs to obtain

$$\frac{d^T}{2}\left|\mathrm{Wg}^U(\mathbb{1}, d) - \frac{1}{d^T}\right|.$$ 
(H.17)

But using Corollary 1 we have $\left|\mathrm{Wg}^U(\mathbb{1}, d) - \frac{1}{d^T}\right| \leq O(T^{7/2}/d^{T+2})$ for $T < \left(\frac{d}{\sqrt{6}}\right)^{4/7}$, and so in total

$$\frac{1}{2}\sum_{\ell \in \mathrm{leaf}(\mathcal{T})}|p^{\mathcal{D}}(\ell) - p_{\mathbb{1},\mathbb{1}}(\ell)| \leq O\left(\frac{T^{7/2}}{d^2}\right).$$ 
(H.18)

**Second term**

Applying Cauchy-Schwarz to the second term on the right-hand side of (H.16), we have the upper bound


(H.19)

The first term can be upper bounded using Hölder's inequality



and we further use the equality


(H.20)

This follows from the fact that $\|A \otimes B\|_1 = \|A\|_1 \|B\|_1 = \mathrm{tr}(A)\,\mathrm{tr}(B)$ if $A$ and $B$ are positive semi-definite. Then (H.19) is upper bounded by


(H.21)

and so summing over leafs we obtain

$$\frac{d^T}{2} \sum_{\sigma \neq \mathbb{1}} |\mathrm{Wg}^U(\sigma^{-1}, d)| \,. \tag{H.22}$$

But this quantity is less than or equal to $O(T^2/d)$ using Lemma 6 of [10]; thus we summarily find

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\sigma \neq \mathbb{1}} |p_{\sigma,\mathbb{1}}(\ell)| \leq O\!\left(\frac{T^2}{d}\right) \,. \tag{H.23}$$

**Third term**

As usual, we apply the Cauchy-Schwarz inequality to the last term on the right-hand side of (H.16) to obtain

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\substack{\tau \neq \mathbb{1} \\ \sigma}} \left| \,\begin{matrix} \rho_0 \\ \sigma \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T} | \!-\! \tau^{-1} \!-\! | \psi_{v_T} \rangle \end{matrix}\, \right| |\mathrm{Wg}^U(\tau\sigma^{-1}, d)| \,. \tag{H.24}$$

Similar to the previous case, we apply Hölder's inequality to the diagrammatic term in the summand as

$$\left| \,\begin{matrix} \rho_0 \\ \sigma \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T} | \!-\! \tau^{-1} \!-\! | \psi_{v_T} \rangle \end{matrix}\, \right| \leq \left\| \,\begin{matrix} \rho_0 \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T} | \!-\! \tau^{-1} \!-\! | \psi_{v_T} \rangle \end{matrix}\, \right\|_1 \|\sigma\|_\infty \tag{H.25}$$

noting again that $\|\sigma\|_\infty = 1$. The other 1-norm term further simplifies to

$$\left\| \,\begin{matrix} \rho_0 \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T} | \!-\! \tau^{-1} \!-\! | \psi_{v_T} \rangle \end{matrix}\, \right\|_1 = \left\| \,\begin{matrix} F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T} | \!-\! \tau^{-1} \!-\! | \psi_{v_T} \rangle \end{matrix}\, \right\|_1 \tag{H.26}$$

where we have used $\|A \otimes B\|_1 = \|A\|_1 \|B\|_1$, where in the above setting $\|B\|_1 = 1$. It is convenient to simplify the remaining 1-norm term for fixed $\tau^{-1}$. To do so, we decompose $\tau^{-1}$ into cycles as $\tau^{-1} = C_1 C_2 \cdots C_{\#(\tau^{-1})}$, and will say that $i \to j$ belongs to the $m$th cycle $C_m$ if $C_m = (\cdots ij \cdots)$. More generally we also say that $i \to j$ belongs to $\tau^{-1}$. Further letting $v_0 = r, v_1, ..., v_{T-1}, v_T = \ell$ be the root-to-leaf path terminating in $\ell$, we leverage the following lemma:

**Lemma 4.** *If $\tau^{-1}$ contains the 1-cycle $T \to T$ then*

$$\left\| \langle \psi_{v_T} | \begin{array}{c} F_\ell^\dagger \quad F_\ell \\ \tau^{-1} \end{array} | \psi_{v_T} \rangle \right\|_1 = \left( \prod_{\substack{i \to j \,\in\, \tau^{-1} \\ i,j \neq T}} \left\| F_{v_j} \quad F_{v_i}^\dagger \right\|_1 \right) \langle \psi_{v_T} | \psi_{v_T} \rangle. \tag{H.27}$$

*Otherwise if $\tau^{-1}$ does not contain the 1-cycle $T \to T$, then it must contain some $\hat{i} \to T \to \hat{j}$ (where possibly $\hat{i} = \hat{j}$) in which case*

$$\left\| \langle \psi_{v_T} | \begin{array}{c} F_\ell^\dagger \quad F_\ell \\ \tau^{-1} \end{array} | \psi_{v_T} \rangle \right\|_1 = \left( \prod_{\substack{i \to j \,\in\, \tau^{-1} \\ i,j \neq T}} \left\| F_{v_j} \quad F_{v_i}^\dagger \right\|_1 \right) \left\| F_{v_{\hat{j}}} \, | \psi_{v_T} \rangle \langle \psi_{v_T} | \, F_{v_{\hat{i}}}^\dagger \right\|_1. \tag{H.28}$$

*Proof.* The identities follow by contracting tensor indices of $F_\ell$, $F_\ell^\dagger$, $\langle \psi_{v_T} |$ and $|\psi_{v_T}\rangle$ according to $\tau^{-1}$, and then using the identity $\|A_1 \otimes A_2 \otimes \cdots \otimes A_n\|_1 = \prod_{i=1}^{n} \|A_i\|_1$. $\square$

Next we simplify the 1-norm terms appearing in (H.27) and (H.28). Since $\|A\|_1 = \|A\|_2$ when $A$ is rank one, we have

$$\left\| F_{v_j} \quad F_{v_i}^\dagger \right\|_1 = \left( F_{v_j} \quad F_{v_i}^\dagger \quad F_{v_i} \quad F_{v_j}^\dagger \right)^{1/2}. \tag{H.29}$$

Equivalently, this is

$$\left\| F_{v_j} \quad F_{v_i}^\dagger \right\|_1 = \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})}. \tag{H.30}$$

Following the same logic, we obtain the equality

$$\left\| F_{v_{\hat{j}}} \, | \psi_{v_T} \rangle \langle \psi_{v_T} | \, F_{v_{\hat{i}}}^\dagger \right\|_1 = \sqrt{\mathrm{tr}(|\psi_{v_T}\rangle\langle\psi_{v_T}| \cdot F_{v_{\hat{i}}}^\dagger F_{v_{\hat{i}}})} \sqrt{\mathrm{tr}(|\psi_{v_T}\rangle\langle\psi_{v_T}| \cdot F_{v_{\hat{j}}}^\dagger F_{v_{\hat{j}}})}. \tag{H.31}$$

Accordingly, (H.27) is equal to

$$\left( \prod_{\substack{i \to j \in \tau^{-1} \\ i,j \neq T}} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})} \right) \langle \psi_{v_T} | \psi_{v_T} \rangle \tag{H.32}$$

and likewise (H.28) is equal to

$$\left( \prod_{\substack{i \to j \in \tau^{-1} \\ i,j \neq T}} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})} \right) \sqrt{\mathrm{tr}(|\psi_{v_T}\rangle\langle\psi_{v_T}| \cdot F_{v_{\hat{i}}}^\dagger F_{v_{\hat{i}}})} \sqrt{\mathrm{tr}(|\psi_{v_T}\rangle\langle\psi_{v_T}| \cdot F_{v_{\hat{j}}}^\dagger F_{v_{\hat{j}}})} . \quad (\text{H}.33)$$

To make the two cases look the same, we simply define $F_{v_T} := |\psi_{v_T}\rangle\langle\psi_{v_T}|$ so that both (H.32) and (H.33) can be written as

$$\prod_{i \to j \in \tau^{-1}} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})} . \quad (\text{H}.34)$$

To further bound (H.34), we decompose the product into cycles as

$$\prod_{m=1}^{\#(\tau^{-1})} \prod_{i \to j \in C_m} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})} \quad (\text{H}.35)$$

and treat each $\prod_{i \to j \in C_m} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})}$ term separately.

Suppose that $C_m$ is a cycle of length $p$ (often denoted by $|C_m| = p$); then we can write it as $C_m = (a_{m,1} a_{m,2} \cdots a_{m,p})$ where $\{a_{m,1}, a_{m,2}, ..., a_{m,p}\} \subseteq \{1, 2, ..., T\}$. Then we can write $\prod_{i \to j \in C_m} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})}$ as

$$\prod_{i=1}^{p} \sqrt{\mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}})} \quad (\text{H}.36)$$

where the $i$ subscripts are treated modulo $p$. Consider two cases:

**Case 1:** *$p$ is even.* We can split up (H.36) into two products as

$$\left( \prod_{i \text{ odd}} \sqrt{\mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}})} \right) \left( \prod_{j \text{ even}} \sqrt{\mathrm{tr}(F_{v_{a_{m,j}}}^\dagger F_{v_{a_{m,j}}} \cdot F_{v_{a_{m,j+1}}}^\dagger F_{v_{a_{m,j+1}}})} \right)$$
$$(\text{H}.37)$$

and using the inequality $ab \leq \frac{1}{2}(a^2 + b^2)$ we obtain the upper bound

$$\frac{1}{2} \prod_{i \text{ odd}} \mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}}) + \frac{1}{2} \prod_{j \text{ even}} \mathrm{tr}(F_{v_{a_{m,j}}}^\dagger F_{v_{a_{m,j}}} \cdot F_{v_{a_{m,j+1}}}^\dagger F_{v_{a_{m,j+1}}}) . \quad (\text{H}.38)$$

Let us define the first term as $\frac{1}{2} R_{m,-}$ and the second term as $\frac{1}{2} R_{m,+}$.

**Case 2:** *p is odd.* Here we opt to split up (H.36) as

$$
\sqrt{\mathrm{tr}(F_{v_{a_{m,p}}}^\dagger F_{v_{a_{m,p}}} \cdot F_{v_{a_{m,1}}}^\dagger F_{v_{a_{m,1}}})} \left( \prod_{\substack{i \text{ odd} \\ 1 \le i \le p-2}} \sqrt{\mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}})} \right)
$$

$$
\times \left( \prod_{j \text{ even}} \sqrt{\mathrm{tr}(F_{v_{a_{m,j}}}^\dagger F_{v_{a_{m,j}}} \cdot F_{v_{a_{m,j+1}}}^\dagger F_{v_{a_{m,j+1}}})} \right) \tag{H.39}
$$

For $A, B$ positive semi-definite we have $\mathrm{tr}(AB) \le \|A\|_2 \|B\|_2 \le \|A\|_1 \|B\|_1 \le \mathrm{tr}(A)\mathrm{tr}(B)$ and so $\sqrt{\mathrm{tr}(F_{v_{a_{m,p}}}^\dagger F_{v_{a_{m,p}}} \cdot F_{v_{a_{m,1}}}^\dagger F_{v_{a_{m,1}}})} \le \sqrt{\mathrm{tr}(F_{v_{a_{m,p}}}^\dagger F_{v_{a_{m,p}}})} \sqrt{\mathrm{tr}(F_{v_{a_{m,1}}}^\dagger F_{v_{a_{m,1}}})}$. Then the above equation is upper bounded by

$$
\sqrt{\mathrm{tr}(F_{v_{a_{m,p}}}^\dagger F_{v_{a_{m,p}}})} \sqrt{\mathrm{tr}(F_{v_{a_{m,1}}}^\dagger F_{v_{a_{m,1}}})} \left( \prod_{\substack{i \text{ odd} \\ 1 \le i \le p-2}} \sqrt{\mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}})} \right)
$$

$$
\times \left( \prod_{j \text{ even}} \sqrt{\mathrm{tr}(F_{v_{a_{m,j}}}^\dagger F_{v_{a_{m,j}}} \cdot F_{v_{a_{m,j+1}}}^\dagger F_{v_{a_{m,j+1}}})} \right) \tag{H.40}
$$

and so using $ab \le \frac{1}{2}(a^2 + b^2)$ we have the further upper bound

$$
\frac{1}{2} \mathrm{tr}(F_{v_{a_{m,p}}}^\dagger F_{v_{a_{m,p}}}) \prod_{\substack{i \text{ odd} \\ 1 \le i \le p-2}} \mathrm{tr}(F_{v_{a_{m,i}}}^\dagger F_{v_{a_{m,i}}} \cdot F_{v_{a_{m,i+1}}}^\dagger F_{v_{a_{m,i+1}}})
$$

$$
+ \frac{1}{2} \mathrm{tr}(F_{v_{a_{m,1}}}^\dagger F_{v_{a_{m,1}}}) \prod_{j \text{ even}} \mathrm{tr}(F_{v_{a_{m,j}}}^\dagger F_{v_{a_{m,j}}} \cdot F_{v_{a_{m,j+1}}}^\dagger F_{v_{a_{m,j+1}}}). \tag{H.41}
$$

We similarly call the first term $\frac{1}{2} R_{m,-}$ and the second term $\frac{1}{2} R_{m,+}$.

Taken together, Case 1 and Case 2 give us the following bound on (H.35):

$$
\prod_{m=1}^{\#(\tau^{-1})} \prod_{i \to j \in C_m} \sqrt{\mathrm{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j})} \le \frac{1}{2^{\#(\tau^{-1})}} \prod_{m=1}^{\#(\tau^{-1})} (R_{m,-} + R_{m,+})
$$

$$
= \frac{1}{2^{\#(\tau^{-1})}} \sum_{i_1,\ldots,i_{\#(\tau^{-1})}=\pm} R_{1,i_1} R_{2,i_2} \cdots R_{\#(\tau^{-1}),i_{\#(\tau^{-1})}} \cdot
$$

$$
\tag{H.42}
$$

Since the $R_{m,\pm}$'s depend implicitly on the leaf $\ell$, we add an $\ell$ superscript as $R_{m,\pm}^\ell$ to make the dependence explicit. The summand $R_{1,i_1}^\ell R_{2,i_2}^\ell \cdots R_{\#(\tau^{-1}),i_{\#(\tau^{-1})}}^\ell$ for fixed indices $i_1, i_2, \ldots, i_{\#(\tau^{-1})}$ has the feature that each $F_{v_i}^\dagger F_{v_i}$ for $i = 1, \ldots, T$ appears exactly once. By virtue of this fact we can establish the following lemma:

**Lemma 5.** *For any fixed set of indices $i_1, i_2, ..., i_{\#(\tau^{-1})} \in \{+, -\}$, we have*

$$\sum_{\ell \in \text{leaf}(\mathcal{T})} R^\ell_{1,i_1} R^\ell_{2,i_2} \cdots R^\ell_{\#(\tau^{-1}), i_{\#(\tau^{-1})}} \le d^{T - \left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor} \tag{H.43}$$

*where $L(\tau^{-1})$ is the length of the longest cycle in $\tau^{-1}$.*

*Proof.* We have the identities

$$\sum_{v \in \text{child}(v_{i-1})} \text{tr}(F_v^\dagger F_v \cdot F_{v_j}^\dagger F_{v_j}) = \text{tr}(F_{v_j}^\dagger F_{v_j}) \tag{H.44}$$

$$\sum_{v \in \text{child}(v_{i-1})} \text{tr}(F_v^\dagger F_v) = d. \tag{H.45}$$

In a slight abuse of notation, we rewrite these as

$$\sum_{v_i} \text{tr}(F_{v_i}^\dagger F_{v_i} \cdot F_{v_j}^\dagger F_{v_j}) = \text{tr}(F_{v_j}^\dagger F_{v_j}) \tag{H.46}$$

$$\sum_{v_i} \text{tr}(F_{v_i}^\dagger F_{v_i}) = d. \tag{H.47}$$

For fixed $i_1, i_2, ..., i_{\#(\tau^{-1})} \in \{+, -\}$, we have

$$R^\ell_{1,i_1} R^\ell_{2,i_2} \cdots R^\ell_{\#(\tau^{-1}), i_{\#(\tau^{-1})}} = \left( \prod_{i \in \mathcal{S}_1} \text{tr}(F_{v_i}^\dagger F_{v_i}) \right) \left( \prod_{(j,j') \in \mathcal{S}_2} \text{tr}(F_{v_j}^\dagger F_{v_j} \cdot F_{v_{j'}}^\dagger F_{v_{j'}}) \right) \tag{H.48}$$

where $\mathcal{S}_1 \subset \{1, ..., T\}$ is the set of indices for which a $\text{tr}(F_{v_i}^\dagger F_{v_i})$ term appears, and $\mathcal{S}_2 \subset \{1, ..., T\} \times \{1, ..., T\}$ is the set of unordered pairs $(j, j')$ for which a $\text{tr}(F_{v_j}^\dagger F_{v_j} \cdot F_{v_{j'}}^\dagger F_{v_{j'}})$ term appears. Note that the size of $\mathcal{S}_1$ is the number of odd-length cycles of $\tau^{-1}$. As noted above, each $v_i$ for $i = 1, ..., T$ appears exactly once in the above expression. Writing the (H.43) as

$$\sum_{v_1} \cdots \sum_{v_{T-1}} \sum_{v_T} \left( \prod_{i \in \mathcal{S}_1} \text{tr}(F_{v_i}^\dagger F_{v_i}) \right) \left( \prod_{(j,j') \in \mathcal{S}_2} \text{tr}(F_{v_j}^\dagger F_{v_j} \cdot F_{v_{j'}}^\dagger F_{v_{j'}}) \right), \tag{H.49}$$

we can perform the inner-most sum over $v_T$ following by the $v_{T-1}$ sum, and so on through the $v_1$ sum. That is, we are summing from the leafs of the tree back up to the root; this order of summation is necessitated because of the adaptive nature of the measurement strategies that we allow. That is, the choice of measurements in the future (i.e. higher depth in the learning tree) depend on measurements made in the past (i.e. lower depth in the learning tree), but not conversely. Leveraging the identities (H.46), (H.47) and the equality $\lceil \frac{x}{2} \rceil + \lfloor \frac{x}{2} \rfloor = x$ for integer $x$, we find that the sum equals

$$d^{\sum_{m=1}^{\#(\tau^{-1})} \lceil \frac{|C_m|}{2} \rceil} = d^{T - \sum_{m=1}^{\#(\tau^{-1})} \lfloor \frac{|C_m|}{2} \rfloor} \le d^{T - \left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor} \tag{H.50}$$

giving the desired bound. $\qquad \square$

An immediate consequence of the above Lemma is that

$$\sum_{\ell \in \text{leaf}(\mathcal{T})} \left( \frac{1}{2^{\#(\tau^{-1})}} \sum_{i_1,\dots,i_{\#(\tau^{-1})}=\pm} R^\ell_{1,i_1} R^\ell_{2,i_2} \cdots R^\ell_{\#(\tau^{-1}),i_{\#(\tau^{-1})}} \right)$$

$$= \frac{1}{2^{\#(\tau^{-1})}} \sum_{i_1,\dots,i_{\#(\tau^{-1})}=\pm} \left( \sum_{\ell \in \text{leaf}(\mathcal{T})} R^\ell_{1,i_1} R^\ell_{2,i_2} \cdots R^\ell_{\#(\tau^{-1}),i_{\#(\tau^{-1})}} \right)$$

$$\leq \frac{1}{2^{\#(\tau^{-1})}} \sum_{i_1,\dots,i_{\#(\tau^{-1})}=\pm} d^{T - \left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor}$$

$$= d^{T - \left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor}. \tag{H.51}$$

Circling back to (H.24), we can combine our bounds to obtain

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} \sum_{\substack{\tau \neq \mathbb{1} \\ \sigma}} |p_{\sigma,\tau}(\ell)| \leq \frac{d^T}{2} \sum_{\sigma} |\text{Wg}^U(\sigma^{-1},d)| \sum_{\tau \neq \mathbb{1}} d^{-\left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor}. \tag{H.52}$$

To bound the right-hand side, we can use $d^T \sum_{\sigma} |\text{Wg}^U(\sigma^{-1},d)| \leq 1 + O(T^2/d)$. Letting $N(T,L)$ denote the number of permutations of $S_T$ whose longest cycle has length $L$, we can write

$$\sum_{\tau \neq \mathbb{1}} d^{-\left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor} = \sum_{L=2}^{T} N(T,L)\, d^{-\left\lfloor \frac{L}{2} \right\rfloor}. \tag{H.53}$$

Here the $L = 1$ case is omitted since this corresponds to the identity permutation. Since $N(T,L) \leq \binom{T}{L} L! = \frac{T!}{(T-L)!} < T^L$, the above sum is upper bounded by

$$\sum_{L=2}^{\infty} T^L d^{-\left\lfloor \frac{L}{2} \right\rfloor} = \frac{(1+T)\frac{T^2}{d}}{1 - \frac{T^2}{d}} = \frac{T^3}{d} + \frac{T^2}{d} + O\left( \frac{T^5}{d^2} \right). \tag{H.54}$$

In summary, if $T \leq o(d^{1/3})$ we have

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} \sum_{\substack{\tau \neq \mathbb{1} \\ \sigma}} |p_{\sigma,\tau}(\ell)| \leq o(1), \tag{H.55}$$

as needed.

Combining the three cases, we find that for $T \leq \Omega(d^{1/3})$ we have

$$\frac{1}{2} \sum_{\ell \in \text{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - (\mathbb{E}_{\mathcal{U}}\, p^{\mathcal{U}}(\ell))| \leq \frac{1}{3} \tag{H.56}$$

which completes the proof. $\qquad \square$

## Product unitary channel versus maximally depolarizing channel

For the next proposition we use the same notations and conventions as we did above.

**Proposition 2.** *For $T \leq \Omega(d^{1/4})$, we have*

$$\frac{1}{2} \sum_{\ell \in leaf(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}_1, \mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]| \leq \frac{1}{3}. \tag{H.57}$$

*Proof.* As before $p^{\mathcal{D}}(\ell) = 1/d^T$, and now we have

$$\mathbb{E}_{\mathcal{U}_1, \mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)] = \sum_{\sigma_1, \sigma_2, \tau_1, \tau_2} \begin{gathered}</gathered} \mathrm{Wg}^U(\tau_1 \sigma_1^{-1}, d^{1/2})\, \mathrm{Wg}^U(\tau_2 \sigma_2^{-1}, d^{1/2}). \tag{H.58}$$

We let $p_{\sigma_1, \sigma_2, \tau_1, \tau_2}(\ell)$ denote the summand of the above. The left-hand side of (H.57) can be upper bounded via the triangle and Cauchy-Schwarz inequalities as

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}_1, \mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]|$$

$$\leq \frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - p_{\mathbb{1}, \mathbb{1}, \mathbb{1}, \mathbb{1}}(\ell)| + \frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\sigma_1 \otimes \sigma_2 \neq \mathbb{1} \otimes \mathbb{1}} |p_{\sigma_1, \sigma_2, \mathbb{1}, \mathbb{1}}(\ell)|$$

$$+ \frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\substack{\tau_1 \otimes \tau_2 \neq \mathbb{1} \otimes \mathbb{1} \\ \sigma_1, \sigma_2}} |p_{\sigma_1, \sigma_2, \tau_1, \tau_2}(\ell)|. \tag{H.59}$$

Similar to the previous Proposition, we will individually bound each term on the right-hand side of (H.59).

**First term**

Applying the Cauchy-Schwarz inequality we have the upper bound

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \left| \begin{gathered}</gathered} \right| \left| \mathrm{Wg}^U(\mathbb{1}, d^{1/2}) \mathrm{Wg}^U(\mathbb{1}, d^{1/2}) - \frac{1}{d^T} \right|. \tag{H.60}$$

Since the argument of the first term is positive, we remove the absolute values; this allows us to explicitly sum over leafs to obtain

$$\frac{d^T}{2}\left|\mathrm{Wg}^U(\mathbb{1}, d^{1/2})\mathrm{Wg}^U(\mathbb{1}, d^{1/2}) - \frac{1}{d^T}\right|. \tag{H.61}$$

The appropriate version of Corollary 1 gives us $\left|\mathrm{Wg}^U(\mathbb{1}, d^{1/2}) - \frac{1}{d^{T/2}}\right| \leq O(T^{7/2}/d^{T/2+1})$ for $T < \left(\frac{d}{6}\right)^{2/7}$, and thus

$$\frac{1}{2}\sum_{\ell \in \mathrm{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - p_{\mathbb{1},\mathbb{1},\mathbb{1},\mathbb{1}}(\ell)| \leq O\left(\frac{T^{7/2}}{d}\right). \tag{H.62}$$

**Second term**

We can apply the Cauchy-Schwarz inequality to the second term on the right-hand side of (H.59) to get the upper bound

$$\frac{1}{2}\sum_{\ell \in \mathrm{leaf}(\mathcal{T})}\sum_{\sigma_1 \otimes \sigma_2 \neq \mathbb{1} \otimes \mathbb{1}} \left| \begin{array}{c} \rho_0 \\ \sigma_1 \\ \sigma_2 \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T}| \quad |\psi_{v_T}\rangle \end{array} \right| |\mathrm{Wg}^U(\sigma_1^{-1}, d^{1/2})|\,|\mathrm{Wg}^U(\sigma_2^{-1}, d^{1/2})| \tag{H.63}$$

The above can be bounded in the same manner as (H.19) in Proposition 1; the proof is the same up through (H.20). Then the analog of (H.21) is

$$\frac{1}{2}\sum_{\ell \in \mathrm{leaf}(\mathcal{T})}\sum_{\sigma_1 \otimes \sigma_2 \neq \mathbb{1} \otimes \mathbb{1}} \begin{array}{c} \rho_0 \\ F_\ell \quad F_\ell^\dagger \\ \langle \psi_{v_T}| \quad |\psi_{v_T}\rangle \end{array} |\mathrm{Wg}^U(\sigma_1^{-1}, d^{1/2})|\,|\mathrm{Wg}^U(\sigma_2^{-1}, d^{1/2})|. \tag{H.64}$$

and summing over leafs we find

$$\frac{d^T}{2}\sum_{\sigma_1 \otimes \sigma_2 \neq \mathbb{1} \otimes \mathbb{1}} |\mathrm{Wg}^U(\sigma_1^{-1}, d^{1/2})|\,|\mathrm{Wg}^U(\sigma_2^{-1}, d^{1/2})|. \tag{H.65}$$

Denoting $\sigma = \sigma_1 \otimes \sigma_2$, the above is equal to

$$2 \cdot \frac{d^{T/2}}{2}|\mathrm{Wg}^U(\mathbb{1}, d^{1/2})|\left(d^{T/2}\sum_{\substack{\sigma \neq \mathbb{1} \\ \sigma \in S_T}} |\mathrm{Wg}^U(\sigma^{-1}, d^{1/2})|\right) + \frac{1}{2}\left(d^{T/2}\sum_{\substack{\sigma \neq \mathbb{1} \\ \sigma \in S_T}} |\mathrm{Wg}^U(\sigma^{-1}, d^{1/2})|\right)^2. \tag{H.66}$$

Since $|\mathrm{Wg}^U(\mathbb{1}, d^{1/2})| = \frac{1}{d^{T/2}} + O(T^{7/2}/d^{T/2+1})$ for $T < \left(\frac{d}{6}\right)^{2/7}$ and the term in the parentheses is less than or equal to $O(T^2/d^{1/2})$ by Lemma 6 of [10], we have in total

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\sigma_1 \otimes \sigma_2 \neq \mathbb{1} \otimes \mathbb{1}} |p_{\sigma_1, \sigma_2, \mathbb{1}, \mathbb{1}}(\ell)| \leq O\left(\frac{T^2}{d^{1/2}}\right). \tag{H.67}$$

**Third term**

Applying the Cauchy-Schwarz inequality to final term in (H.59) we find

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\substack{\tau_1 \otimes \tau_2 \neq \mathbb{1} \otimes \mathbb{1} \\ \sigma_1 \otimes \sigma_2}} \left| \begin{array}{c} \boxed{\rho_0} \\ \boxed{\sigma_1} \\ \boxed{\sigma_2} \\ \boxed{F_\ell} \quad \boxed{F_\ell^\dagger} \\ \langle \psi_{v_T} | \boxed{\tau_1^{-1}} | \psi_{v_T} \rangle \\ \boxed{\tau_2^{-1}} \end{array} \right| |\mathrm{Wg}^U(\tau_1 \sigma_1^{-1}, d^{1/2})| \, |\mathrm{Wg}^U(\tau_2 \sigma_2^{-1}, d^{1/2})|. \tag{H.68}$$

If we label $\sigma = \sigma_1 \otimes \sigma_2$ and $\tau = \tau_1 \otimes \tau_2$, the proof proceeds identically to the third case of Proposition 1 up through (H.51). Then the new analog of (H.52) is

$$\sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\substack{\tau_1 \otimes \tau_2 \neq \mathbb{1} \otimes \mathbb{1} \\ \sigma_1, \sigma_2}} |p_{\sigma_1, \sigma_2, \tau_1, \tau_2}(\ell)|$$

$$\leq d^T \sum_{\sigma_1, \sigma_2} |\mathrm{Wg}^U(\sigma_1^{-1}, d^{1/2})| \, |\mathrm{Wg}^U(\sigma_2^{-1}, d^{1/2})| \sum_{\tau_1 \otimes \tau_2 \neq \mathbb{1} \otimes \mathbb{1}} d^{-\left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor}. \tag{H.69}$$

The right-hand side can be bounded in part using $d^{T/2} \sum_\sigma |\mathrm{Wg}^U(\sigma^{-1}, d^{1/2})| \leq 1 + O(T^2/d^{1/2})$. Since

$$\sum_{\substack{\tau_1 \otimes \tau_2 \neq \mathbb{1} \otimes \mathbb{1} \\ \tau_1, \tau_2 \in S_T}} d^{-\left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor} \leq \sum_{\substack{\tau \neq \mathbb{1} \\ \tau \in S_T}} d^{-\left\lfloor \frac{L(\tau^{-1})}{2} \right\rfloor} \leq \frac{T^3}{d} + \frac{T^2}{d} + O\left(\frac{T^5}{d^2}\right) \tag{H.70}$$

where the last bound comes from Proposition 1, we find that if $T \leq o(d^{1/4})$ then

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} \sum_{\substack{\tau_1 \otimes \tau_2 \neq \mathbb{1} \\ \sigma_1, \sigma_2}} |p_{\sigma_1, \sigma_2, \tau_1, \tau_2}(\ell)| \leq o(1). \tag{H.71}$$

Putting all three cases together, we see that for $T \leq \Omega(d^{1/4})$ we have

$$\frac{1}{2} \sum_{\ell \in \mathrm{leaf}(\mathcal{T})} |p^{\mathcal{D}}(\ell) - \mathbb{E}_{\mathcal{U}_1, \mathcal{U}_2}[p^{\mathcal{U}_1 \otimes \mathcal{U}_2}(\ell)]| \leq \frac{1}{3}. \tag{H.72}$$

$\square$

# Bibliography

[1] Our parallelized dynamics code is available open-source as the package `dynamite`: `https://dynamite.readthedocs.io/` DOI:10.5281/zenodo.3606826.

[2] Scott Aaronson. "Shadow tomography of quantum states". In: *SIAM Journal on Computing* 49.5 (2019), STOC18–368.

[3] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". In: *Physical Review A* 70.5 (2004), p. 052328.

[4] David Aasen et al. "Milestones Toward Majorana-Based Quantum Computing". In: *Physical Review X* 6 (3 2016), p. 031016.

[5] Dmitry A. Abanin, Wojciech De Roeck, and Fran çois Huveneers. "Exponentially Slow Heating in Periodically Driven Many-Body Systems". In: *Physical Review Letters* 115 (25 2015), p. 256803. DOI: `10.1103/PhysRevLett.115.256803`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.115.256803`.

[6] Dmitry A Abanin and Eugene Demler. "Measuring entanglement entropy of a generic many-body system with a quantum switch". In: *Physical Review Letters* 109.2 (2012), p. 020504.

[7] Paul J Ackerman and Ivan I Smalyukh. "Static three-dimensional topological solitons in fluid chiral ferromagnets and colloids". In: *Nature Materials* 16.4 (2017), p. 426.

[8] Kartiek Agarwal and Ning Bao. "Toy model for decoherence in the black hole information problem". In: *Physical Review D* 102.8 (2020), p. 086017.

[9] Adhip Agarwala. "Topological insulators in amorphous systems". In: *Excursions in Ill-Condensed Quantum Matter*. Springer, 2019, pp. 61–79.

[10] Dorit Aharonov, Jordan Cotler, and Xiao-Liang Qi. "Quantum algorithmic measurement". In: *arXiv:2101.04634* (2021).

[11] Dorit Aharonov et al. "A polynomial-time classical algorithm for noisy random circuit sampling". In: *arXiv:2211.03999* (2022).

[12] Monika Aidelsburger et al. "Measuring the Chern number of Hofstadter bands with ultracold bosonic atoms". In: *Nature Physics* 11.2 (2015), pp. 162–166.

[13] K Aikawa et al. "Bose-Einstein condensation of erbium". In: *Physical Review Letters* 108.21 (2012), p. 210401.

[14] Francesco Albarelli et al. "Resource theory of quantum non-Gaussianity and Wigner negativity". In: *Physical Review A* 98 (5 2018), p. 052350. DOI: `10.1103/PhysRevA.98.052350`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.98.052350`.

[15] J Aldegunde, Hong Ran, and Jeremy M Hutson. "Manipulating ultracold polar molecules with microwave radiation: The influence of hyperfine structure". In: *Physical Review A* 80.4 (2009), p. 043410.

[16] J Aldegunde et al. "Hyperfine energy levels of alkali-metal dimers: Ground-state polar molecules in electric and magnetic fields". In: *Physical Review A* 78.3 (2008), p. 033434.

[17] Igor L Aleiner, Lara Faoro, and Lev B Ioffe. "Microscopic model of quantum butterfly effect: out-of-time-order correlators and traveling combustion waves". In: *Annals of Physics* 375 (2016), pp. 378–406.

[18] Rafael N Alexander, Glen Evenbly, and Israel Klich. "Exact holographic tensor networks for the Motzkin spin chain". In: *arXiv:1806.09626* (2018).

[19] A Alexandradinata, Aleksandra Nelson, and Alexey A Soluyanov. "The actually robust surface signature of a Hopf insulator: Bulk-to-boundary flow of Berry curvature beyond the anomaly inflow paradigm". In: *arXiv:1910.10717* (2019).

[20] Ahmed Almheiri et al. "Black holes: complementarity or firewalls?" In: *J. High Energy Phys.* 02 (2013), p. 62.

[21] José Raúl González Alonso, Nicole Yunger Halpern, and Justin Dressel. "Out-of-time-ordered-correlator quasiprobabilities robustly witness scrambling". In: *Physical Review Letters* 122.4 (2019), p. 040404.

[22] K. Alps et al. "Fourier-transform spectroscopy and deperturbation analysis of the spin-orbit coupled A1Îč+ and b3Îă states of KRb". In: *The Journal of Chemical Physics* 144.14 (2016), p. 144310.

[23] Ehud Altman et al. "Quantum simulators: Architectures and opportunities". In: *PRX Quantum* 2.1 (2021), p. 017003.

[24] A. Ambainis and J. Emerson. "Quantum t-designs: t-wise Independence in the Quantum World". In: *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)* (2007), pp. 129–140.

[25] Sanghun An et al. "Braiding of Abelian and non-Abelian anyons in the fractional quantum Hall effect". In: *arXiv:1112.3400* (2011).

[26] Loïc Anderegg et al. "An optical tweezer array of ultracold molecules". In: *Science* 365.6458 (2019), pp. 1156–1158.

[27] Philip W Anderson. "The resonating valence bond state in La2CuO4 and superconductivity". In: *science* 235.4793 (1987), pp. 1196–1198.

[28] Faidon Andreadakis, Namit Anand, and Paolo Zanardi. "Scrambling of Algebras in Open Quantum Systems". In: *arXiv:2206.02033* (2022).

[29] Daniel Arovas, J. R. Schrieffer, and Frank Wilczek. "Fractional Statistics and the Quantum Hall Effect". In: *Physical Review Letters* 53 (7 1984), pp. 722–723.

[30] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510.

[31] J. K. Asbóth, B. Tarasinski, and P. Delplace. "Chiral symmetry and bulk-boundary correspondence in periodically driven one-dimensional systems". In: *Physical Review B - Condensed Matter and Materials Physics* 90.12 (2014), pp. 1–7. ISSN: 1550235X. DOI: `10.1103/PhysRevB.90.125143`. eprint: `1405.1709`.

[32] Eyal Bairey, Itai Arad, and Netanel H Lindner. "Learning a local Hamiltonian from local measurements". In: *Physical Review Letters* 122.2 (2019), p. 020504.

[33] T. Bakkegaard et al. "Realization of efficient quantum gates with a superconducting qubit-qutrit circuit". en. In: *Scientific Reports* 9.1 (Sept. 2019). Number: 1 Publisher: Nature Publishing Group, p. 13389.

[34] N Balakrishnan. "Perspective: Ultracold molecules and the dawn of cold controlled chemistry". In: *The Journal of Chemical Physics* 145.15 (2016), p. 150901.

[35] C. J. Ballance et al. "High-Fidelity Quantum Logic Gates Using Trapped-Ion Hyperfine Qubits". In: *Physical Review Letters* 117 (6 2016), p. 060504. DOI: `10.1103/PhysRevLett.117.060504`.

[36] Masashi Ban, Masahide Sasaki, and Masahiro Takeoka. "Continuous variable teleportation as a generalized thermalizing quantum channel". In: *J. Phys. A: Math. Gen.* 35.28 (2002), p. L401.

[37] Ning Bao and Yuta Kikuchi. "Hayden-Preskill decoding from noisy Hawking radiation". In: *Journal of High Energy Physics* 2021.2 (2021), pp. 1–22.

[38] Ning Bao et al. "Traversable wormholes as quantum channels: exploring CFT entanglement structure and channel capacity in holography". In: 2018.11 (2018), p. 71.

[39] Oliver Bär and Isabel Campos. "Global anomalies in chiral lattice gauge theory". In: *Nuclear Physics B-Proceedings Supplements* 83 (2000), pp. 594–599.

[40] R. Barends et al. "Superconducting quantum circuits at the surface code threshold for fault tolerance". In: *Nature* 508.7497 (2014), 500âĂŞ503.

[41] M Barkeshli, NY Yao, and CR Laumann. "Continuous preparation of a fractional Chern insulator". In: *Physical Review Letters* 115.2 (2015), p. 026802.

[42] Jacob Baron et al. "Order of magnitude smaller limit on the electric dipole moment of the electron". In: *Science* 343.6168 (2014), pp. 269–272.

[43] Daniel Barredo et al. "Synthetic three-dimensional atomic structures assembled atom by atom". In: *Nature* 561.7721 (2018), pp. 79–82.

[44] M. D. Barrett et al. "Sympathetic cooling of $^9Be^+$ and $^{24}Mg^+$ for quantum logic". In: *Physical Review A* 68 (4 2003), p. 042302. DOI: `10.1103/PhysRevA.68.042302`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.68.042302`.

[45] MD Barrett et al. "Deterministic quantum teleportation of atomic qubits". In: *Nature* 429.6993 (2004), pp. 737–739.

[46] Jean Baum et al. "Multiple-quantum dynamics in solid state NMR". In: *The Journal of Chemical Physics* 83.5 (1985), pp. 2015–2025.

[47] Helle Bechmann-Pasquinucci and Asher Peres. "Quantum cryptography with 3-state systems". In: *Physical Review Letters* 85.15 (2000), p. 3313.

[48] Charles H Bennett et al. "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels". In: *Physical Review Letters* 70.13 (1993), p. 1895.

[49] Gregory Bentsen, Yingfei Gu, and Andrew Lucas. "Fast scrambling on sparse graphs". In: *Proceedings of the National Academy of Sciences* 116.14 (2019), pp. 6689–6694.

[50] Emil J Bergholtz and Zhao Liu. "Topological flat band models and fractional Chern insulators". In: *International Journal of Modern Physics B* 27.24 (2013), p. 1330017.

[51] B Andrei Bernevig. "Topological insulators and topological superconductors". In: *Topological Insulators and Topological Superconductors*. Princeton university press, 2013.

[52] Hannes Bernien et al. "Probing many-body dynamics on a 51-atom quantum simulator". In: *Nature* 551.7682 (2017), pp. 579–584.

[53] M.V. Berry. "Quantal phase factors accompanying adiabatic changes". In: *Proc. R. Soc. Lond. A* 392 (1984), pp. 45–57.

[54] M.V. Berry. "The Adiabatic Phase and Pancharatnam's Phase for Polarized Light". In: *Journal of Modern Optics* 34.11 (1987), pp. 1401–1407.

[55] Aranya Bhattacharya et al. "Operator growth and Krylov construction in dissipative open quantum systems". In: *arXiv:2207.05347* (2022).

[56] R. Bianchetti et al. "Control and Tomography of a Three Level Superconducting Artificial Atom". In: *Physical Review Letters* 105 (22 2010), p. 223601. DOI: 10.1103/PhysRevLett.105.223601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.223601.

[57] RMW van Bijnen and T Pohl. "Quantum magnetism and topological ordering via enhanced Rydberg-dressing near Forster-resonances". In: *arXiv:1411.3118* (2014).

[58] John W Blanchard, Dmitry Budker, and Andreas Trabesinger. "Lower than low: Perspectives on zero-to ultralow-field nuclear magnetic resonance". In: *Journal of Magnetic Resonance* 323 (2021), p. 106886.

[59] R. Blatt and D. Wineland. "Entangled states of trapped atomic ions". In: *Nature* 453 (2008), pp. 1008–1014. URL: https://www.nature.com/articles/nature07125.

[60] Immanuel Bloch. "Ultracold quantum gases in optical lattices". In: *Nature physics* 1.1 (2005), pp. 23–30.

[61] Machiel S Blok et al. "Quantum information scrambling on a superconducting qutrit processor". In: *Physical Review X* 11.2 (2021), p. 021010.

[62] Robin Blume-Kohout and Peter S Turner. "The curious nonexistence of gaussian 2-designs". In: *Commun. Math. Phys.* 326.3 (2014), pp. 755–771.

[63] Justin G Bohnet et al. "Quantum spin dynamics and entanglement generation with hundreds of trapped ions". In: *Science* 352.6291 (2016), pp. 1297–1301.

[64] Keith D Bonin and Vitaly V Kresin. *Electric-Dipole Polarizabilities of Atoms, Molecules, and Clusters.* Singapore: World Scientific, 1997.

[65] Fausto Borgonovi and Felix M Izrailev. "Emergence of correlations in the process of thermalization of interacting bosons". In: *arXiv:1806.00435* ().

[66] Jens M Boss et al. "One-and two-dimensional nuclear magnetic resonance spectroscopy with a diamond quantum sensor". In: *Physical Review Letters* 116.19 (2016), p. 197601.

[67] R. Bott and R. Seeley. "Some remarks on the paper of Callias: "Axial anomalies and index theorems on open spaces"". In: *Comm. Math. Phys.* 62.3 (1978), pp. 235–245.

[68] Fernando GSL Brandao, Aram W Harrow, and Michał Horodecki. "Local random quantum circuits are approximate polynomial-designs". In: *Commun. Math. Phys.* 346.2 (2016), pp. 397–434.

[69] Sergey B Bravyi and A Yu Kitaev. "Quantum codes on a lattice with boundary". In: *quant-ph/9811052* (1998).

[70] Adam R Brown and Leonard Susskind. "Second law of quantum complexity". In: *Physical Review D* 97.8 (2018), p. 086015.

[71] Adam R. Brown et al. *Quantum Gravity in the Lab: Teleportation by Size and Traversable Wormholes.* 2019. arXiv: `1911.06314 [quant-ph]`.

[72] John M Brown and Alan Carrington. *Rotational spectroscopy of diatomic molecules.* Cambridge University Press, 2003.

[73] Winton Brown and Omar Fawzi. "Decoupling with Random Quantum Circuits". In: *Comm. Math. Phys.* 340.3 (2015), pp. 867–900. DOI: `10.1007/s00220-015-2470-1`.

[74] Todd Brun, Igor Devetak, and Min-Hsiu Hsieh. "Correcting quantum errors with entanglement". In: *Science* 314.5798 (2006), pp. 436–439.

[75] Dagmar Bruss and Chiara Macchiavello. "Optimal eavesdropping in cryptography with three-dimensional quantum states". In: *Physical Review Letters* 88.12 (2002), p. 127901.

[76] Colin D Bruzewicz et al. "Trapped-ion quantum computing: Progress and challenges". In: *Applied Physics Reviews* 6.2 (2019), p. 021314.

[77] Tiff Brydges et al. "Probing Rényi entanglement entropy via randomized measurements". In: *Science* 364.6437 (2019), pp. 260–263.

[78] Lars Buitinck et al. "API design for machine learning software: experiences from the scikit-learn project". In: *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 2013, pp. 108–122.

[79] Marin Bukov, Luca D'Alessio, and Anatoli Polkovnikov. "Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to Floquet engineering". In: *Advances in Physics* 64.2 (2015), pp. 139–226.

[80] Stephen S. Bullock, Dianne P. OâĂŹLeary, and Gavin K. Brennen. "Asymptotically Optimal Quantum Circuits for $d$-Level Systems". In: *Physical Review Letters* 94.23 (June 2005). Publisher: American Physical Society, p. 230502.

[81] Daniel Burgarth and Ashok Ajoy. "Evolution-free Hamiltonian parameter estimation through Zeeman markers". In: *Physical Review Letters* 119.3 (2017), p. 030402.

[82] Daniel Burgarth and Koji Maruyama. "Indirect Hamiltonian identification through a small gateway". In: *New Journal of Physics* 11.10 (2009), p. 103019.

[83] Daniel Burgarth, Koji Maruyama, and Franco Nori. "Coupling strength estimation for spin chains despite restricted access". In: *Physical Review A* 79.2 (2009), p. 020305.

[84] Kevin E Cahill and Roy J Glauber. "Ordered expansions in boson amplitude operators". In: *Physical Review* 177.5 (1969), p. 1857.

[85] Jianming Cai et al. "A large-scale quantum simulator on a diamond surface at room temperature". In: *Nature Physics* 9.3 (2013), pp. 168–173.

[86] F. E. Camino, Wei Zhou, and V. J. Goldman. "Realization of a Laughlin quasiparticle interferometer: Observation of fractional statistics". In: *Physical Review B* 72 (7 2005), p. 075342.

[87] Earl T. Campbell. "Enhanced Fault-Tolerant Quantum Computing in d-Level Systems". In: *Physical Review Letters* 113.23 (2014). ISSN: 1079-7114. DOI: 10.1103/physrevlett.113.230501. URL: http://dx.doi.org/10.1103/PhysRevLett.113.230501.

[88] Earl T Campbell, Hussain Anwar, and Dan E Browne. "Magic-state distillation in all prime dimensions using quantum Reed-Muller codes". In: *Physical Review X* 2.4 (2012), p. 041021.

[89] Paweł Caputa, Tokiro Numasawa, and Alvaro Veliz-Osorio. "Out-of-time-ordered correlators and purity in rational conformal field theories". In: *Progr. Theor. Exp. Phys* 11 (2016).

[90] David Carpentier et al. "Topological index for periodically driven time-reversal invariant 2D systems". In: *Physical Review Letters* 114.10 (2015), p. 106806.

[91] M. Cetina et al. "Quantum Gates on Individually-Addressed Atomic Qubits Subject to Noisy Transverse Motion". In: *arXiv:2007.06768* (2020).

[92]   Shira Chapman et al. "Toward a Definition of Complexity for Quantum Field The-
       ory States". In: *Physical Review Letters* 120 (12 2018), p. 121602. DOI: `10.1103/`
       `PhysRevLett.120.121602`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.`
       `120.121602`.

[93]   Shirshendu Chatterjee and Partha S. Dey. "Multiple phase transitions in long-range
       first-passage percolation on square lattices". In: *Communications on Pure and Applied
       Mathematics* 69.2 (2016), pp. 203–256.

[94]   Jorge Chávez-Carlos et al. "Quantum and Classical Lyapunov Exponents in Atom-
       Field Interaction Systems". In: *arXiv:1807.10292* (2018).

[95]   Sitan Chen et al. "A hierarchy for replica quantum advantage". In: *arXiv:2111.05874*
       (2021).

[96]   Sitan Chen et al. "Exponential separations between learning with and without quan-
       tum memory". In: *arXiv:2111.05881* (2021).

[97]   Xiao Chen and Tianci Zhou. "Quantum chaos dynamics in long-range power law
       interaction systems". In: *Physical Review B* 100.6 (2019), p. 064305.

[98]   Ching-Kai Chiu et al. "Classification of topological quantum matter with symmetries".
       In: *Reviews of Modern Physics* 88 (2016), p. 035005.

[99]   Soonwon Choi, Norman Y Yao, and Mikhail D Lukin. "Dynamical engineering of
       interactions in qudit ensembles". In: *Physical Review Letters* 119.18 (2017), p. 183603.

[100]  Soonwon Choi et al. "Observation of discrete time-crystalline order in a disordered
       dipolar many-body system". In: *Nature* 543.7644 (2017), p. 221. DOI: `10.1038/`
       `nature21426`.

[101]  T. Choi et al. "Optimal Quantum Control of Multimode Couplings between Trapped
       Ion Qubits for Scalable Entanglement". In: *Physical Review Letters* 112 (19 2014),
       p. 190502. DOI: `10.1103/PhysRevLett.112.190502`. URL: `http://link.aps.org/`
       `doi/10.1103/PhysRevLett.112.190502`.

[102]  Jerry M. Chow et al. "Simple All-Microwave Entangling Gate for Fixed-Frequency
       Superconducting Qubits". In: *Physical Review Letters* 107.8 (2011).

[103]  Alessio Ciamei et al. "The RbSr $2\,\Sigma+$ ground state investigated via spectroscopy of
       hot and ultracold molecules". In: *Physical Chemistry Chemical Physics* 20.41 (2018),
       pp. 26221–26240.

[104]  J. I. Cirac and P. Zoller. "Quantum Computation with Cold Trapped Ions". In: *Phys-
       ical Review Letters* 74 (1995), pp. 4091–4094.

[105]  J Ignacio Cirac and Peter Zoller. "Goals and opportunities in quantum simulation".
       In: *Nature Physics* 8.4 (2012), p. 264.

[106]  William R Clements et al. "Optimal design for universal multiport interferometers".
       In: *Optica* 3.12 (2016), pp. 1460–1465.

[107]   Richard Cleve, Daniel Gottesman, and Hoi-Kwong Lo. "How to share a quantum secret". In: *Physical Review Letters* 83.3 (1999), p. 648.

[108]   Benoît Collins and Sho Matsumoto. "Weingarten calculus via orthogonality relations: new applications". In: *arXiv:1701.04493* (2017).

[109]   Ivan Corwin. "The Kardar–Parisi–Zhang equation and universality class". In: *Random matrices: Theory and applications* 1.01 (2012), p. 1130001.

[110]   Marius V. Costache et al. "Lateral metallic devices made by a multiangle shadow evaporation technique". In: *Journal of Vacuum Science & Technology B* 30.4 (2012), 04E105. DOI: `10.1116/1.4722982`. eprint: `https://doi.org/10.1116/1.4722982`. URL: `https://doi.org/10.1116/1.4722982`.

[111]   Jordan Cotler, Thomas Schuster, and Masoud Mohseni. "Information-theoretic Hardness of Out-of-time-order Correlators". In: *arXiv:2208.02256* (2022).

[112]   Jordan Cotler and Frank Wilczek. "Quantum overlapping tomography". In: *Physical Review Letters* 124.10 (2020), p. 100401.

[113]   Jordan Cotler et al. "Chaos, complexity, and random matrices". In: *Journal of High Energy Physics* 2017.11 (2017), pp. 1–60.

[114]   Jordan S Cotler, Dawei Ding, and Geoffrey R Penington. "Out-of-time-order operators and the butterfly effect". In: *Annals of Physics* 396 (2018), pp. 318–333.

[115]   Jordan S Cotler et al. "Black holes and random matrices". In: *J. High Energy Phys.* 05 (2017), p. 118.

[116]   Jacob P Covey. *Enhanced optical and electric manipulation of a quantum gas of KRb molecules.* Springer, 2018.

[117]   Jacob P Covey et al. "An approach to spin-resolved molecular gas microscopy". In: *New Journal of Physics* 20.4 (2018), p. 043031.

[118]   Fernando M Cucchietti et al. "Decoherence and the Loschmidt echo". In: *Physical Review Letters* 91.21 (2003), p. 210403.

[119]   Kunzhe Dai et al. "Demonstration of quantum permutation parity determine algorithm in a superconducting qutrit". en. In: *Chinese Physics B* 27.6 (June 2018). Publisher: IOP Publishing, p. 060305.

[120]   Luca D'Alessio et al. "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics". In: *Advances in Physics* 65.3 (2016), pp. 239–362.

[121]   AJ Daley et al. "Measuring entanglement growth in quench dynamics of bosons in an optical lattice". In: *Physical Review Letters* 109.2 (2012), p. 020505.

[122]   Christoph Dankert et al. "Exact and approximate unitary 2-designs and their application to fidelity estimation". In: *Physical Review A* 80.1 (2009), p. 012304.

[123] Giacomo Mauro D'Ariano, Lorenzo Maccone, and Massimiliano Federico Sacchi. "Homodyne tomography and the reconstruction of quantum states of light". In: *Quantum Information With Continuous Variables of Atoms and Light*. World Scientific, 2007, pp. 141–158.

[124] K. M. Davis et al. "Writing waveguides in glass with a femtosecond laser". In: *Opt. Lett.* 21.21 (1996), pp. 1729–1731.

[125] Luigi De Marco et al. "A degenerate Fermi gas of polar molecules". In: *Science* 363.6429 (2019), pp. 853–856.

[126] S. Debnath et al. "Demonstration of a small programmable quantum computer with atomic qubits". In: *Nature* 536.7614 (Aug. 2016), pp. 63–66.

[127] Shantanu Debnath. "A Programmable Five Qubit Quantum Computer Using Trapped Atomic Ions". PhD thesis. University of Maryland, College Park, 2016.

[128] Christian L Degen, F Reinhard, and Paola Cappellaro. "Quantum sensing". In: *Reviews of modern physics* 89.3 (2017), p. 035002.

[129] J Deiglmayr et al. "Permanent dipole moment of LiCs in the ground state". In: *Physical Review A* 82.3 (2010), p. 032503.

[130] D-L Deng, S-T Wang, and L-M Duan. "Systematic construction of tight-binding Hamiltonians for topological insulators and superconductors". In: *Physical Review B* 89.7 (2014), p. 075126.

[131] D-L Deng et al. "Hopf insulators and their topologically protected surface states". In: *Physical Review B* 88.20 (2013), p. 201105.

[132] Dong-Ling Deng et al. "Probe knots and Hopf insulators with ultracold atoms". In: *Chinese Physics Letters* 35.1 (2018), p. 013701.

[133] Yao-Min Di and Hai-Rui Wei. *Elementary gates for ternary quantum logic circuit*. 2011. arXiv: 1105.5485 [quant-ph].

[134] Yao-Min Di and Hai-Rui Wei. "Synthesis of multivalued quantum logic circuits by elementary gates". In: *Physical Review A* 87.1 (Jan. 2013). Publisher: American Physical Society, p. 012325.

[135] Carlo Di Franco, Mauro Paternostro, and MS Kim. "Hamiltonian tomography in an access-limited setting without state initialization". In: *Physical Review Letters* 102.18 (2009), p. 187203.

[136] Shiqian Ding et al. "Cross-Kerr Nonlinearity for Phonon Counting". In: *Physical Review Letters* 119 (19 2017), p. 193602. DOI: 10.1103/PhysRevLett.119.193602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.193602.

[137] Florian Dolde et al. "High-fidelity spin entanglement using optimal control". In: *Nature Communications* 5 (Feb. 2014).

[138] Federico D Domínguez and Gonzalo A Álvarez. "Dynamics of quantum information scrambling under decoherence effects measured via active spin clusters". In: *Physical Review A* 104.6 (2021), p. 062406.

[139] Federico D Domínguez et al. "Decoherence scaling transition in the dynamics of quantum information scrambling". In: *Physical Review A* 104.1 (2021), p. 012402.

[140] A. Dunsworth et al. "Characterization and reduction of capacitive loss induced by sub-micron Josephson junction fabrication in superconducting qubits". In: *Applied Physics Letters* 111.2 (2017), p. 022601. ISSN: 1077-3118. DOI: 10.1063/1.4993577. URL: http://dx.doi.org/10.1063/1.4993577.

[141] Sepehr Ebadi et al. "Quantum phases of matter on a 256-atom programmable quantum simulator". In: *Nature* 595.7866 (2021), pp. 227–232.

[142] A. Elben et al. "Rényi Entropies from Random Quenches in Atomic Hubbard and Spin Models". In: *Physical Review Letters* 120 (5 2018), p. 050406. DOI: 10.1103/PhysRevLett.120.050406. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.050406.

[143] Shmuel Elitzur and VP Nair. "Nonperturbative anomalies in higher dimensions". In: *Nucl. Phys.* 243.Print-84-0257 (IAS, PRINCETON) (1984), p. 205.

[144] Dominic V. Else et al. "Improved Lieb-Robinson bound for many-body Hamiltonians with power-law interactions". In: *Physical Review A* 101 (2 2020), p. 022333. DOI: 10.1103/PhysRevA.101.022333. URL: https://link.aps.org/doi/10.1103/PhysRevA.101.022333.

[145] Nic Ezzell et al. "Dynamical decoupling for superconducting qubits: a performance survey". In: *arXiv:2207.03670* (2022).

[146] F. F. Fanchini, J. E. M. Hornos, and R. d. J. Napolitano. "Continuously decoupling single-qubit operations from a perturbing thermal bath of scalar bosons". In: *Physical Review A* 75.2 (Feb. 2007). Publisher: American Physical Society, p. 022329.

[147] A. Fedorov et al. "Implementation of a Toffoli gate with superconducting circuits". In: *Nature* 481 (Jan. 2012).

[148] MD Feit and JA Fleck Jr. "Wave packet dynamics and chaos in the Hénon–Heiles system". In: *J. Chem. Phys.* 80.6 (1984), pp. 2578–2584.

[149] Igor Ferrier-Barbut et al. "Scissors mode of dipolar quantum droplets of dysprosium atoms". In: *Physical Review Letters* 120.16 (2018), p. 160402.

[150] Austin G Fowler et al. "Surface codes: Towards practical large-scale quantum computation". In: *Physical Review A* 86.3 (2012), p. 032324.

[151] M. Freedman et al. "Topological Quantum Computation". In: *Bulletin of the American Mathematical Society* 40 (1 2002), pp. 31–38.

[152] Jürg Fröhlich. "Statistics of Fields, the Yang-Baxter Equation, and the Theory of Knots and Links". English. In: *Nonperturbative Quantum Field Theory*. Ed. by G. 't Hooft et al. Vol. 185. NATO ASI Series. Springer US, 1988, pp. 71–100.

[153] Michel Fruchart. "Complex classes of periodically driven topological lattice systems". In: *Physical Review B* 93.11 (2016), p. 115429.

[154] Liang Fu. "Topological crystalline insulators". In: *Physical Review Letters* 106.10 (2011), p. 106802.

[155] Liang Fu and C. L. Kane. "Superconducting Proximity Effect and Majorana Fermions at the Surface of a Topological Insulator". In: *Physical Review Letters* 100 (9 2008), p. 096407.

[156] Liang Fu, Charles L Kane, and Eugene J Mele. "Topological insulators in three dimensions". In: *Physical Review Letters* 98.10 (2007), p. 106803.

[157] Takahiro Fukui, Takanori Fujiwara, and Yasuhiro Hatsugai. "Topological meaning of Z2 numbers in time reversal invariant systems". In: *Journal of the Physical Society of Japan* 77.12 (2008), pp. 1–4. ISSN: 00319015. DOI: `10.1143/JPSJ.77.123705`. eprint: `0809.4532`.

[158] J. P. Gaebler et al. "High-Fidelity Universal Gate Set for $^9Be^+$ Ion Qubits". In: *Physical Review Letters* 117 (6 2016), p. 060505. DOI: `10.1103/PhysRevLett.117.060505`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.117.060505`.

[159] Ping Gao and Daniel Louis Jafferis. "A Traversable Wormhole Teleportation Protocol in the SYK Model". In: *arXiv:1911.07416* (2019).

[160] Ping Gao, Daniel Louis Jafferis, and Aron C Wall. "Traversable wormholes via a double trace deformation". In: *Journal of High Energy Physics* 2017.12 (2017), p. 151.

[161] Ping Gao and Hong Liu. "Regenesis and quantum traversable wormholes". In: *arXiv:1810.01444* (2018).

[162] Yvonne Y. Gao et al. "Entanglement of bosonic modes through an engineered exchange interaction". en. In: *Nature* 566.7745 (Feb. 2019), pp. 509–512.

[163] Martin Gärttner et al. "Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet". In: *Nature Physics* 13.8 (2017), p. 781.

[164] Alexander L Gaunt et al. "Bose-Einstein condensation of atoms in a uniform potential". In: *Physical Review Letters* 110.20 (2013), p. 200406.

[165] Marco G Genoni and Matteo GA Paris. "Quantifying non-Gaussianity for quantum information". In: *Physical Review A* 82.5 (2010), p. 052341.

[166] Marco G Genoni, Matteo GA Paris, and Konrad Banaszek. "Quantifying the non-Gaussian character of a quantum state by quantum relative entropy". In: *Physical Review A* 78.6 (2008), p. 060303.

[167] Antonio A Gentile et al. "Learning models of quantum systems from experiments". In: *Nature Physics* 17.7 (2021), pp. 837–843.

[168] A. W. Glaetzle et al. "Designing Frustrated Quantum Magnets with Laser-Dressed Rydberg Atoms". In: *Physical Review Letters* 114 (2015), p. 173002.

[169] Pranav Gokhale et al. "Asymptotic Improvements to Quantum Circuits via Qutrits". In: *Proceedings of the 46th International Symposium on Computer Architecture - ISCA '19* (2019). arXiv: 1905.10481, pp. 554–566. DOI: `10.1145/3307650.3322253`. URL: `http://arxiv.org/abs/1905.10481`.

[170] Alexey V Gorshkov, Kaden RA Hazzard, and Ana Maria Rey. "Kitaev honeycomb and other exotic spin models with polar molecules". In: *Molecular Physics* 111.12-13 (2013), pp. 1908–1916.

[171] Daniel Gottesman. "The Heisenberg representation of quantum computers". In: *quant-ph/9807006* (1998).

[172] Daniel Gottesman, Alexei Kitaev, and John Preskill. "Encoding a qubit in an oscillator". In: *Physical Review A* 64.1 (2001), p. 012310.

[173] Arseni Goussev et al. "Loschmidt echo". In: *arXiv:1206.6348* (2012).

[174] T. M. Graham et al. "Rydberg mediated entanglement in a two-dimensional neutral atom qubit array". In: *Physical Review Letters* 123 (2019), p. 230501.

[175] D. Gross. "Hudson's theorem for finite-dimensional quantum systems". In: *J. Math. Phys.* 47.12 (2006), p. 122107. DOI: `10.1063/1.2393152`. URL: `https://doi.org/10.1063/1.2393152`.

[176] David Gross, Koenraad Audenaert, and Jens Eisert. "Evenly distributed unitaries: On the structure of unitary designs". In: *J. Math. Phys.* 48.5 (2007), p. 052104.

[177] David Gross, Sepehr Nezami, and Michael Walter. "Schur-Weyl Duality for the Clifford Group with Applications: Property Testing, a Robust Hudson Theorem, and de Finetti Representations". In: *arXiv:1712.08628* ().

[178] Frédéric Grosshans and Philippe Grangier. "Continuous variable quantum cryptography using coherent states". In: *Physical Review Letters* 88.5 (2002), p. 057902.

[179] Lov K. Grover. "Quantum Mechanics Helps in Searching for a Needle in a Haystack". In: *Physical Review Letters* 79 (2 1997), pp. 325–328. DOI: `10.1103/PhysRevLett.79.325`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.79.325`.

[180] Adolfo G Grushin, Álvaro Gómez-León, and Titus Neupert. "Floquet fractional Chern insulators". In: *Physical Review Letters* 112.15 (2014), p. 156801.

[181] Yingfei Gu and Alexei Kitaev. "On the relation between the magnitude and exponent of OTOCs". In: 2019.2 (2019), p. 75.

[182] Yingfei Gu, Xiao-Liang Qi, and Douglas Stanford. "Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models". In: *Journal of High Energy Physics* 2017.5 (2017), p. 125.

[183] Zhenghao Gu et al. "Floquet Spectrum and Transport through an Irradiated Graphene Ribbon". In: *Physical Review Letters* 107, 216601 (2011), p. 216601.

[184] Jonathan Guglielmon et al. "Photonic realization of a transition to a strongly driven Floquet topological phase". In: *Physical Review A* 97 (3 2018), p. 031801.

[185] Mingyang Guo et al. "Creation of an ultracold gas of ground-state dipolar na 23 rb 87 molecules". In: *Physical Review Letters* 116.20 (2016), p. 205303.

[186] Jeongwan Haah, Robin Kothari, and Ewin Tang. "Optimal learning of quantum Hamiltonians from high-temperature Gibbs states". In: *arXiv:2108.04842* (2021).

[187] Felix M Haehl and Moshe Rozali. "Fine Grained Chaos in AdS 2 Gravity". In: *Physical Review Letters* 120.12 (2018), p. 121601.

[188] Felix M Haehl, Alexandre Streicher, and Ying Zhao. "Six-point functions and collisions in the black hole interior". In: *arXiv:2105.12755* (2021).

[189] Mohammad Hafezi et al. "Imaging topological edge states in silicon photonics". In: *Nat. Photon.* 7 (2013), pp. 907–912.

[190] F. D. M. Haldane and S. Raghu. "Possible Realization of Directional Optical Waveguides in Photonic Crystals with Broken Time-Reversal Symmetry". In: *Physical Review Letters* 100 (1 2008), p. 013904.

[191] F Duncan M Haldane. "Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the" parity anomaly"". In: *Physical Review Letters* 61.18 (1988), p. 2015.

[192] Oskar Hallatschek and Daniel S Fisher. "Acceleration of evolutionary spread by long-range dispersal". In: *Proceedings of the National Academy of Sciences* 111.46 (2014), E4911–E4919.

[193] B. I. Halperin. "Statistics of Quasiparticles and the Hierarchy of Fractional Quantized Hall States". In: *Physical Review Letters* 52 (18 1984), pp. 1583–1586.

[194] Nicole Yunger Halpern. "Jarzynski-like equality for the out-of-time-ordered correlator". In: *Physical Review A* 95.1 (2017), p. 012120.

[195] RH Hardin. "Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations". In: *SIAM Review (Chronicles)* 15 (1973), p. 423.

[196] Daniel Harlow. "Jerusalem lectures on black holes and quantum information". In: *Reviews of Modern Physics* 88.1 (2016), p. 015002.

[197] Joseph Harris, Bin Yan, and Nikolai A Sinitsyn. "Benchmarking Information Scrambling". In: *Physical Review Letters* 129.5 (2022), p. 050602.

[198] Aram W Harrow and Richard A Low. "Random quantum circuits are approximate 2-designs". In: *Commun. Math. Phys.* 291.1 (2009), pp. 257–302.

[199] Thomas Hartman and Juan Maldacena. "Time evolution of entanglement entropy from black hole interiors". In: *J. High Energy Phys.* 05 (2013), p. 014. DOI: `10.1007/JHEP05(2013)014`.

[200] Koji Hashimoto, Keiju Murata, and Ryosuke Yoshii. "Out-of-time-order correlators in quantum mechanics". In: *J. High Energy Phys.* 10 (2017), p. 138.

[201] S. W. Hawking. "Breakdown of predictability in gravitational collapse". In: *Physical Review D* 14 (10 1976), pp. 2460–2473.

[202] Patrick Hayden and John Preskill. "Black holes as mirrors: quantum information in random subsystems". In: *Journal of High Energy Physics* 2007.09 (2007), p. 120.

[203] Patrick Hayden et al. "Holographic duality from random tensor networks". In: *J. High Energy Phys.* 2016.11 (2016), p. 9.

[204] Kaden RA Hazzard et al. "Far-from-equilibrium quantum magnetism with ultracold polar molecules". In: *Physical Review Letters* 110.7 (2013), p. 075301.

[205] Kaden RA Hazzard et al. "Many-body dynamics of dipolar molecules in an optical lattice". In: *Physical Review Letters* 113.19 (2014), p. 195302.

[206] Yan He and Chih-Chun Chien. "Non-Hermitian three-dimensional two-band Hopf insulator". In: *arXiv:2003.06715* (2020).

[207] Yan He and Chih-Chun Chien. "Three-dimensional two-band Floquet topological insulator with Z 2 index". In: *Physical Review B* 99.7 (2019), p. 075120.

[208] Reinier W Heeres et al. "Cavity state manipulation using photon-number selective phase gates". In: *Physical Review Letters* 115.13 (2015), p. 137002.

[209] WK Hensinger et al. "T-junction ion trap array for two-dimensional ion shuttling, storage, and manipulation". In: *Applied Physics Letters* 88.3 (2006), p. 034101.

[210] Martin Hermann and M Seravi. *Nonlinear ordinary differential equations*. Springer, 2016.

[211] "Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence". In: *Journal of High Energy Physics* 2015.6 (2015), p. 149.

[212] Heinz Hopf. "About the images of the three-dimensional Sphere on the spherical surface". In: *Mathematical Annals* 104.1 (1931), pp. 637–665.

[213] Pavan Hosur et al. "Chaos in quantum channels". In: *Journal of High Energy Physics* 2016.2 (2016), p. 4.

[214] Chang-Yu Hou, Claudio Chamon, and Christopher Mudry. "Electron Fractionalization in Two-Dimensional Graphenelike Structures". In: *Physical Review Letters* 98 (18 2007), p. 186809.

[215] Haiping Hu, Chao Yang, and Erhai Zhao. "Quench dynamics of Hopf insulators". In: *Physical Review B* 101.15 (2020), p. 155131.

[216] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Information-theoretic bounds on quantum advantage in machine learning". In: *Physical Review Letters* 126.19 (2021), p. 190505.

[217] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Predicting many properties of a quantum system from very few measurements". In: *Nature Physics* 16.10 (2020), pp. 1050–1057.

[218] Hsin-Yuan Huang et al. "Provably efficient machine learning for quantum many-body problems". In: *Science* 377.6613 (2022), eabk3333.

[219] Hsin-Yuan Huang et al. "Quantum advantage in learning from experiments". In: *arXiv:2112.00778* (2021).

[220] Hsin-Yuan Huang et al. "Quantum advantage in learning from experiments". In: *Science* 376.6598 (2022), pp. 1182–1186. DOI: `10.1126/science.abn7293`. eprint: `https://www.science.org/doi/pdf/10.1126/science.abn7293`.

[221] Eric R Hudson et al. "Cold molecule spectroscopy for constraining the evolution of the fine structure constant". In: *Physical review letters* 96.14 (2006), p. 143004.

[222] Thomas Iadecola, Thomas Schuster, and Claudio Chamon. "Non-Abelian Braiding of Light". In: *Physical Review Letters* 117 (7 2016), p. 073901.

[223] Poolad Imany et al. In: *npj Quantum Information* 5.1 (2019), pp. 1–10. DOI: `10.1038/s41534-019-0173-8`.

[224] Rajibul Islam et al. "Measuring entanglement entropy in a quantum many-body system". In: *Nature* 528.7580 (2015), pp. 77–83.

[225] D. A. Ivanov. "Non-Abelian Statistics of Half-Quantum Vortices in *p*-Wave Superconductors". In: *Physical Review Letters* 86 (2 2001), pp. 268–271.

[226] M. Izutsu, Y. Nakai, and T. Sueta. "Operation mechanism of the single-mode optical-waveguide Y junction". In: *Opt. Lett.* 7.3 (1982), pp. 136–138.

[227] R. Jackiw and P. Rossi. "Zero modes of the vortex-fermion system". In: *Nucl. Phys. B* 190.4 (1981), pp. 681 –691.

[228] D. Jaksch et al. "Fast Quantum Gates for Neutral Atoms". In: *Physical Review Letters* 85 (2000), p. 2208.

[229] Rodolfo A Jalabert and Horacio M Pastawski. "Environment-independent decoherence rate in classically chaotic systems". In: *Physical Review Letters* 86.12 (2001), p. 2490.

[230] V John et al. "Techniques for the reconstruction of a distribution from a finite number of its moments". In: *Chemical Engineering Science* 62.11 (2007), pp. 2890–2904.

[231] Sonika Johri, Damian S. Steiger, and Matthias Troyer. "Entanglement spectroscopy on a quantum computer". In: *Physical Review B* 96 (19 2017), p. 195136. DOI: `10.1103/PhysRevB.96.195136`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.96.195136`.

[232] Chaitanya Joshi, Elinor K Irish, and Timothy P Spiller. "Qubit-flip-induced cavity mode squeezing in the strong dispersive regime of the quantum Rabi model". In: *Sci. Rep.* 7 (2017), p. 45587.

[233] Manoj K Joshi et al. "Quantum information scrambling in a trapped-ion quantum simulator with tunable range interactions". In: *arXiv:2001.02176* (2020).

[234] Petar Jurcevic et al. *Demonstration of quantum volume 64 on a superconducting quantum computing system.* 2020. arXiv: 2008.08571 [quant-ph].

[235] Abhinav Kandala et al. "Error mitigation extends the computational reach of a noisy quantum processor". In: *Nature* 567.7749 (2019), 491âĂŞ495.

[236] Abhinav Kandala et al. "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets". In: *Nature* 549.7671 (2017), 242âĂŞ246. ISSN: 1476-4687. DOI: 10.1038/nature23879. URL: http://dx.doi.org/10.1038/nature23879.

[237] Charles L Kane and Eugene J Mele. "Z 2 topological order and the quantum spin Hall effect". In: *Physical Review Letters* 95.14 (2005), p. 146802.

[238] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. "Dynamic scaling of growing interfaces". In: *Physical Review Letters* 56.9 (1986), p. 889.

[239] Torsten Karzig et al. "Scalable designs for quasiparticle-poisoning-protected topological quantum computation with Majorana zero modes". In: *Physical Review B* 95 (23 2017), p. 235305.

[240] Adam M Kaufman et al. "Quantum thermalization through entanglement in an isolated many-body system". In: *Science* 353.6301 (2016), pp. 794–800.

[241] H. Kaufmann et al. "Fast ion swapping for quantum-information processing". In: *Physical Review A* 95 (5 2017), p. 052319. DOI: 10.1103/PhysRevA.95.052319. URL: https://link.aps.org/doi/10.1103/PhysRevA.95.052319.

[242] Steven Kay. "Fundamentals of statistical signal processing: estimation theory". In: *Technometrics* 37.4 (1993), pp. 465–466.

[243] Ricardo Kennedy. "Topological Hopf-Chern insulators and the Hopf superconductor". In: *Physical Review B* 94.3 (2016), p. 035137.

[244] Ricardo Kennedy and Charles Guggenheim. "Homotopy theory of strong and weak topological insulators". In: *Physical Review B* 91.24 (2015), p. 245148.

[245] Curt W von Keyserlingk and Shivaji L Sondhi. "Phase structure of one-dimensional interacting Floquet systems. I. Abelian symmetry-protected topological phases". In: *Physical Review B* 93.24 (2016), p. 245145.

[246] Vedika Khemani, Ashvin Vishwanath, and David A Huse. "Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws". In: *Physical Review X* 8.3 (2018), p. 031057.

[247] Vedika Khemani et al. "Phase structure of driven quantum systems". In: *Physical Review Letters* 116.25 (2016), p. 250401.

[248] A Yu Kitaev. "Fault-tolerant quantum computation by anyons". In: *Annals of Physics* 303.1 (2003), pp. 2–30.

[249] Alexei Kitaev. *A simple model of quantum holography.* 2015.

[250] Alexei Kitaev. "Hidden correlations in the Hawking radiation and thermal noise". In: *Talk given at the Fundamental Physics Prize Symposium.* Vol. 10. 2014.

[251] Alexei Kitaev. "Periodic table for topological insulators and superconductors". In: *AIP conference proceedings.* Vol. 1134. 1. American Institute of Physics. 2009, pp. 22–30.

[252] Alexei Kitaev and S Josephine Suh. "The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual". In: 2018.5 (2018), p. 183.

[253] A.Yu. Kitaev. "Fault-tolerant quantum computation by anyons". In: *Ann. Phys. (N.Y.)* 303.1 (2003), pp. 2 –30.

[254] Takuya Kitagawa et al. "Topological characterization of periodically driven quantum systems". In: *Physical Review B* 82.23 (2010), p. 235114.

[255] J. Kobayashi et al. "Prospects for narrow-line cooling of KRb molecules in the rovibrational ground state". In: *Physical Review A* 89 (2 2014), p. 021401.

[256] Bryce Kobrin et al. "Many-Body Chaos in the Sachdev-Ye-Kitaev Model". In: *arXiv:2002.05725* (2020).

[257] Bryce Kobrin et al. "Many-body chaos in the sachdev-ye-kitaev model". In: *Physical Review Letters* 126.3 (2021), p. 030602.

[258] Jens Koch et al. "Charge-insensitive qubit design derived from the Cooper pair box". In: *Physical Review A* 76.4 (2007). ISSN: 1094-1622. DOI: `10.1103/physreva.76.042319`. URL: `http://dx.doi.org/10.1103/PhysRevA.76.042319`.

[259] Michael H. Kolodrubetz et al. "Topological Floquet-Thouless Energy Pump". In: *Physical Review Letters* 120 (15 2018), p. 150601. DOI: `10.1103/PhysRevLett.120.150601`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.120.150601`.

[260] Markus König et al. "Quantum spin Hall insulator state in HgTe quantum wells". In: *Science* 318.5851 (2007), pp. 766–770.

[261] Simcha Korenblit et al. "Quantum simulation of spin models on an arbitrary lattice with trapped ions". In: *New Journal of Physics* 14.9 (2012), p. 095024.

[262] Georg Köstenberger. "Weingarten Calculus". In: *arXiv:2101.00921* (2021).

[263] S. Kotochigova and E. Tiesinga. "Controlling polar molecules in optical lattices". In: *Physical Review A* 73 (4 2006), p. 041405.

[264] S. Kotochigova, E. Tiesinga, and P. S. Julienne. "Photoassociative formation of ultracold polar KRb molecules". In: *Eur. Phys. J. D* 31 (2004), p. 189.

[265]   Svetlana Kotochigova and David DeMille. "Electric-field-dependent dynamic polariz-
        ability and state-insensitive conditions for optical trapping of diatomic polar molecules".
        In: *Physical Review A* 82 (6 2010), p. 063421.

[266]   Ioanna Kourkoulou and Juan Maldacena. "Pure states in the SYK model and nearly-
        AdS _2 gravity". In: *arXiv:1707.02325* (2017).

[267]   Mikhail G Kozlov and Leonti N Labzowsky. "Parity violation effects in diatomics". In:
        *Journal of Physics B: Atomic, Molecular and Optical Physics* 28.10 (1995), p. 1933.

[268]   Stefan Krastanov et al. "Universal control of an oscillator with dispersive coupling to
        a qubit". In: *Physical Review A* 92.4 (2015), p. 040303.

[269]   J. M. Kreikebaum, K. P. O'Brien, and I. Siddiqi. *Improving wafer-scale Joseph-
        son junction resistance variation in superconducting quantum coherent circuits.* 2019.
        arXiv: `1909.09165` [`quant-ph`].

[270]   Georg Kucsko et al. "Critical thermalization of a disordered dipolar spin system in
        diamond". In: *Physical Review Letters* 121.2 (2018), p. 023601.

[271]   R. Kueng and D. Gross. "Qubit stabilizer states are complex projective 3-designs".
        In: *ArXiv* abs/1510.02767 (2015).

[272]   P. Kurpiers et al. "Deterministic quantum state transfer and remote entanglement
        using microwave photons". In: *Nature* 558 (June 2018).

[273]   H. Labuhn et al. "Tunable two-dimensional arrays of single Rydberg atoms for real-
        izing quantum Ising models". In: *Nature* 534.7609 (2016), pp. 667–70.

[274]   Kevin A Landsman et al. "Verified quantum information scrambling". In: *Nature*
        567.7746 (2019), pp. 61–65.

[275]   B. P. Lanyon et al. "Manipulating Biphotonic Qutrits". In: *Physical Review Letters*
        100 (Feb. 2008).

[276]   Radek Lapkiewicz et al. "Experimental non-classicality of an indivisible quantum
        system". In: *Nature* 474 (June 2011).

[277]   AI Larkin and Yu N Ovchinnikov. "Quasiclassical method in the theory of supercon-
        ductivity". In: *Sov Phys JETP* 28.6 (1969), pp. 1200–1205.

[278]   Nima Lashkari et al. "Towards the fast scrambling conjecture". In: *J. High Energy
        Phys.* 04 (2013), p. 22. DOI: `10.1007/JHEP04(2013)022`.

[279]   David D Laws, Hans-Marcus L Bitter, and Alexej Jerschow. "Solid-state NMR spec-
        troscopic methods in chemistry". In: *Angewandte Chemie International Edition* 41.17
        (2002), pp. 3096–3129.

[280]   Ching Hua Lee et al. "Floquet mechanism for non-Abelian fractional quantum hall
        states". In: *Physical Review Letters* 121.23 (2018), p. 237401.

[281]   Tony E Lee. "Floquet engineering from long-range to short-range interactions". In:
        *Physical Review A* 94.4 (2016), p. 040701.

[282] J.M. Leinaas and J. Myrheim. "On the theory of identical particles". English. In: *Il Nuovo Cimento B* 37.1 (1977), pp. 1–23.

[283] Gabriela B Lemos et al. "Experimental observation of quantum chaos in a beam of light". In: *Nat. Commun.* 3 (2012), p. 1211.

[284] S. Lepoutre et al. "Out-of-equilibrium quantum magnetism and thermalization in a spin-3 many-body dipolar lattice system". In: *Nature Communications* 10.1 (2019). ISSN: 2041-1723. DOI: `10.1038/s41467-019-09699-5`. URL: `http://dx.doi.org/10.1038/s41467-019-09699-5`.

[285] Sylvain de Léséleuc et al. "Observation of a symmetry-protected topological phase of interacting bosons with Rydberg atoms". In: *Science* 365.6455 (2019), pp. 775–780.

[286] Eyal Leviatan et al. "Quantum thermalization dynamics with matrix-product states". In: *arXiv:1702.08894* (2017).

[287] H. Levine et al. "High-fidelity control and entanglement of Rydberg atom qubits". In: *Physical Review Letters* 121 (2018), p. 123603.

[288] H. Levine et al. "Parallel implementation of high-fidelity multi-qubit gates with neutral atoms". In: *arXiv:1908.06101* (2019).

[289] Jun Li et al. "Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator". In: *Physical Review X* 7.3 (2017), p. 031011.

[290] Linhu Li, Ching Hua Lee, and Jiangbin Gong. "Realistic Floquet Semimetal with Exotic Topological Linkages between Arbitrarily Many Nodal Loops". In: *Physical Review Letters* 121 (3 2018), p. 036401.

[291] Yaodong Li, Xiao Chen, and Matthew PA Fisher. "Measurement-driven entanglement transition in hybrid quantum circuits". In: *Physical Review B* 100.13 (2019), p. 134306.

[292] Yaodong Li, Xiao Chen, and Matthew PA Fisher. "Quantum Zeno effect and the many-body entanglement transition". In: *Physical Review B* 98.20 (2018), p. 205136.

[293] Elliott H Lieb and Derek W Robinson. "The finite group velocity of quantum spin systems". In: *Statistical mechanics*. Springer, 1972, pp. 425–431.

[294] Henry W Lin, Juan Maldacena, and Ying Zhao. "Symmetries Near the Horizon". In: *arXiv:1904.12820* (2019).

[295] Henry W Lin and Leonard Susskind. "Complexity Geometry and Schwarzian Dynamics". In: *arXiv:1911.02603* (2019).

[296] Jianhua Lin. "Divergence measures based on the Shannon entropy". In: *IEEE Transactions on Information theory* 37.1 (1991), pp. 145–151.

[297] Netanel H. Lindner, Gil Refael, and Victor Galitski. "Floquet topological insulator in semiconductor quantum wells". In: *Nature Physics* 7 (2011), pp. 490–495.

[298] Norbert M Linke et al. "Measuring the Rényi entropy of a two-site Fermi-Hubbard model on a trapped ion quantum computer". In: *Physical Review A* 98.5 (2018), p. 052334.

[299] Chunxiao Liu, Farzan Vafa, and Cenke Xu. "Symmetry-protected topological Hopf insulator and its generalizations". In: *Physical Review B* 95.16 (2017), p. 161116.

[300] Jing Liu et al. "Quantum Fisher information matrix and multiparameter estimation". In: *Journal of Physics A: Mathematical and Theoretical* 53.2 (2019), p. 023001.

[301] LR Liu et al. "Building one molecule from a reservoir of two atoms". In: *Science* 360.6391 (2018), pp. 900–903.

[302] Seth Lloyd and Samuel L Braunstein. "Quantum computation over continuous variables". In: *Physical Review Letters* 82.8 (1999), p. 1784.

[303] J Loschmidt. "volume 73 of Sitzungsber". In: *Kais. Akad. Wiss. Wien, Math. Naturwiss* (1876), pp. 128–142.

[304] Igor Lovchinsky et al. "Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic". In: *Science* 351.6275 (2016), pp. 836–841.

[305] Mingwu Lu, Seo Ho Youn, and Benjamin L Lev. "Trapping ultracold dysprosium: a highly magnetic gas for dipolar physics". In: *Physical Review Letters* 104.6 (2010), p. 063001.

[306] Frane Lunić et al. "Exact solutions of a model for synthetic anyons in a noninteracting system". In: *Physical Review B* 101.11 (2020), p. 115139.

[307] MingXing Luo and XiaoJun Wang. "Universal quantum computation with qudits". In: *Science China Physics, Mechanics & Astronomy* 57.9 (2014), pp. 1712–1717.

[308] Yi-Han Luo et al. "Quantum Teleportation in High Dimensions". In: *Physical Review Letters* 123.7 (2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.070505. URL: http://dx.doi.org/10.1103/PhysRevLett.123.070505.

[309] Alexander I Lvovsky and Michael G Raymer. "Continuous-variable optical quantum-state tomography". In: *Reviews of Modern Physics* 81.1 (2009), p. 299.

[310] Lukas J Maczewsky et al. "Observation of photonic anomalous Floquet topological insulators". In: *Nature Communications* 8 (2017), p. 13756.

[311] I. S. Madjarov et al. "High-fidelity entanglement and detection of alkaline-earth Rydberg atoms". In: *Nature Physics* (2020).

[312] Easwar Magesan and Jay M. Gambetta. "Effective Hamiltonian models of the cross-resonance gate". In: *Physical Review A* 101.5 (2020).

[313] Juan Maldacena. "The large-N limit of superconformal field theories and supergravity". In: *International Journal of Theoretical Physics* 38.4 (1999), pp. 1113–1133.

[314] Juan Maldacena and Xiao-Liang Qi. "Eternal traversable wormhole". In: *arXiv:1804.00491* (2018).

[315] Juan Maldacena, Stephen H Shenker, and Douglas Stanford. "A bound on chaos". In: *Journal of High Energy Physics* 2016.8 (2016), p. 106.

[316] Juan Maldacena and Douglas Stanford. "Remarks on the sachdev-ye-kitaev model". In: *Physical Review D* 94.10 (2016), p. 106002.

[317] Juan Maldacena, Douglas Stanford, and Zhenbin Yang. "Diving into traversable wormholes". In: *Fortschritte der Physik* 65.5 (2017).

[318] Juan Maldacena and Leonard Susskind. "Cool horizons for entangled black holes". In: *Fortschritte der Physik* 61.9 (2013), pp. 781–811.

[319] K. M. Maller et al. "Rydberg-blockade controlled-NOT gate and entanglement in a two-dimensional array of neutral-atom qubits". In: *Physical Review A* 92 (2015), p. 022336.

[320] Paulina Marian and Tudor A Marian. "Relative entropy is an exact measure of non-Gaussianity". In: *Physical Review A* 88.1 (2013), p. 012322.

[321] G Edward Marti et al. "Imaging Optical Frequencies with 100 $\mu$ Hz Precision and 1.1 $\mu$ m Resolution". In: *Physical Review Letters* 120.10 (2018), p. 103201.

[322] John Martyn and Brian Swingle. "Product spectrum ansatz and the simplicity of thermal states". In: *Physical Review A* 100.3 (2019), p. 032107.

[323] Kyle Matsuda et al. "Resonant collisional shielding of reactive molecules using electric fields". In: *Science* 370.6522 (2020), pp. 1324–1327.

[324] David C. McKay et al. "Efficient Z gates for quantum computing". In: *Physical Review A* 96.2 (2017). ISSN: 2469-9934.

[325] Eric J Meier, Fangzhao Alex An, Bryce Gadway, et al. "Exploring quantum signatures of chaos on a Floquet synthetic lattice". In: *Physical Review A* 100.1 (2019), p. 013623.

[326] Xiao Mi et al. "Information scrambling in computationally complex quantum circuits". In: *arXiv:2101.08870* (2021).

[327] A Micheli et al. "Cold polar molecules in two-dimensional traps: Tailoring interactions with external fields for novel quantum phases". In: *Physical Review A* 76.4 (2007), p. 043604.

[328] Roderich Moessner and SL Sondhi. "Equilibration and order in quantum Floquet matter". In: *Nature Physics* 13.5 (2017), p. 424. DOI: `10.1038/nphys4106`.

[329] Masoud Mohseni, Ali T Rezakhani, and Daniel A Lidar. "Quantum-process tomography: Resource analysis of different strategies". In: *Physical Review A* 77.3 (2008), p. 032322.

[330] Klaus Molmer and Anders Sorensen. "Multiparticle Entanglement of Hot Trapped Ions". In: *Physical Review Letters* 82 (9 1999), pp. 1835–1838. DOI: `10.1103/PhysRevLett.82.1835`. URL: `http://link.aps.org/doi/10.1103/PhysRevLett.82.1835`.

[331]   Peter K Molony et al. "Creation of ultracold rb 87 cs 133 molecules in the rovibrational ground state". In: *Physical Review Letters* 113.25 (2014), p. 255301.

[332]   C Monroe et al. "Programmable Quantum Simulations of Spin Systems with Trapped Ions". In: *arXiv:1912.07845* (2019).

[333]   Christopher Monroe and Jungsang Kim. "Scaling the ion trap quantum processor". In: *Science* 339.6124 (2013), pp. 1164–1169.

[334]   Joel E Moore and Leon Balents. "Topological invariants of time-reversal-invariant band structures". In: *Physical Review B* 75.12 (2007), p. 121306.

[335]   Joel E Moore, Ying Ran, and Xiao-Gang Wen. "Topological surface states in three-dimensional magnetic insulators". In: *Physical Review Letters* 101.18 (2008), p. 186805.

[336]   Takahiro Morimoto, Hoi Chun Po, and Ashvin Vishwanath. "Floquet topological phases protected by time glide symmetry". In: *Physical Review B* 95.19 (2017), pp. 1–17. ISSN: 24699969. DOI: `10.1103/PhysRevB.95.195155`. eprint: `1703.02553`.

[337]   Steven A Moses et al. "Creation of a low-entropy quantum gas of polar molecules in an optical lattice". In: *Science* 350.6261 (2015), pp. 659–662.

[338]   Steven A Moses et al. "New frontiers for quantum gases of polar molecules". In: *Nature Physics* 13.1 (2017), p. 13.

[339]   Klaus Mÿlmer and Anders Sÿrensen. "Multiparticle Entanglement of Hot Trapped Ions". In: *Physical Review Letters* 82.9 (Mar. 1999), pp. 1835–1838.

[340]   Adam Nahum, Sagar Vijay, and Jeongwan Haah. "Operator spreading in random unitary circuits". In: *Physical Review X* 8.2 (2018), p. 021014.

[341]   Adam Nahum et al. "Quantum entanglement growth under random unitary dynamics". In: *Physical Review X* 7.3 (2017), p. 031016.

[342]   R. K. Naik et al. "Random access quantum information processors using multimode circuit quantum electrodynamics". In: *Nature Communications* 8.1 (Dec. 2017), pp. 1–7.

[343]   Mikio Nakahara. *Geometry, topology and physics*. CRC press, 2018.

[344]   James Nakamura et al. "Aharonov–Bohm interference of fractional quantum Hall edge modes". In: *Nature Physics* 15 (2019), pp. 563–569.

[345]   Rahul Nandkishore and David A Huse. "Many-body localization and thermalization in quantum statistical mechanics". In: *Annu. Rev. Condens. Matter Phys.* 6.1 (2015), pp. 15–38.

[346]   Frederik Nathan and Mark S Rudner. "Topological singularities and the general classification of Floquet–Bloch systems". In: *New Journal of Physics* 17.12 (2015), p. 125014.

[347]   Chetan Nayak et al. "Non-Abelian anyons and topological quantum computation". In: *Reviews of Modern Physics* 80 (3 2008), pp. 1083–1159.

[348] Brian Neyenhuis. "Ultracold Polar KRb Molecules in Optical Lattices". PhD thesis. University of Colorado at Boulder, Jan. 2012.

[349] Brian Neyenhuis et al. "Anisotropic polarizability of ultracold polar k 40 rb 87 molecules". In: *Physical Review Letters* 109.23 (2012), p. 230403.

[350] Sepehr Nezami et al. "Quantum gravity in the lab: teleportation by size and traversable wormholes, Part II". In: *arXiv:2102.01064* (2021).

[351] K-K Ni et al. "A high phase-space-density gas of polar molecules". In: *Science* 322.5899 (2008), pp. 231–235.

[352] Kang-Kuen Ni, Till Rosenband, and David D Grimes. "Dipolar exchange quantum logic gate with polar molecules". In: *Chemical Science* 9.33 (2018), pp. 6830–6838.

[353] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. 2002.

[354] Thomas E O'Brien et al. "Quantum computation of molecular structure using data from challenging-to-classically-simulate nuclear magnetic resonance experiments". In: *arXiv:2109.02163* (2021).

[355] Nissim Ofek et al. "Extending the lifetime of a quantum bit with error correction in superconducting circuits". en. In: *Nature* 536.7617 (Aug. 2016), pp. 441–445.

[356] S. Olmschenk et al. "Manipulation and detection of a trapped $Yb^+$ hyperfine qubit". In: *Physical Review A* 76 (5 2007), p. 052314. DOI: 10.1103/PhysRevA.76.052314. URL: http://link.aps.org/doi/10.1103/PhysRevA.76.052314.

[357] S. Olmschenk et al. "Quantum Teleportation Between Distant Matter Qubits". In: *Science* 323.5913 (2009), pp. 486–489.

[358] Román Orús. "A practical introduction to tensor networks: Matrix product states and projected entangled pair states". In: *Annals of Physics* 349 (2014), pp. 117–158.

[359] S Ospelkaus et al. "Controlling the hyperfine state of rovibronic ground-state polar molecules". In: *Physical Review Letters* 104.3 (2010), p. 030402.

[360] Tomoki Ozawa et al. "Topological photonics". In: *Reviews of Modern Physics* 91 (1 2019), p. 015006.

[361] Don N. Page. "Average entropy of a subsystem". In: *Physical Review Letters* 71 (1993), pp. 1291–1294.

[362] Shriya Pai, Michael Pretko, and Rahul M Nandkishore. "Localization in fractonic random circuits". In: *Physical Review X* 9.2 (2019), p. 021003.

[363] S. Pancharatnam. "Generalized theory of interference, and its applications. Part I. Coherent pencils". In: *Proc. Ind. Acad. Sci. A* 44 (1956), pp. 247–262.

[364] Jee Woo Park, Sebastian A Will, and Martin W Zwierlein. "Ultracold dipolar gas of fermionic na 23 k 40 molecules in their absolute ground state". In: *Physical Review Letters* 114.20 (2015), p. 205302.

[365] Daniel E Parker et al. "A universal operator growth hypothesis". In: *Physical Review X* 9.4 (2019), p. 041017.

[366] A. Pashov et al. "Coupling of the $X^1\Sigma^+$ and $a^3\Sigma^+$ states of KRb". In: *Physical Review A* 76 (2 2007), p. 022511.

[367] Aavishkar A Patel and Subir Sachdev. "Quantum chaos on a critical Fermi surface". In: *Proc. Natl. Acad. Sci.* 114.8 (2017), pp. 1844–1849.

[368] David Peter et al. "Topological bands with a Chern number C= 2 by dipolar exchange interactions". In: *Physical Review A* 91.5 (2015), p. 053617.

[369] Michael J. Peterer et al. "Coherence and Decay of Higher Energy Levels of a Superconducting Transmon Qubit". In: *Physical Review Letters* 114 (1 2015), p. 010501.

[370] Stefano Pirandola et al. "Advances in quantum teleportation". In: *Nature photonics* 9.10 (2015), pp. 641–652.

[371] Stephan Plugge, Étienne Lantagne-Hurtubise, and Marcel Franz. "Revival dynamics in a traversable wormhole". In: *Physical Review Letters* 124.22 (2020), p. 221601.

[372] Lev Semenovič Pontrjagin. "Smooth manifolds and their applications in homotopy theory". In: *Topological Library: Part 1: Cobordisms and Their Applications*. World Scientific, 2007, pp. 1–130.

[373] I.-D. Potirniche et al. "Floquet Symmetry-Protected Topological Phases in Cold-Atom Systems". In: *Physical Review Letters* 119 (2017), p. 123601.

[374] A. Potts et al. "CMOS compatible fabrication methods for submicron Josephson junction qubits". In: *IEE Proceedings - Science, Measurement and Technology* 148.5 (2001), pp. 225–228. ISSN: 1350-2344. DOI: 10.1049/ip-smt:20010395.

[375] John Preskill. "Quantum computing in the NISQ era and beyond". In: *Quantum* 2 (2018), p. 79.

[376] John Preskill. "The Physics of Quantum Information". In: *arXiv:2208.08064* (2022).

[377] HM Price and NR Cooper. "Mapping the Berry curvature from semiclassical dynamics in optical lattices". In: *Physical Review A* 85.3 (2012), p. 033620.

[378] Timothy Proctor et al. "What randomized benchmarking actually measures". In: *Physical Review Letters* 119.13 (2017), p. 130502.

[379] Xiao-Liang Qi and Alexandre Streicher. "Quantum epidemiology: operator growth, thermal effects, and SYK". In: 2019.8 (2019), p. 12.

[380] Xiao-Liang Qi, Yong-Shi Wu, and Shou-Cheng Zhang. "Topological quantization of the spin Hall effect in two-dimensional paramagnetic semiconductors". In: *Physical Review B* 74.8 (2006), p. 085308.

[381] Xiao-Liang Qi et al. "Measuring operator size growth in quantum quench experiments". In: *arXiv:1906.00524* (2019).

[382] Marco Túlio Quintino et al. "Probabilistic exact universal quantum circuits for transforming unitary operations". In: *Physical Review A* 100.6 (2019), p. 062339.

[383] Marco Túlio Quintino et al. "Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations". In: *Physical Review Letters* 123.21 (2019), p. 210502.

[384] Tibor Rakovszky, CW von Keyserlingk, and Frank Pollmann. "Dissipation-assisted operator evolution method for capturing hydrodynamic transport". In: *Physical Review B* 105.7 (2022), p. 075131.

[385] Tibor Rakovszky, Frank Pollmann, and CW von Keyserlingk. "Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation". In: *Physical Review X* 8.3 (2018), p. 031058.

[386] Josef Rammensee, Juan Diego Urbina, and Klaus Richter. "Many-Body Quantum Interference and the Saturation of Out-of-Time-Order Correlators". In: *Physical Review Letters* 121 (12 2018), p. 124101.

[387] N. Read and Dmitry Green. "Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect". In: *Physical Review B* 61 (15 2000), pp. 10267–10297.

[388] Mikael C Rechtsman et al. "Photonic Floquet topological insulators". In: *Nature* 496.7444 (2013), p. 196.

[389] Michael Reck et al. "Experimental realization of any discrete unitary operator". In: *Physical Review Letters* 73.1 (1994), p. 58.

[390] Nicolas Regnault and B Andrei Bernevig. "Fractional chern insulator". In: *Physical Review X* 1.2 (2011), p. 021014.

[391] Ji-Gang Ren et al. "Ground-to-satellite quantum teleportation". In: *Nature* 549.7670 (2017), p. 70.

[392] S Revathi and V Balakrishnan. "Effective diffusion constant for inhomogeneous diffusion". In: *Journal of Physics A: Mathematical and General* 26.21 (1993), p. 5661.

[393] Mark Riebe et al. "Deterministic quantum teleportation with atoms". In: *Nature* 429.6993 (2004), pp. 734–737.

[394] Chad Rigetti and Michel Devoret. "Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies". In: *Physical Review B* 81 (13 2010), p. 134507.

[395] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. "Thermalization and its mechanism for generic isolated quantum systems". In: *Nature* 452.7189 (2008), p. 854.

[396] D. Riste et al. "Millisecond charge-parity fluctuations and induced decoherence in a superconducting transmon qubit". In: *Nature Communications* 4.1 (2013).

[397] Daniel A Roberts, Douglas Stanford, and Alexandre Streicher. "Operator growth in the SYK model". In: 2018.6 (2018), p. 122.

[398] Daniel A Roberts, Douglas Stanford, and Leonard Susskind. "Localized shocks". In: *Journal of High Energy Physics* 2015.3 (2015), pp. 1–27.

[399] Daniel A Roberts and Beni Yoshida. "Chaos and complexity by design". In: *Journal of High Energy Physics* 2017.4 (2017), p. 121.

[400] S. Rosenblum et al. "Fault-tolerant detection of a quantum error". In: *Science* 361 (July 2018).

[401] I. Roth et al. "Recovering Quantum Gates from Few Average Gate Fidelities". In: *Physical Review Letters* 121 (17 2018), p. 170502.

[402] Aidan Roy and A. J. Scott. "Unitary designs and codes". In: *Des. Codes Cryptogr.* 53.1 (2009), pp. 13–31. DOI: 10.1007/s10623-009-9290-2.

[403] Rahul Roy. "Topological phases and the quantum spin Hall effect in three dimensions". In: *Physical Review B* 79.19 (2009), p. 195322.

[404] Rahul Roy and Fenner Harper. "Periodic table for floquet topological insulators". In: *Physical Review B* 96.15 (2017), p. 155118.

[405] Efim B Rozenbaum, Sriram Ganeshan, and Victor Galitski. "Lyapunov exponent and out-of-time-ordered correlator's growth rate in a chaotic system". In: *Physical Review Letters* 118.8 (2017), p. 086801.

[406] Mark S Rudner and Netanel H Lindner. "Floquet topological insulators: from band structure engineering to novel non-equilibrium quantum phenomena". In: *arXiv:1909.02008* (2019).

[407] Mark S Rudner et al. "Anomalous edge states and the bulk-edge correspondence for periodically driven two-dimensional systems". In: *Physical Review X* 3.3 (2013), p. 031005.

[408] Nicholas J Russell et al. "Direct dialling of Haar random unitary matrices". In: *New Journal of Physics* 19.3 (2017), p. 033007.

[409] Shinsei Ryu and Tadashi Takayanagi. "Holographic derivation of entanglement entropy from the anti–de sitter space/conformal field theory correspondence". In: *Physical Review Letters* 96.18 (2006), p. 181602.

[410] Subir Sachdev and Jinwu Ye. "Gapless spin-fluid ground state in a random quantum Heisenberg magnet". In: *Physical Review Letters* 70.21 (1993), p. 3339.

[411] LE Sadler et al. "Spontaneous symmetry breaking in a quenched ferromagnetic spinor Bose–Einstein condensate". In: *Nature* 443.7109 (2006), pp. 312–315.

[412] MS Safronova et al. "Search for new physics with atoms and molecules". In: *Reviews of Modern Physics* 90.2 (2018), p. 025008.

[413] Jeremy M Sage et al. "Optical production of ultracold polar molecules". In: *Physical Review Letters* 94.20 (2005), p. 203001.

[414] Y. Salath et al. "Digital Quantum Simulation of Spin Models with Circuit Quantum Electrodynamics". In: *Physical Review X* 5.2 (2015).

[415] Olof Salberger et al. "Deformed Fredkin spin chain with extensive entanglement". In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.6 (2017), p. 063103.

[416] Claudia M Sánchez, Ana Karina Chattah, and Horacio M Pastawski. "Emergent decoherence induced by quantum chaos in a many-body system: A Loschmidt echo observation through NMR". In: *Physical Review A* 105.5 (2022), p. 052232.

[417] Claudia M Sánchez et al. "Clustering and decoherence of correlated spins under double quantum dynamics". In: *Physical Review A* 90.4 (2014), p. 042122.

[418] Claudia Marina Sánchez et al. "Perturbation independent decay of the loschmidt echo in a many-body system". In: *Physical Review Letters* 124.3 (2020), p. 030601.

[419] Barry C Sanders, Stephen D Bartlett, and Hubert de Guise. "From qubits to continuous-variable quantum computation". In: *arXiv:0208008* (2002).

[420] Mohan Sarovar et al. "Detecting crosstalk errors in quantum information processors". In: *Quantum* 4 (2020), p. 321.

[421] Rahul Sawant et al. "Ultracold polar molecules as qudits". In: *New Journal of Physics* 22.1 (2020), p. 013027.

[422] Clement Sayrin et al. "Real-time quantum feedback prepares and stabilizes photon number states". In: *Nature* 477 (Sept. 2011).

[423] Frank Schindler et al. "Higher-order topological insulators". In: *Science Advances* 4.6 (2018), eaat0346.

[424] Romana Schirhagl et al. "Nitrogen-vacancy centers in diamond: nanoscale sensors for physics and biology". In: *Annual Reviews of Physical Chemistry* 65.1 (2014), pp. 83–105.

[425] Andreas P Schnyder et al. "Classification of topological insulators and superconductors in three spatial dimensions". In: *Physical Review B* 78.19 (2008), p. 195125.

[426] J. A. Schreier et al. "Suppressing charge noise decoherence in superconducting charge qubits". In: *Physical Review B* 77.18 (2008).

[427] Thomas Schuster and Norman Y Yao. "Operator Growth in Open Quantum Systems". In: *arXiv:2208.12272* (2022).

[428] Thomas Schuster et al. "Floquet Hopf insulators". In: *Physical Review Letters* 123.26 (2019), p. 266803.

[429] Thomas Schuster et al. "Many-body quantum teleportation via operator spreading in the traversable wormhole protocol". In: *Physical Review X* 12.3 (2022), p. 031013.

[430] Thomas Schuster et al. "Realizing Hopf Insulators in Dipolar Spin Systems". In: *arXiv:1901.08597* (2019).

[431] Yasuhiro Sekino and Leonard Susskind. "Fast scramblers". In: *Journal of High Energy Physics* 2008.10 (2008), p. 065.

[432] Gordon W Semenoff. "Chiral symmetry breaking in graphene". In: *Phys. Scr.* T146 (2012), p. 014016.

[433] C Senko et al. "Realization of a quantum integer-spin chain with controllable interactions". In: *Physical Review X* 5.2 (2015), p. 021026.

[434] Eyob A. Sete, John M. Martinis, and Alexander N. Korotkov. "Quantum theory of a bandpass Purcell filter for qubit readout". In: *Physical Review A* 92 (1 2015), p. 012325. DOI: 10.1103/PhysRevA.92.012325. URL: https://link.aps.org/doi/10.1103/PhysRevA.92.012325.

[435] Sarah Sheldon et al. "Procedure for systematically tuning up cross-talk in the cross-resonance gate". In: *Physical Review A* 93.6 (2016).

[436] Stephen H Shenker and Douglas Stanford. "Black holes and the butterfly effect". In: *Journal of High Energy Physics* 2014.3 (2014), pp. 1–25.

[437] Stephen H Shenker and Douglas Stanford. "Stringy effects in scrambling". In: 2015.5 (2015), p. 132.

[438] Rajiah Simon. "Peres-Horodecki separability criterion for continuous variable systems". In: *Physical Review Letters* 84.12 (2000), p. 2726.

[439] Brian Skinner, Jonathan Ruhman, and Adam Nahum. "Measurement-induced phase transitions in the dynamics of entanglement". In: *Physical Review X* 9.3 (2019), p. 031009.

[440] J Smith et al. "Many-body localization in a quantum simulator with programmable random disorder". In: *Nature Physics* 12.10 (2016), pp. 907–911.

[441] Akira Sone and Paola Cappellaro. "Hamiltonian identifiability assisted by a single-probe measurement". In: *Physical Review A* 95.2 (2017), p. 022335.

[442] Douglas Stanford. "Many-body chaos at weak coupling". In: *Journal of High Energy Physics* 2016.10 (2016), pp. 1–18.

[443] A. J. Stone. *The Theory of Intermolecular Forces*. London: Clarendon Press, 1996.

[444] Vincent Paul Su. "Variational Preparation of the Sachdev-Ye-Kitaev Thermofield Double". In: *arXiv:2009.04488* (2020).

[445] Leonard Susskind. "Complexity and Newton's Laws". In: *arXiv:1904.12819* (2019).

[446] Leonard Susskind. "Dear Qubitzers, GR= QM". In: *arXiv:1708.03040* (2017).

[447] Leonard Susskind. "The typical-state paradox: diagnosing horizons with complexity". In: *Fortsch. Phys.* 64.1 (2016), pp. 84–91.

[448]  Leonard Susskind. "Why do things fall?" In: *arXiv:1802.01198* (2018).

[449]  Leonard Susskind and Ying Zhao. "Teleportation Through the Wormhole". In: *arXiv:1707.04354 [hep-th]* (2017).

[450]  Brian Swingle and Nicole Yunger Halpern. "Resilience of scrambling measurements". In: *arXiv:1802.01587* (2018).

[451]  Brian Swingle et al. "Measuring the scrambling of quantum information". In: *Physical Review A* 94.4 (2016), p. 040302.

[452]  Sergey V Syzranov et al. "Emergent Weyl excitations in systems of polar particles". In: *Nature communications* 7.1 (2016), pp. 1–7.

[453]  Alexander Szameit and Stefan Nolte. "Discrete optics in femtosecond-laser-written photonic structures". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 43.16 (2010), p. 163001.

[454]  Ryuji Takagi and Quntao Zhuang. "Convex resource theory of non-Gaussianity". In: *Physical Review A* 97 (6 2018), p. 062337.

[455]  Tetsu Takekoshi et al. "Ultracold dense samples of dipolar RbCs molecules in the rovibrational and hyperfine ground state". In: *Physical Review Letters* 113.20 (2014), p. 205301.

[456]  Matthias Tarnowski et al. "Measuring topology from dynamics by obtaining the Chern number from a linking number". In: *Nature Communications* 10.1 (2019), pp. 1–13.

[457]  Jeffrey C. Y. Teo and C. L. Kane. "Topological defects and gapless modes in insulators and superconductors". In: *Physical Review B* 82 (11 2010), p. 115120.

[458]  Yi Hong Teoh et al. "Machine learning design of a trapped-ion quantum spin simulator". In: *Quantum Science and Technology* 5.2 (2020), p. 024001.

[459]  Thomas Theis et al. "Direct and cost-efficient hyperpolarization of long-lived nuclear spin states on universal 15N2-diazirine molecular tags". In: *Science Advances* 2.3 (2016), e1501438.

[460]  David J Thouless et al. "Quantized Hall conductance in a two-dimensional periodic potential". In: *Physical Review Letters* 49.6 (1982), p. 405.

[461]  Paraj Titum et al. "Anomalous floquet-anderson insulator as a nonadiabatic quantized charge pump". In: *Physical Review X* 6.2 (2016), p. 021013.

[462]  Marija Todorić et al. "Quantum Hall Effect with Composites of Magnetic Flux Tubes and Charged Particles". In: *Physical Review Letters* 120.26 (2018), p. 267201.

[463]  Akram Touil and Sebastian Deffner. "Information Scrambling versus Decoherence: Two Competing Sinks for Entropy". In: *PRX Quantum* 2.1 (2021), p. 010306.

[464]  Minh C Tran et al. "Hierarchy of linear light cones with long-range interactions". In: *arXiv:2001.11509* (2020).

[465] A Trautmann et al. "Dipolar quantum mixtures of erbium and dysprosium atoms". In: *Physical Review Letters* 121.21 (2018), p. 213601.

[466] Vinay Tripathi, Mostafa Khezri, and Alexander N. Korotkov. "Operation and intrinsic error budget of a two-qubit cross-resonance gate". In: *Physical Review A* 100.1 (2019).

[467] Daniel C Tsui, Horst L Stormer, and Arthur C Gossard. "Two-dimensional magneto-transport in the extreme quantum limit". In: *Physical Review Letters* 48.22 (1982), p. 1559.

[468] Shih-Kuang Tung et al. "Ultracold mixtures of atomic 6 Li and 133 Cs with tunable interactions". In: *Physical Review A* 87.1 (2013), p. 010702.

[469] F Nur Ünal, André Eckardt, and Robert-Jan Slager. "Hopf characterization of two-dimensional Floquet topological insulators". In: *Physical Review Research* 1.2 (2019), p. 022003.

[470] Giovannetti V. et al. "Ultimate classical communication rates of quantum optical channels". In: *Nat Photon* 8.10 (Oct. 2014), pp. 796–800.

[471] Giacomo Valtolina et al. "Dipolar evaporation of reactive molecules to below the Fermi temperature". In: *Nature* 588.7837 (2020), pp. 239–243.

[472] Victor Veitch et al. "The resource theory of stabilizer quantum computation". In: *New Journal of Physics* 16.1 (2014), p. 013009.

[473] Benoît Vermersch et al. "Probing scrambling using statistical correlations between randomized measurements". In: *Physical Review X* 9.2 (2019), p. 021061.

[474] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272.

[475] Brian Vlastakis et al. "Deterministically encoding quantum information using 100-photon Schrödinger cat states". In: *Science* 342.6158 (2013), pp. 607–610.

[476] CW Von Keyserlingk et al. "Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws". In: *Physical Review X* 8.2 (2018), p. 021013.

[477] David J Wales and Jonathan PK Doye. "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms". In: *The Journal of Physical Chemistry A* 101.28 (1997), pp. 5111–5116.

[478] Xiangang Wan et al. "Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates". In: *Physical Review B* 83.20 (2011), p. 205101.

[479] Ce Wang et al. "Scheme to Measure the Topological Number of a Chern Insulator from Quench Dynamics". In: *Physical Review Letters* 118 (18 2017), p. 185701.

[480] Jianwei Wang et al. "Experimental quantum Hamiltonian learning". In: *Nature Physics* 13.6 (2017), pp. 551–555.

[481]   Sheng-Tao Wang, Dong-Ling Deng, and Lu-Ming Duan. "Hamiltonian tomography for quantum many-body systems with arbitrary couplings". In: *New Journal of Physics* 17.9 (2015), p. 093017.

[482]   Zheng Wang et al. "Observation of unidirectional backscattering-immune topological electromagnetic states". In: *Nature* 461.7265 (2009), pp. 772–775.

[483]   John Watrous. *The theory of quantum information*. Cambridge University Press, 2018.

[484]   Zak Webb. "The Clifford group forms a unitary 3-design". In: *arXiv:1510.02769* (2015).

[485]   Christian Weedbrook et al. "Gaussian quantum information". In: *Reviews of Modern Physics* 84 (2 2012), pp. 621–669. DOI: `10.1103/RevModPhys.84.621`.

[486]   Ken Xuan Wei, Chandrasekhar Ramanathan, and Paola Cappellaro. "Exploring localization in nuclear spin chains". In: *Physical Review Letters* 120.7 (2018), p. 070501.

[487]   Ken Xuan Wei et al. "Emergent prethermalization signatures in out-of-time ordered correlations". In: *Physical Review Letters* 123.9 (2019), p. 090605.

[488]   Xiao-Gang Wen. "Topological order: From long-range entangled quantum matter to a unified origin of light and electrons". In: *International Scholarly Research Notices* (2013).

[489]   Christopher David White et al. "Quantum dynamics of thermalizing systems". In: *Physical Review B* 97.3 (2018), p. 035127.

[490]   Steven R White. "Density matrix formulation for quantum renormalization groups". In: *Physical Review Letters* 69.19 (1992), p. 2863.

[491]   Nathan Wiebe et al. "Quantum Hamiltonian learning using imperfect quantum resources". In: *Physical Review A* 89.4 (2014), p. 042314.

[492]   Frank Wilczek. "Magnetic Flux, Angular Momentum, and Statistics". In: *Physical Review Letters* 48 (17 1982), pp. 1144–1146.

[493]   Frank Wilczek. "Quantum Mechanics of Fractional-Spin Particles". In: *Physical Review Letters* 49 (14 1982), pp. 957–959.

[494]   Mark M Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[495]   R. L. Willett et al. "Magnetic-Field-Tuned Aharonov-Bohm Oscillations and Evidence for Non-Abelian Anyons at $\nu = 5/2$". In: *Physical Review Letters* 111 (18 2013), p. 186401.

[496]   J. Wilson et al. "Trapped arrays of alkaline earth Rydberg atoms in optical tweezers". In: *arXiv:1912.08754* (2018).

[497]   Martin Wimmer et al. "Experimental measurement of the Berry curvature from anomalous transport". In: *Nature Physics* 13.6 (2017), pp. 545–550.

[498]   Edward Witten. "An SU(2) anomaly". In: *Physics Letters B* 117.5 (1982), pp. 324–328.

[499] Edward Witten. "Current algebra, baryons, and quark confinement". In: *Nuclear Physics, Section B* 223.2 (1983), pp. 433–444.

[500] Edward Witten. "Quantum field theory and the Jones polynomial". English. In: *Commun. Math. Phys* 121.3 (1989), pp. 351–399. ISSN: 0010-3616.

[501] K. Wright et al. "Benchmarking an 11-qubit quantum computer". In: *Nat. Commun.* 10 (2019), p. 5464.

[502] Jingxiang Wu and Timothy H. Hsieh. "Variational Thermal Quantum Simulation via Thermofield Double States". In: *Physical Review Letters* 123 (22 2019), p. 220502.

[503] T. Xia et al. "Randomized Benchmarking of Single-Qubit Gates in a 2D Array of Neutral-Atom Qubits". In: *Physical Review Letters* 114 (2015), p. 100503.

[504] Shenglong Xu and Brian Swingle. "Accessing scrambling using matrix product operators". In: *Nature Physics* 16.2 (2020), pp. 199–204.

[505] Shenglong Xu and Brian Swingle. "Locality, quantum fluctuations, and scrambling". In: *arXiv:1805.05376* (2018).

[506] Shenglong Xu and Brian Swingle. "Scrambling Dynamics and Out-of-Time Ordered Correlators in Quantum Many-Body Systems: a Tutorial". In: *arXiv:2202.07060* (2022).

[507] Bo Yan et al. "Observation of dipolar spin-exchange interactions with lattice-confined polar molecules". In: *Nature* 501 (2013), pp. 521–525.

[508] Bo Yan et al. "Realizing a lattice spin model with polar molecules". In: *arXiv preprint arXiv:1305.5598* (2013).

[509] Zhongbo Yan et al. "Nodal-link semimetals". In: *Physical Review B* 96.4 (2017), p. 041103.

[510] Yi Yang et al. "Synthesis and observation of non-Abelian gauge fields in real space". In: *Science* 365.6457 (2019), 1021–ĂŞ1025.

[511] N. Y. Yao et al. "A quantum dipolar spin liquid". In: *Nature Physics* 14.4 (Apr. 2018), pp. 405–410.

[512] Norman Y Yao et al. "Interferometric approach to probing fast scrambling". In: *arXiv:1607.01801* (2016).

[513] Norman Y Yao et al. "Realizing fractional Chern insulators in dipolar spin systems". In: *Physical Review Letters* 110.18 (2013), p. 185302.

[514] Amnon Yariv. *Optical electronics; 4th ed.* The Holt, Rinehart and Winston series in electrical engineering. 1991.

[515] Beni Yoshida. "Firewalls vs. scrambling". In: 2019.10 (2019), p. 132.

[516] Beni Yoshida. "Observer-dependent black hole interior from operator collision". In: *arXiv:1910.11346* (2019).

[517] Beni Yoshida and Alexei Kitaev. "Efficient decoding for the Hayden-Preskill protocol". In: *arXiv:1710.03363* (2017).

[518] Beni Yoshida and Norman Y Yao. "Disentangling scrambling and decoherence via quantum teleportation". In: *Physical Review X* 9.1 (2019), p. 011006.

[519] Bin Yu. "Assouad, Fano, and Le Cam". In: *Festschrift for Lucien Le Cam*. Springer, 1997, pp. 423–435.

[520] X.-X. Yuan et al. "Observation of Topological Links Associated with Hopf Insulators in a Solid-State Quantum Simulator". In: *Chinese Physics Letters* 34.6, 060302 (2017), p. 060302.

[521] Michael P Zaletel and Frank Pollmann. "Isometric tensor network states in two dimensions". In: *Physical Review Letters* 124.3 (2020), p. 037201.

[522] J. Zeiher et al. "Coherent Many-Body Spin Dynamics in a Long-Range Interacting Ising Chain". In: *Physical Review X* 7 (2017), p. 041063.

[523] Johannes Zeiher et al. "Coherent many-body spin dynamics in a long-range interacting Ising chain". In: *Physical Review X* 7.4 (2017), p. 041063.

[524] Bei Zeng et al. *Quantum information meets quantum matter*. Springer, 2019.

[525] Haijun Zhang et al. "Topological insulators in Bi 2 Se 3, Bi 2 Te 3 and Sb 2 Te 3 with a single Dirac cone on the surface". In: *Nature Physics* 5.6 (2009), p. 438.

[526] J Zhang et al. "Observation of a discrete time crystal". In: *Nature* 543.7644 (2017), p. 217.

[527] Jun Zhang and Mohan Sarovar. "Quantum Hamiltonian identification from measurement time traces". In: *Physical Review Letters* 113.8 (2014), p. 080401.

[528] Pengfei Zhang, Yingfei Gu, and Alexei Kitaev. "An obstacle to sub-AdS holography for SYK-like models". In: *arXiv:2012.01620* (2020).

[529] Yong-Liang Zhang, Yichen Huang, Xie Chen, et al. "Information scrambling in chaotic systems with dissipation". In: *Physical Review B* 99.1 (2019), p. 014303.

[530] Hengyun Zhou et al. "Quantum metrology with strongly interacting spin systems". In: *Physical Review X* 10.3 (2020), p. 031003.

[531] Tianci Zhou et al. "Operator Lévy Flight: Light Cones in Chaotic Long-Range Interacting Systems". In: *Physical Review Letters* 124 (18 2020), p. 180601.

[532] Xiangfa Zhou et al. "Unconventional states of bosons with the synthetic spin–orbit coupling". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 46.13 (2013), p. 134001.

[533] D. Zhu et al. "Generation of Thermofield Double States and Critical Ground States with a Quantum Computer". In: *arXiv:1906.02699* (2019).

[534] Huangjun Zhu. "Multiqubit Clifford groups are unitary 3-designs". In: *Physical Review A* 96 (6 2017), p. 062336.

[535] S.-L. Zhu, C. Monroe, and L.-M. Duan. "Arbitrary-speed quantum gates within large ion crystals through minimum control of laser beams". In: *Europhysics Letters* 73 (4 2006), p. 485.

[536] Shi-Liang Zhu, C. Monroe, and L.-M. Duan. "Trapped Ion Quantum Computation with Transverse Phonon Modes". In: *Physical Review Letters* 97 (5 2006), p. 050505.

[537] Quntao Zhuang, Peter W. Shor, and Jeffrey H. Shapiro. "Resource theory of non-Gaussian operations". In: *Physical Review A* 97 (5 2018), p. 052317. DOI: `10.1103/PhysRevA.97.052317`.

[538] Quntao Zhuang and Biao Wu. "Equilibration of quantum chaotic systems". In: *Physical Review E* 88.6 (2013), p. 062147.

[539] Quntao Zhuang and Biao Wu. "Quantum thermalization and equilibrium state with multiple temperatures". In: *Laser Phys. Lett.* 11.8 (2014), p. 085501.

[540] Quntao Zhuang et al. "Scrambling and complexity in phase space". In: *Physical Review A* 99.6 (2019), p. 062334.

[541] Philip Zupancic et al. "Ultra-precise holographic beam shaping for microscopic quantum control". In: *Optics express* 24.13 (2016), pp. 13881–13893.