UC San Diego

Title

Reject Local Bases, Embrace Entanglement Features

Permalink

Author

Akhtar, Ahmed Ali

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Reject Local Bases, Embrace Entanglement Features

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Physics

by

Ahmed Ali Akhtar

Committee in charge:

        Professor Yi-Zhuang You, Chair
        Professor Tarun Grover
        Professor Ken Intriligator
        Professor John McGreevy
        Professor David Meyer

2023

The Dissertation of Ahmed Ali Akhtar is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

For my parents, who raised me to love science.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Professor Yi-Zhuang You, for being continually and effortlessly inspiring, inquisitive, and insightful.

I would also like to acknowledge some of my friends, mentors, and teachers: Hong-Ye, Namit, Jeff, John, and Tarun. I would like to thank Peter and Avaneesh for countless late night conversations about physics. I would also like to thank Aidee for their support throughout the years.

VITA

2017   Bachelor of Arts, Princeton University

2020–2023 Ph.D. Candidate
University of California San Diego

2023   Doctor of Philosophy, University of California San Diego

ABSTRACT OF THE DISSERTATION

Reject Local Bases, Embrace Entanglement Features

by

Ahmed Ali Akhtar

Doctor of Philosophy in Physics

University of California San Diego, 2023

Professor Yi-Zhuang You, Chair

The nonlocality of quantum information is an intrinsic consequence of the interactions in many-body systems. Quantifying it can reveal how information between parties propogates across space and time, but in general it is hard to keep track of the exponentially-large set of bipartite entanglement data. In this dissertation, we exploit statistic symmetries about local-basis invariant ensembles to efficiently characterize the bipartite entanglement data in the form of the average purity, or entanglement features. We are able to systematically study the entanglement features of a variety of locally-scrambled ensembles, and determine their entanglement velocities, volume-law coefficients, butterfly velocities, etc. Next, we propose a two-parameter, bond dimension two, matrix-product state ansatz for characterizing the full, multiregion entanglement

data of typical locally-scrambled state ensembles. We are able to compare this two-parameter ansatz to a variety of different models and show that it performs well at capturing the full evolution from product to Page state. We also study the meaning of the parameters physically and how they determine different aspects of the multiregion entanglement. Next, we adapt these techniques to classical shadow tomography, where the goal is to construct efficient, classical approximations of quantum states. We show that the reconstruction map is solely a function of the entanglement features of the snapshot state ensemble and we utilize our efficient description of the entanglement data to perform shallow circuit tomography scalably and flexibly. We show that shallow circuit tomography can perform better than Pauli or Clifford measurements for predicting quasi-local operators. We also demonstrate how shallow circuit tomography can serve as a more feasible alternative to Clifford measurements for fidelity estimation, a desirable task on current quantum devices. Lastly, we study the interplay between measurements and entanglement generating unitaries by studying the entanglement features of classical snapshot states produced from hybrid circuits, and show that such schemes are optimal at the phase transition in the purity.

Chapter 1

Primer on Entanglement, Measurement, and Randomness

Much of this disseratation revolves around three recurring themes: entanglement, measurements, and randomness. Therefore, in this chapter, we will introduce some of the core concepts and technical machinery related to these concepts, focusing specifically on how they will be applied in the following chapters. All three of these concepts come together when we discuss an efficient prediction scheme based on forming classical estimators of quantum states.

The profound ramifications of quantum entanglement have been explored at least as early as 1935, when Einstein, Podolsky, and Rosen (EPR) took to answering the question, "Can Quantum-Mechanical Description of Physical Reality be Considered Complete[1]?" The crux of their argument rests on the existence of non-factorizable states in a tensor product Hilbert space, and goes as follows: consider a pair of maximally entangled particles (e.g. in a Bell-pair state) separated at a very large distance. If property Q of particle A is measured at time t_0, the state w.r.t. Q of the particle B must be determined at all times $t > t_0$. This implies the existence of a so-called "element of reality" associated to particle B, and by locality, this element of reality must have a well-defined value for Q for times $t < t_0$. Since quantum mechanics does not admit a description of this hidden variable prior to t_0, it must be an incomplete theory [56]. Decades later,

---

[1]The authors are taking inspiration from Gödel's incompleteness theorems and thinking about physics as a formal, axiomatic system. In this context, *completeness* means every element of physical reality must be describable in the language of the theory.

Bell went on to show that while one could construct local, hidden variable theories to reproduce the singlet correlations in the EPR paradox, no local, hidden variable theory could reproduce the correlations of measurements along arbitrary axes that are predicted by quantum mechanics and verified in experiment [18]. In other words, the non-locality of quantum information, i.e. entanglement, is an intrinsic feature of the real world, not a bug.

Nevertheless, as entanglement is a major source of confusion for physicists committed to a classical description of reality, so too is measurement. Although incorporating the non-linear collapse of the wavefunction into the formalism of quantum mechanics, via e.g. positive-operator valued measurement operators, is clunky, it works. Still, the philosophical ramifications has spawned various fascinating interpretations of quantum mechanics such as the Everettian view, which asserts that every measurement signifies a branching point of different universes [217]. Measurement plays an important role in the EPR paradox as well: one of the premises is that being measurable is closely tied to something being a real or physical [190]. Measurement is also interesting because, when performed locally, it tends to disentangle degrees of freedom. This is in contrast to generic time evolution, which due to interactions tends to generate entanglement between nearby degrees of freedom. The competition between measurement and entanglement-generating unitaries will be an important theme in this disseratation.

That leaves randomness. At the heart of all physics is randomness. It is baked into quantum uncertainty, the essential randomness of a wavefunction. Frequently, one voluntarily introduces randomness into a quantum system to study the generic, basis-independent, or universal properties of a family of systems. Far from being obfuscating, it may be injected to simplify the physics by erasing or "averaging-out" the irrelevant information, and has been utilized to study the universal dynamical behavior of entanglement [37, 127, 154, 216, 243]. In this chapter, we'll learn how to average over random unitaries and apply that to the study of purity dynamics. Before doing that, we'll need a brief primer on what entanglement is all about.

Let us establish some notational practices that will be followed in this thesis:

- We follow bra-ket notation i.e. states are denoted $|\psi\rangle$ and their adjoint $\langle\psi|$; inner-products are written as $\langle\psi|\phi\rangle$; norms, e.g. $\langle\psi|\psi\rangle$, are denoted with double-bars $\||\psi\rangle\|^2$; etc.

- Recurring symbols:

  - $\mathcal{H}$ - Hilbert space.

  - N - number of qudits.

  - d - qudit dimension. For qubits, $d = 2$.

  - Ω - refers to lattice of qudits, or the set of qudits $\{1...N\}$. Capital english letters refer to subsets $A, B, C \subset \Omega$.

  - S - entanglement measure.

  - X, Y, Z denote the Pauli matrices. Pauli strings are denoted the symbol P. The subscript i, e.g. X_i, P_i, denotes a Pauli matrix at site i.

- Renyi-entropies are denoted as $S^{(n)}$. Entanglement-entropy is denoted as $S^{(1)}$. The symbol S without any super-script denotes a general, unspecified entanglement measure.

- Purity is denoted with a capital "W", whereas Pauli weight is generally denoted with a lowercase "w". These quantities are related by a local basis transformation. For entanglement features, the symbol σ usually refers to an element of the permutation group S_2, whereas σ in the context of classical shadow tomography refers to a snapshot state.

1.1 An brief introduction to entanglement in many-body systems

Evidently, entanglement characterizes how quantum information is shared between different parties. When a pure state $|\psi\rangle$ can be factored between two parties A and B s.t. $|\psi\rangle = |\chi\rangle_A |\chi\rangle_B$, measuring A tells you nothing about the B (and vice versa). In this situation, we say the state is a product state, and the two parties don't share any quantum information. On

the other hand, when the wavefunction cannot be factored between parties s.t. $|\psi\rangle \neq |\chi\rangle_A |\chi\rangle_B$, measuring A should in principle tell you something about B; namely, the information that is shared between the two. One way to think about the "factorizability" of a wavefunction $|\psi\rangle \in \mathcal{H}$ on a product space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is by considering its Schmidt decomposition between the factor spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Given a bipartition A, B of the full system (e.g. A, B could be complementary subsets of qubits situated on a lattice), and $|\psi\rangle \in \mathcal{H} \cong \mathcal{H}_A \otimes \mathcal{H}_B$, we may decompose $|\psi\rangle$ as a sum over $m = \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)$ disentangled basis states,

$$|\psi\rangle = \sum_{i=1}^{m} \alpha_i |u_i\rangle_A |v_i\rangle_B, \tag{1.1}$$

where $\mathcal{B}_A = \{|u_1\rangle_A, ..., |u_m\rangle_A\}$ and $\mathcal{B}_B = \{|v_1\rangle_B, ..., |v_m\rangle_B\}$ are orthonormal basis kets on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. The Schmidt decomposition may be derived using the Singular-Value Decomposition of the state viewed as a matrix [195]. The Schmidt coefficients $\{\alpha_i\}_{i=1...m}$ are non-negative, ordered, real numbers: $\alpha_1 \geq \alpha_2 \geq ... \geq 0$. Furthermore, their squares $\{\alpha_i^2\}_{i=1...m}$ form the diagonal elements of ρ_A (ρ_B) in the $\mathcal{B}_A$ ($\mathcal{B}_B$) basis. Therefore, they are constrained by the normalization of $|\psi\rangle$ to satisfy $\sum_i^m \alpha_i^2 = 1$. This decomposition is unique up to phase changes in the basis states, which in principle could be highly entangled within their respective regions.

From the Schmidt decomposition, we see that $|\psi\rangle$ is a product state, i.e. disentangled, if and only if $\alpha_1 = 1$. We also see that "spreading out" the Schmidt weights α_i^2 makes the state less "factorizable" or more entangled. For the Bell pair states, the weight is equally distributed over all Schmidt coefficients, $\alpha_1 = \alpha_2 = 1/\sqrt{2}$. These states must have the maximal entanglement since measuring a one qubit tells you everything about both qubits i.e. they mutually share two bits of information. This suggests that the entanglement could be characterized by the entropy of the Schmidt weights α_i^2 viewed as probabilities, and motivates the definition of the Renyi entanglement-entropies[151] $S^{(n)}(A)$:

$$S^{(n)}(A) := \frac{1}{1-n} \log\left(\mathrm{Tr}\,\rho_A^n\right). \tag{1.2}$$

4

It is easy to check that $S^{(n)}(A) = H_n(p_A)$ where H_n is the classical Renyi entropy of order n and p_A are the diagonal elements of ρ_A in the Schmidt basis, which can be viewed as a probability distribution due to the normalization of the Schmidt weights, $(p_A)_i = \alpha_i^2$. Equation 1.2 seems ill-defined at $n = 1$, but it reduces to the Von-Neumann entanglement-entropy by taking the limit $n \to 1$ and using L'hopital's rule.[2]

$$S^{(1)}(A) := \lim_{n \to 1} S^{(n)}(A) := -\lim_{n \to 1} \frac{\mathrm{Tr}\,\rho_A^n \log(\rho_A)}{\mathrm{Tr}\,\rho_A^n} = -\mathrm{Tr}(\rho_A \log \rho_A). \tag{1.3}$$

In Fig. 1.1, we consider the ($\log 2$ normalized) Renyi entropy for a two-qubit system $S^{(n)}(A)$, where the subsystem A is one of the qubits, in terms of first Schmidt coefficient α. The first Schmidt coefficient uniquely characterizes the entropy because by state normalization the second Schmidt coefficient must be $1 - \alpha$. When ρ_A is pure state, it is a projector onto the wavefunction and therefore has one non-trivial, unital eigenvalue. This situation corresponds to $\alpha = 0$ or 1. Since a pure quantum state is a complete description of a system, pure states cannot share any information with their environment, so our measure of entanglement must vanish when $\alpha = 0, 1$. The simplest measure of entanglement would be 0 if ρ_A is pure, and 1 if it is mixed–this is precisely the zeroeth Renyi entropy. The next Renyi entropy, $n = 1$, is the average of the log Schmidt weights. The second Renyi entropy, $S^{(2)}(A)$, also characterizes how mixed the reduced density matrix is, and is of particular interest because it only depends on the second moment of the ρ_A, meaning that it can be treated much more easily analytically. In this thesis, we will focus predominantly on the second Renyi entropy or its exponential, $W^{(2)}(A) = \exp(-S^{(2)}(A))$, denoted the purity.

The Renyi and entanglement-entropy satisfy a number of different properties that make them amenable to our intuition about entanglement. For example, both Renyi and Von Neumann entropy are additive for products of independent states i.e. $S^{(n)}(\rho_A \otimes \rho_B) = S^{(n)}(\rho_A) + S^{(n)}(\rho_B)$. Intuitively, the entanglement of producted states should simply add together, because there is

---

[2]In the derivation below, the log of a density matrix is defined on its support.
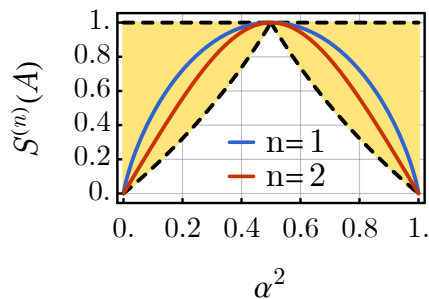
Figure 1.1. We plot the $\log 2$ normalized Renyi entropy of a two-qubit system as a function of the Schmidt weight α^2. For all Renyi entropies, the entropy is zero at $\alpha = 0, 1$, and increases monotonically until $\alpha = 1/\sqrt{2}$, where they take value 1, and then decreases symmetrically back to zero. The midpoint corresponds to a maximally-entangled state on the two sites, which after tracing out one of the qubits gives a maximally-mixed reduced density matrix $\rho = \mathbb{1}/2$. The top boundary is $S^{(0)}(A)$ and the bottom boundary is $S^{(\infty)}(A)$, and the Renyi entropy decreases monotonically in n for fixed α.

no redundant information shared between the two. However, the entanglement-entropy some stronger constraints, such as strong sub-additivity [164]. In Tab. 1.1, we compare and contrast some different properties satisfied by the Renyi entropies. In the following, we will use $S^{(n)}$ to refer specifically to a Renyi entropy measure, and we will use S to denote an arbitrary entanglement measure.

In condensed matter, we are interested in the physics of many-qubit systems, and so we want to extend our notions of entanglement to many-particle systems. All the physics in this disseration will take place on a lattice of N qudits $\Omega = \{1...N\}$. The Hilbert space $\mathscr{H}$ is a tensor product of N d-dimensional Hilbert spaces $\mathscr{H}_i \equiv \mathbb{C}^d$, with $i = 1...N$,

$$\mathscr{H} = \otimes_{i=1}^{N} \mathscr{H}_i, \tag{1.4}$$

so that the overall dimension D of the full Hilbert space $\mathscr{H}$ scales exponentially in the system size: $D = d^N$. We will often be interested in the limit where N is large, because that is the setting of many condensed-matter systems; we will also frequently want to study $d = 2$, which corresponds to a lattice of qubits, and also $d \to \infty$ as an analyitcally tractable limit.

Each of the 2^N sub-regions $A \subset \Omega$ has a corresponding entanglement with its complement. The entanglement of the trivial region O is always zero. Consider a simple, contiguous sub-region A with volume $|A|$: in one dimension, this would be an interval; in two dimensions, a disk; in three dimensions, a sphere; and so on. In Fig. 1.2 (a), we see the depiction of a square, contiguous region in a two-dimensional lattice. We can ask: how does the entanglement of region A scale with its size? In a many-body quantum system, we are free to explore the full space-time structure of the entanglement. The space-time structure of the entanglement is highly dependent on the state (or phase) under consideration. Below, we consider some examples of different kinds of entanglement one encounters in various physical settings.
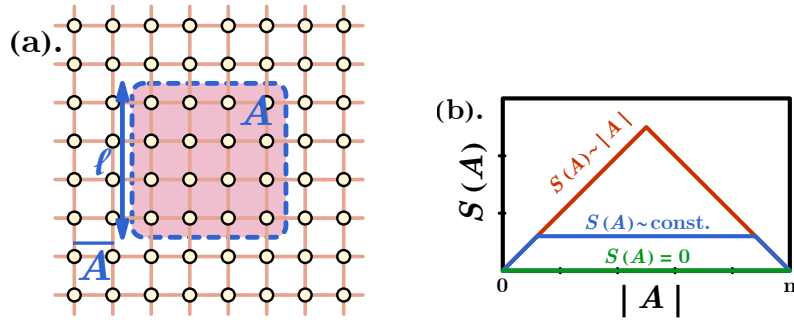
Figure 1.2. (a). A square sub-region A on a two-dimensional lattice. Yellow dots represent different qudits. In an area law phase, information shared between A and $\bar{A}$ happens primarily through the qubits on the boundary. In this case, we say the state is short-range entangled and the $S \sim l$ where l is the side-length. In a volume-law state, the entanglement scales as $S \sim l^2$. We say the system has long-range entanglement in this case, because qubits deep in the bulk of A share information with qubits in $\bar{A}$. (b). We show the stages of entanglement scaling that one typically observes in quenched evolution: product state (green), area-law state (blue), and volume-law state (red). The x-axis is the size $|A|$ of a contiguous subregion A in a one-dimensional system, and the y-axis is the entanglement of A. Typically, one prepares the state initially in a product state, which by construction has zero entanglement for every sub-region $A \subset \Omega$. Hence, the green line is at zero for all values of $|A|$. Then, as the system evolves, it moves through an area-law scaling for which the entanglement scales like the boundary size of A. In one dimension, an interval always has the same boundary size (2) irregardless of $|A|$, hence the entanglement levels off to a constant for the blue line. Finally, the entanglement scaling enters a volume-law phase, for which $S \propto |A|$. Hence the system is linear in $|A|$ for the red line.

- **Product-states:** A product state is a tensor product of local states $|\psi\rangle = \otimes_{i=1}^{N} |\psi_i\rangle$. Product-states carry no entanglement for any sub-region, $S(A) = 0 \forall A \subset \Omega$, because individual

qubits share no information between themselves. Though entanglement is generically expected in quantum systems, there are nonetheless many examples of physical states which are product-states (or close to them).

– Ground states of trivial Hamiltonians, e.g. $H_0 = -\sum_i X_i$. This Hamiltonian has ground state $|+\rangle^{\otimes N}$ where $|+\rangle$ is the $+1$-eigenstate of Pauli matrix X.

– Initial configurations of quenched experiments. Here, one starts from a random product state and records its evolution by a chaotic Hamiltonian. Though such a state has no entanglement, it typically represents a highly-energetic, complicated superposition of eigenstates.

– Mean-field approximated states. Such states can often serve as simple approximations of the correlated eigenstates of interacting Hamiltonians.

• **Area-law entangled states:** These states are ubiquitious in quantum systems. Such states arise when quantum correlations are short-ranged, so that information is only shared between nearby degrees of freedom. For a given entanglement cut, the enclosed sub-region A only shares information with its complement $\bar{A}$ from those qubits along its boundary ∂A, hence $S \sim |\partial A|$. Fig. 1.2 (b) depicts area-law scaling in one dimension.

– Unique ground states of one-dimensional, gapped Hamiltonians [91]. It is widely believed that low-energy states of gapped, local Hamiltonians generically follow an area-law in any dimension. Similarly, it has been shown that any states with finite correlation length in one dimension satisfy an area law [26].

– Short-time evolved product states. For evolution generated by a gapped, local Hamiltonian, entanglement is only generated at the boundary of a cut. Likewise, any area-law states evolved by a short-depth, local unitary circuit must also satisfy an area law [221].

- Many-body localized states [161]. Area-law entanglement is an identifying feature of many-body localization, because typical thermal states will have volume-law scaling.

- All matrix-product states and their higher-dimensional equivalents, or states which are well-described by the aforementioned states [49].

- Free bosonic and fermionic models which are gapped [178].

- Quantum fields propogating in the vicinity of black holes [25]. This is closely related to the famous Bekenstein bound, which says that a black hole's entropy is proportional to its surface area.

- Generic quantum states with positive coefficients [79]. This suggests that entanglement is closely related to the sign structure of the underlying state, and motivates the definition of other measures of entanglement such as the negativity [246].

- **Volume-law entangled states:** Such states typically exhibit long-range correlations, which enable information to be shared deep within an entanglement region A with qubits in its complement $\bar{A}$. Fig. 1.2 (b) depicts volume-law scaling in one dimension.

  - In analogy with thermal, classical entropy, the entanglement entropy of a thermal Gibbs' state scales as a volume law. This can be easily shown the Hamiltonian eigenbasis.

  - Quenched evolution under chaotic Hamiltonians tends to flow towards volume law configurations. Similarly, eigenstates of generic Hamiltonians at finite energy density exhibit a volume law [22, 161, 211].

  - Random states on n qubits generically exhibit volume law scaling, which also can be easily shown using Haar averaging techniques.

This is not a complete classification of the sorts of entanglement structure one observes in physical systems: quantum critical or gapless phenomena will exhibit scaling which is between

9

an area-law and volume law. For example, it is known in conformal field theories that there is logarithmic scaling of the entanglement entropy [30]. In mixed circuits where there is competition between measurement and entangling unitaries, the boundary of the area-law and volume-law phases exhibits logarithmic corrections to the the entanglement scaling [106, 112, 123, 198].

Measurements are interesting because they connect quantum systems which exhibit properties such as superposition and entanglement to the observed, marcoscopic classical world. In doing so, they distenangle a local degree of freedom of the system. Let us now briefly review how measurement works at the technical, state level.

1.2 The formalism of measurements

Let's recap the rules of elementary quantum mechanics in a paragraph. We can associate to any closed physical system a Hilbert space $\mathscr{H}$, and to every state of the physical system a vector $|\psi\rangle \in \mathscr{H}$. We will refer to the state of the system and its corresponding representation in $\mathscr{H}$ interchangeably. Time evolution occurs via unitary operators, i.e. $|\psi\rangle \to U|\psi\rangle$, and the generator of time translations is the Hamiltonian H of the system.

$$H|\psi\rangle = i\partial_t|\psi\rangle \tag{1.5}$$

For every physical observable (e.g. energy, momentum, position, etc.), there is a corresponding hermitian operator $O \in \mathscr{L}(\mathscr{H}), O = O^\dagger$. For simplicity we will again conflate the two. The eigenvalues λ_i of O are the allowed measurement outcomes, and the corresponding eigenvectors are $|\lambda_i\rangle$. The measurement postulate says that if we measure O, the probability of observing outcome λ_i is given by $|\langle\lambda_i|\psi\rangle|^2$, and after measurement the state "collapses" to the corresponding eigenvector: $|\psi\rangle \to |\lambda_i\rangle$ [164, 190, 197, 201].

Without doing any calcuations, one can see how local measurements should in principle reduce entanglement. If the measurement outcome was b for some local observable, then in the eigenbasis of this local measurement operator, the wavefunction is $|\psi\rangle = |b\rangle \otimes \ldots$, where

10

... is some state supported on the remaining degrees of freedom. The measured qubit will not contribute any entanglement to any region that it is included in, and is thereby disentangled from the rest of the system.

Often, we don't have access to the full quantum system but only a subset of it. In this case, rather than being given by a vector, the state of the system is defined by a density matrix $\rho \in \mathcal{H}$. To have a proper state interpretation, the density matrix must be hermitian, have unit trace, and be positive semidefinite.

$$\text{(i). } \rho = \rho^\dagger. \quad \text{(ii). } \text{Tr}\,\rho = 1. \quad \text{(iii). } \rho \succeq 0. \tag{1.6}$$

It's clear that every pure state $|\psi\rangle$ can be represented as a density matrix ρ, via $\rho = |\psi\rangle\langle\psi|$, satisfying the three properties above. However, density matrices allow for more nuanced descriptions of physical systems, namely probabilistic ensembles of pure states e.g. $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $\sum_i p_i = 1$. In the density matrix setting, we can revise the basic postulates of quantum mechanics, e.g. time evolution translates into $\rho \to U\rho U^\dagger$, $[H, \rho] = i\partial_t \rho$, etc.

Similarly, measurements may be generalized in terms of collections of measurement operators $\{M_m | M_m \in \mathscr{L}(\mathscr{H})\}$. Here, the index m refers to a particular measurement outcome. The probabilty of getting the outcome associated to m is

$$p(m) = \text{Tr}(M_m^\dagger M_m \rho). \tag{1.7}$$

Here, the operator $M_m^\dagger M_m$ defines a positive semi-definite measure on the space of density matrices. After the measuring outcome m, the new state of the system is

$$\rho \to M_m \rho M_m^\dagger / \text{Tr}(M_m^\dagger M_m \rho). \tag{1.8}$$

The probabilities must be normalized, so that for any state ρ, we should have $\sum_m p(m) = \text{Tr}(\sum_m M_m^\dagger M_m \rho) = 1$. From this, we get the completeness relation of the measurement operators:

$$\sum_m M_m^\dagger M_m = \mathbb{I}. \tag{1.9}$$

In the case of *projective* measurements, the measurement operators M_m are orthogonal projectors. The idea is that any hermition operator M defines an observable and a set of measurement outcomes given by the spectral decomposition of M, $M = \sum_m m M_m$ where m are the eigenvalues of M and M_m is the projector onto corresponding eigenspace. Then, the probability of measuring outcome m is simply $\mathrm{Tr}(M_m \rho) = \mathrm{Tr}(M_m^\dagger M_m \rho) = p(m)$, in agreement with the above generalization. Similarly, one can easily check the remaining properties are satisfied [164].

As an example, consider an arbitrary, single qubit projective measurement defined by $M = \vec{v} \cdot \vec{\sigma} \equiv v_x X + v_y Y + v_z Z$, where the vector $\vec{v}$ is normalized such that $\vec{v} \cdot \vec{v} = 1$. The possible measurement outcomes are $m = \pm 1$ with projectors $M_\pm = (\mathbb{1} \pm \vec{v} \cdot \vec{\sigma})/2$. We can always write the single qubit state in terms of Pauli matrices as well, with $\rho = (\mathbb{1} + \vec{u} \cdot \vec{\sigma})/2$, where $|\vec{u}| = \sqrt{\vec{u} \cdot \vec{u}} \leq 1$ is required to ensure the state is positive semi-definite. Then, the probability of getting any particular outcome is

$$p(m = \pm 1) = \frac{1}{4} \mathrm{Tr}\left((\mathbb{1} \pm \vec{v} \cdot \vec{\sigma})(\mathbb{1} + \vec{u} \cdot \vec{\sigma})\right) = \frac{1}{2}(1 \pm \vec{u} \cdot \vec{v}). \tag{1.10}$$

This makes sense intuitively, as when $\vec{u} = \vec{v}$, the state and observable are aligned and we can only get outcome $m = 1$. On the other hand, when $\vec{u} = -\vec{v}$, we can only get outcome $m = -1$ because they are anti-aligned. The average state after measurement $\mathcal{N}(\rho)$ is given by a weighted sum based on probability of the different measurement outcomes. If we measured the state but didn't record the outcome, then this weighted average $\mathcal{N}(\rho)$ is our best description of the state of the system.

$$\mathcal{N}(\rho) = \sum_m M_m p(m) = \frac{1}{2}(\mathbb{1} + (\vec{u} \cdot \vec{v})\vec{v} \cdot \vec{\sigma}) \tag{1.11}$$

Thus, the expected state evolves by $\vec{u} \to \vec{u}' = (\vec{v} \cdot \vec{u})\vec{v} = (\vec{v}\vec{v}^T)\vec{u}$. By applying the above quantum channel repeatedly, we see that the act of measuring alone seemingly neither"collapses" the state

along the measurement axis, nor has the effect of purifying the quantum state. That is, repeatedly measuring a state won't purify the state, if we don't record the measurement outcomes. We can quantify this by studying the second Renyi entropy of the state ρ under this channel,

$$S^{(2)}(\rho) = -\log(\frac{1 + |\vec{u}|^2}{2}), \tag{1.12}$$

where $S^{(2)}(\rho) = -\log \text{Tr}(\rho^2)$. Using Eq. (1.11), the entanglement of the average state after measurement is

$$S^{(2)}(\mathcal{N}(\rho)) = -\log(\frac{1 + (\vec{u} \cdot \vec{v})^2}{2}), \tag{1.13}$$

whereas the entanglement of any individual state after measurement is of course zero. There are several unintuitive observations in order. First, the formula above implies that a pure state, when subjected to the averaging channel $\mathcal{N}$, can *increase* $S^{(2)}$. This is quite unintuitive, and speaks to a shortcoming of the entanglement entropies in quantifying entanglement in mixed states. In this example, a single qubit passed through a single qubit quantum channel that does not depend on any external degrees of freedom, so how could it suddenly share quantum information with something else? The confusion arises because the entanglement-entropies sometimes confuse entropy with entanglement. In fact, Eq. (1.13) is precisely the entropy one would get from just the distribution of measurement outcomes $p(m = \pm 1)$.

Nevertheless, upon $\mathcal{N}$, the *purity* of the state $\text{Tr}(\rho^2)$ will generically decrease i.e. become more mixed or impure. This is consistent with our intuition, since the state becomes a mixture of different pure state measurement outcomes. In fact, a totally pure state can become totally-mixed if the measurement axis is orthogonal to the polarization of the state, since in this case $\vec{u} \cdot \vec{v} = 0$ so $\text{Tr}(\mathcal{N}(\rho)^2) = 1/2$. Furthermore, we see that the purity is not generically a markovian quantity. We can see this by choosing $\vec{u}$ to be a unit vector parallel to $\vec{v}$ in one case and orthogonal to $\vec{v}$ in another: in the first case, applying $\mathcal{N}$ doesn't affect the purity so both input and output states remain pure; in the second case, the state starts off pure and then becomes totally mixed.

In the next section, we will see how we can overcome the non-markovianness of the purity by introducing random local unitaries.

1.3   A random measure on the unitary group

Employing randomness can often be a useful trick for simplifying physics. This is obvious to anyone who has studied statistical mechanics, where we are often able to ignore specific details of a system in favor of statistical properties of ensembles. This is also especially useful when a particular system one wants to study is "typical" in that its properties may be calculated by instead averaging over an ensemble to which it belongs. Wigner applied this reasoning to study the energy levels of heavy atomic nuclei, which have complicated Hamiltonians of very large dimension. By assuming the Hamiltonians were typical for a random, Gaussian-distributed, hermitian matrix ensemble, he was able to determine the spectral properties of a wide range of atomic nuclei with a high degree of precision [222]. It has been conjectured that Hamiltonians whose classical counterpart is chaotic are typical e.g. their eigenvalue spectrum resembles that of a random Gaussian Unitary Ensemble (GUE) matrix [23].

The previous appraoch, highlighted by Wigner, uses "typicality" to replace a detailed object, e.g. a specific Hamiltonian, by a random sample. Then to study the properties of the detailed object, one argues that properties of it are close to the average properties of the random ensemble. Rather than starting with a specific state and assuming typicality, one could be interested in the properties of an ensemble that are independent of some irrelevant details. For example, suppose we want to study properties of an ensemble of states that is *independent* of local basis information. One way to accomplish that is by averaging over local bases–if we've done that properly, then we've have eliminated any local basis dependence.

To average over local basis information, we need a measure on the group of unitary matrices, $U(d)$. We will use the *Haar* measure, whose defining property is that it is left (or right) multiplication invariant i.e. $dU = d(UV) = d(VU)$. Using this definition, it is possible to get

a handle on how to perform averages over the unitary group. If we can calculate the moments of the Haar unitary ensemble, we should be able to average over any group function, so let us restrict ourselves to evaluating integrals of the form

$$\mathbb{E}_k[X] := \int dU\, U^{\otimes k} X (U^\dagger)^k, \tag{1.14}$$

where $X \in \mathscr{L}((\mathbb{C}^d)^{\otimes k})$. Since $\mathbb{E}_k[X]$ commutes with $V^{\otimes k}$ for $V \in U(d)$, as a consequence of Schur Weyl duality, it should be possible to expand $\mathbb{E}_k[X]$ in terms of permutation group elements $\sigma \in S_k$. We will use this fact in the derivation of the Weingarten functions. Let us determine the form of the map $\mathbb{E}_k$ for each k:

- **k $= 1$**. Using left-multiplication symmetry, we see that $\mathbb{E}_1[X]$ commutes with every unitary,

$$V\mathbb{E}_1[X]V^\dagger = \int dU\, (VU)X(VU)^\dagger = \int dU\, UXU^\dagger = \mathbb{E}_1[X], \tag{1.15}$$

  therefore using Schur's lemma must be proportional to the identity $\mathbb{E}_1[X] = c\mathbb{1}$. Since $\mathrm{Tr}\,\mathbb{E}_1[X] = \mathrm{Tr}\,X$, it follows that

$$\mathbb{E}_1[X] = \mathrm{Tr}(X)/d. \tag{1.16}$$

- **k > 1**. We can expand the RHS in the permutation basis elements $\tau \in S_k$,

$$\mathbb{E}_k[X] = \sum_\tau c(k,d,X)_\tau \tau, \tag{1.17}$$

  where the coefficients $c(k,d)_\tau$ can in principle depend on k, d and X. If we now take the trace of both sides with $\sigma^{-1} \in S_k$, we get that

$$\mathrm{Tr}(\sigma^{-1}X) = \sum_\tau c(k,d,X)_\tau \mathrm{Tr}(\sigma^{-1}\tau) := \sum_\tau c(k,d,X)_\tau W_{\mathbb{1}}[\sigma,\tau]. \tag{1.18}$$

The matrix $W_{\mathbb{1}}[\sigma, \tau] \in \mathbb{R}^{k! \times k!}$ counts the number of cycles in $\sigma^{-1}\tau \in S_k$ i.e. $W_{\mathbb{1}}[\sigma, \tau] = d^{\text{num. of cycles in } \sigma^{-1}\tau}$. This matrix is only invertible for $d \geq k$, but nevertheless its Moore-Penrose pseudo-inverse $W_{\mathbb{1}}^{-1}$ is always defined. The matrix elements $W_{\mathbb{1}}^{-1}[\sigma, \tau]$ of the inverse are known as the Weingarten coefficients $Wg(\sigma^{-1}\tau)$ and are given in terms of d and k, and $\sigma, \tau \in S_k$. The exact coefficients are given by [51]

$$Wg(\sigma) = \frac{1}{(k!)^2} \sum_{\lambda} \frac{\chi^{\lambda}(1)\chi^{\lambda}(\sigma)}{s_{\lambda,d}(1)}, \qquad (1.19)$$

where λ are partitions of k, χ^{λ} is the corresponding character of S_k, and $s_{\lambda,d}$ is the Schur polynomial of λ. The exact formula will not be of much use to us, as we will instead take the route of deriving $Wg(\sigma)$ by taking the pseudo-inverse of $W_{\mathbb{1}}$, which is much less complicated analytically for small k. Putting it all together, we get

$$\mathbb{E}_k[X] := \sum_{\sigma, \tau} \text{Tr}(\sigma^{-1}X)Wg(\sigma^{-1}\tau)\tau. \qquad (1.20)$$

The table of the exact, analytic Weingaren coefficients for $d \geq k$ can be found up to $k = 6$ in Tab. 1.2.

- **k = 2**. As an example, let us work out the case $k = 2$. There are only two elements in S_2: the identity permutation $e = (1)(2)$ and the swap permutation $s = (12)$. The cycle matrix $W_{\mathbb{1}}$ is then a two by two matrix with the following elements:

$$W_{\mathbb{1}} = \begin{pmatrix} d^2 & d \\ d & d^2 \end{pmatrix} = d^2 \mathbb{1} + dX \implies W_{\mathbb{1}}^{-1} = \frac{1}{d^2-1}\begin{pmatrix} 1 & -\frac{1}{d} \\ -\frac{1}{d} & 1 \end{pmatrix} = \frac{\mathbb{1} - X/d}{d^2-1} \quad (1.21)$$

Hence the Weingarten function evaluates to $Wg(e) = \frac{1}{d^2-1}$ and $Wg(s) = \frac{-1/d}{d^2-1}$. Plugging

this into Eq. (1.20), we get

$$\int dU U^{\otimes 2} X (U^\dagger)^2 = \frac{e(\text{Tr}(X) - \text{Tr}(Xs)/d) + s(\text{Tr}(Xs) - \text{Tr}(X)/d)}{d^2 - 1}. \qquad (1.22)$$

This expression can also be re-written in any complete operator basis $\mathscr{O} = \{O_0, \ldots O_{d^2-1}\}$. Suppose that the basis $\mathscr{O}$ is hermitian and orthonormal under the Hilbert-Schmidt norm i.e. $\text{Tr}(O_i O_j) = d^2 \delta_{ij}$. Furthermore, suppose that only one basis element has non-zero trace so that $\text{Tr}\, O_i = 0$ for $i > 0$. This is equivalent to saying $O_0 = \mathbb{1}$. For $d = 2$, we see that the (sign-normalized) Pauli group elements satisfy these conditions as a orthonormal basis on $d = 2^n$ dimensional Hilbert spaces. If we now view Eq. (1.22) as a superoperator, we can expand it in the operator basis $\mathscr{O}$ and analyze its components. It's clear that if $\mathbb{E}_2[O_i \otimes O_j] \propto \delta_{ij}$, since the traces eliminate any pair of basis elements which don't match. By the same logic, $\text{Tr}(O_i \otimes O_j \mathbb{E}_2[X]) \propto \delta_{ij}$ as well. Therefore, there are only $(d^2)^2$ expansion coefficients for $\mathbb{E}_2$,

$$t^{(2)}_{O_i, O_j} := \text{Tr}(O_i^{\otimes 2} \mathbb{E}_2[O_j^{\otimes 2}])/d^2 = \delta_{i,0} \delta_{j,0} + \frac{(1 - \delta_{i,0})(1 - \delta_{j,0})}{d^2 - 1}. \qquad (1.23)$$

Eq. (1.23) says that under a Haar unitary scrambling channel, any non-trivial basis operator is equally scrambled amongst all the others. In a qubit tensor product space, $d = 2^l$ where l is the locality of the unitary gate. In the case where we have $l = 2$ qubits, the above equation says that for any non-trivial Pauli string e.g. $P = XI, ZY, ZZ$, etc.,

$$\int dU U^{\otimes 2} P^{\otimes 2} (U^\dagger)^2 = \frac{1}{4^l - 1} \sum_{P' \neq \mathbb{1}\mathbb{1}} (P')^{\otimes 2}. \qquad (1.24)$$

Equipped with this new knowledge, let us return to the single-qubit measurement example in Sec. 1.3. Suppose we modify Eq. (1.11) so that instead of measuring along a fixed axis, we measure along a random axis on the unit 3-sphere. We wish to understand how the channel

changes on average, and how the its second Renyi entropy is modified. We can choose as our random, projective measurement operator $M_V = VZV^\dagger$ where $V \in U(2)$ is a single qubit unitary chosen randomly according to the Haar measure. When conjugated with this rotation, this Pauli Z transforms as $SO(3)$ into a different Pauli vector $\vec{v} \cdot \vec{\sigma}$. Here, the choice of Z is arbitrary and could be replaced with any other Pauli vector without loss of generality. The eigenvalues of $M_V = VZV^\dagger$ are still $m = \pm 1$ and the orthogonal projectors are

$$M_\pm = (\mathbb{1} \pm M_V)/2 = \begin{cases} V|0\rangle\langle 0|V^\dagger & m = 1 \\ V|1\rangle\langle 1|V^\dagger & m = -1 \end{cases}. \tag{1.25}$$

Hence, the averaging channel as a function of V becomes

$$\mathcal{N}_V(\rho) = \sum_{b=0,1} V|b\rangle\langle b|V^\dagger \operatorname{Tr}(\rho V|b\rangle\langle b|V^\dagger). \tag{1.26}$$

Let's also denote the average as the *measurement channel*:

$$\mathcal{M}(\rho) = \sum_{b=0,1} \int dV \, V|b\rangle\langle b|V^\dagger \operatorname{Tr}(\rho V|b\rangle\langle b|V^\dagger). \tag{1.27}$$

The effect of $\mathcal{N}_V$ on average is to impurify the state i.e. reduce its purity, while not affecting its direction of polarization. To see this, we can evaluate the components of this map in the Pauli basis $\frac{1}{2}\operatorname{Tr}(P\mathcal{M}(P'))$ using Eq. (1.24). We see that it is diagonal in this basis, and its eigenvalues only depend on the support of its input: $\delta_{P,P'}(t_{P,\mathbb{1}}^{(2)} + t_{P,Z}^{(2)})$. Plugging in the result from Eq. (1.23),

$$\mathcal{M}(P) = \begin{cases} P & P = \mathbb{1} \\ P/3 & P = X,Y,Z \end{cases}. \tag{1.28}$$

Therefore, for a given state ρ, under $\mathcal{M}$, its polarization vector transforms as $\vec{u} \to \vec{u}/3$. Hence $\mathcal{M}$, unlike the previous one for fixed $\vec{u}$, is actually invertible. The second Renyi entropy also

18

evolves more predictably, so that states with the same initially entropy have matching entropies after the application of $\mathscr{M}$:

$$S^{(2)}(\mathscr{M}(\rho)) = -\log\left(\frac{1 + |\vec{u}|^2/9}{2}\right). \tag{1.29}$$

With out new formulas to compute unitary averages, let us now calculate the average purity $W_{\mathscr{E}}(A)$ of different types of random state ensembles $\mathscr{E}$ on n qubits. Given a random state ensemble, we can define the *entanglement feature*, or average purity, of this ensemble in terms of its second moment and the appropriate swap boundary conditions,

$$W_{\mathscr{E}}(A) := \mathbb{E}_{\sigma \in \mathscr{E}} \operatorname{Tr}(\sigma_A^2) = \mathbb{E}_{\sigma \in \mathscr{E}} \operatorname{Tr}(\sigma^{\otimes 2} \mathbb{X}_A), \tag{1.30}$$

where the operator $\mathbb{X}_A := \prod_{i \in A} (\mathrm{SWAP})_i \prod_{i \notin A} \mathbb{1}^{\otimes 2}$ acts on the doubled Hilbert space as a swap operator between the replicas for $i \in A$ and as the identity for $i \notin A$. Let us start with a random product state, which we can define in terms of a random pure state ensemble $\mathscr{E}_0$, consisting of states $\sigma = \otimes_{i=1}^n \sigma_i$, $\sigma_i = U_i |0\rangle\langle 0| U_i^\dagger$, where the probability of sampling σ is defined by the Haar measure on each individual Hilbert space $p(\sigma) = \prod_i dU_i$. Let $W_{\mathscr{E}_0}(A)$ denote the average purity of this ensemble over a subregion A. It's clear that since this state is a product state on every site, $\operatorname{Tr}(\sigma_i^2) = \operatorname{Tr}(\sigma_i) = 1$, so this state ensemble has trivial entanglement feature,

$$W_{\mathscr{E}_0}(A) = 1 \quad \forall A \subset \Omega. \tag{1.31}$$

Alternatively, we can consider a different random state ensemble $\mathscr{E}_\infty$ on n qudits formed by acting with a large unitary $U \in U(d^n)$ on a fixed initial state according to the Haar measure. We can think of such a state as the infinite time limit of a random brick-wall local unitary quantum circuit, and so the entanglement features of this state will appear as a fixed point when we study

entanglement dynamics in random systems. Using Eq. (1.22),

$$W_{\mathscr{E}_\infty}(A) = \int dU \, \mathrm{Tr}\left(\mathbb{X}_A U^{\otimes 2} \sigma_0^{\otimes 2} (U^\dagger)^{\otimes 2}\right) = \frac{d^n}{d^n + 1}(d^{-|A|} + d^{-|\bar{A}|}). \tag{1.32}$$

The volume law coefficient η of this state characterizes how the entropy of a subsystem grows with system size. As we can see, the volume law coefficient of this state is maximal since a single qubit cannot contribute more than $\log d$ in entropy.

$$\eta := -\lim_{|A| \to \infty} \frac{\partial_{|A|} \log W_{\mathscr{E}_\infty}(A)}{\log d} \leq 1 \tag{1.33}$$

With this definition, it is clear that $\eta = 1$ for a random Haar state. A typical Haar random state in the whole Hilbert space has volume law entanglement, and so we see that area law entangled states more rare when considering the whole Hilbert space. The fact that area-law states are so ubiquitous in the low-energy physics of many-body systems is a blessing in disguise, as it opens up the possibility of efficient simulation.

Now that we've gotten some practice with the basic themes of this dissertation, we will outline the proceeding chapters so the reader can have a roadmap of what to expect. In Chapter 2, we lay out a formalism for understanding purity dynamics in quantum many body systems. The idea is that we can map dynamics of the average purity of an ensemble of states onto the dynamics of Ising spin systems, where each Ising spin configuration corresponds to a particular subregion or subset of lattice sites. The state ensemble maps onto its corresponding entanglement-feature state, whose components are the bipartite purities of the state ensemble. In the case where the ensemble is local basis independent, the average purity dynamics are markovian. Therefore, we can then commute the evolution of the state ensemble under a random channel ensemble to the application of a transfer matrix on the entanglement feature state. The transfer matrix may be used to define the entanglement feature Hamiltonian, whose spectral

properties determine many aspects of how quantum information propogates in the system. We conclude the chapter with some examples of different types of locally-scrambled evolution and their entanglement dynamics.

Next, in Chapter 3, we explore the entanglement structure of the entanglement feature states themselves. We find that their sign structure suggests that the states themselves have little entanglement, and therefore should have efficient matrix-product state representation. We propose a simple, bond-dimension two, two parameter ansatz that characterizes the typical states that one finds throughout the phase, e.g., product states, area-law states, volume law states, and their intermediaries. We study the entanglement "spectrum" and relate the the butterfly velocity of the system to its entanglement "gap" and compare our formula to previously studied results.

In Chapter 4, we shift gears slightly and explain the core principles of classical shadow tomography [100, 102], a general prediction scheme for quantum states that requires few measurements. As we'll see, the measurement channel of classical shadow tomography, which defines the prediction scheme, only depends on the average purity dynamics of the random snapshot state ensemble. This fact will be used to decompose the measurement channel as an matrix product operator in the shallow circuit limit, and furthermore we'll see how to exploit the low-entanglement of entanglement feature states explored in Chapter 3 to do efficient prediction of quasi-local operators.

In Chapter 5, we apply concepts from classical shadow tomography to study the measurement induced phase transition in a more practical settling. The measurement-induced phase transition in the second Renyi entropy was using the entanglement-feature formalism in [63], where the authors mapped the phase transition in the average purity to the Ising phase transition. In this paper, we consider hybrid circuits in a classical shadow setting to see how randomized, mid-circuit measurements affect the Pauli weight dynamics and ultimately the predictive properties of the underlying scheme. We study the phase transition in the purity by observing the base at which the shadow-norm, which characterizes the efficiency of the classical shadow scheme at predicting different types of operators, scales. Through simple arguments, we are able to argue

that the base undergoes a phase transition and is minimized (optimal) at the critical point.

Lastly, in Chapter 6, we summarize the key results of this dissertation and outline future directions of research.

Table 1.1. Properties of entanglement measures.

	Renyi Entropy	Von-Neumann Entropy	Negativity
Vanishes on pure states	X	X	X
Unitarily invariant	X	X	X
Continuous in ρ	X	X	
Additive on tensor product states $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$	X	X	X
Maximal on maximally-mixed state	X	X	
Convex function of ρ i.e. $S(\sum_i p_i \rho_i) \leq \sum_i p_i S(\rho_i)$	X	X	
Entanglement monotone i.e. $S(\mathcal{N}(\rho)) \leq S(\rho)$ for $\mathcal{N}$ any LOCC operation.			X
Ease of computation	Easy	Hard	Hard

Table 1.2. Exact, analytical formulas for the Weingarten functions up to $k = 6$.

Weingarten Functions	S_3	S_4	S_5
e	$\frac{d^2-2}{d(d^4-5d^2+4)}$	$\frac{d^4-8d^2+6}{d^2(d^2(d^2-7)^2-36)}$	$\frac{d^4-20d^2+78}{d^9-30d^7+273d^5-820d^3+576d}$
(12)	$-\frac{1}{d^4-5d^2+4}$	$-\frac{1}{d^5-10d^3+9d}$	$-\frac{d^4-14d^2+24}{d^2(d^8-30d^6+273d^4-820d^2+576)}$
$(12)(34)$	N.A.	$\frac{d^2+6}{d^2(d^2(d^2-7)^2-36)}$	$\frac{d^2-2}{d(d^8-30d^6+273d^4-820d^2+576)}$
(123)	$\frac{2}{d^5-5d^3+4d}$	$\frac{2d^2-3}{d^2(d^2(d^2-7)^2-36)}$	$\frac{2}{d(d^6-21d^4+84d^2-64)}$
$(123)(45)$	N.A.	N.A.	$-\frac{2(d^2+12)}{d^2(d^8-30d^6+273d^4-820d^2+576)}$
(1234)	N.A.	$-\frac{5}{d(d^2(d^2-7)^2-36)}$	$\frac{24-5d^2}{d^2(d^8-30d^6+273d^4-820d^2+576)}$
(12345)	N.A.	N.A.	$\frac{14}{d^9-30d^7+273d^5-820d^3+576d}$

Weingarten Functions	S_6
e	$\frac{d^8-41d^6+458d^4-1258d^2+240}{d^2(d^2-1)^2(d^8-54d^6+969d^4-6676d^2+14400)}$
(12)	$-\frac{d^4-24d^2+38}{d(d^2-1)^2(d^6-45d^4+564d^2-1600)}$
$(12)(34)$	$\frac{d^4-3d^2+10}{d^2(d^2-1)^2(d^6-38d^4+361d^2-900)}$
$(12)(34)(56)$	$-\frac{d^4+d^2+358}{d(d^2-1)^2(d^8-54d^6+969d^4-6676d^2+14400)}$
(123)	$\frac{2d^6-51d^4+229d^2-60}{d^2(d^2-1)^2(d^8-54d^6+969d^4-6676d^2+14400)}$
$(123)(45)$	$\frac{-2d^2-13}{d(d^2-1)^2(d^6-45d^4+564d^2-1600)}$
$(123)(456)$	$\frac{4(d^4+29d^2-90)}{d^2(d^2-1)^2(d^8-54d^6+969d^4-6676d^2+14400)}$
(1234)	$\frac{13-5d^2}{d(d^2-1)^2(d^6-38d^4+361d^2-900)}$
$(1234)(56)$	$\frac{5(d^4+15d^2+8)}{d^2(d^2-1)^2(d^8-54d^6+969d^4-6676d^2+14400)}$
(12345)	$\frac{14(d^2-10)}{d^2(d^{10}-55d^8+1023d^6-7645d^4+21076d^2-14400)}$
(123456)	$\frac{42}{-d^{11}+55d^9-1023d^7+7645d^5-21076d^3+14400d}$

Chapter 2

The Entanglement Feature Formalism under Locally-Scrambled Evolution

2.1 Introduction

Quantum entanglement dynamics[29, 94, 115, 119, 143] is an emerging field that ties several interesting topics together, including non-equilibrium and driven quantum systems[57, 150, 181], many-body localization and thermalization[4, 17, 120, 147, 162, 206], quantum chaos and holography[19, 71, 96, 149, 218]. The central theme is to understand the production and propagation of quantum entanglement in quantum many-body systems. For pure states, the amount of quantum entanglement between a subsystem A and its environment $\bar{A}$ can be quantified by the (Rényi) *entanglement entropy* (EE) $S^{(n)}(A) = \frac{1}{1-n} \log \mathrm{Tr}_A \rho_A^n$ where $\rho_A = \mathrm{Tr}_{\bar{A}} |\Psi\rangle\langle\Psi|$ is the reduced density matrix of subsystem A. Various quantum information measures (such as mutual and tripartite information) can be constructed from the EE over different regions. Here, we would like to focus on the 2nd Rényi entropies $S^{(2)}(A)$ and establish their dynamic equations under quantum evolution.

As a quantum state $|\Psi\rangle$ evolves in time, its EE's $S^{(2)}(A)$ over different regions A will also change with respect to time in general. It is desired to understand how the unitary evolution of the quantum state induces the dynamics of quantum entanglement. There have been several works on the entanglement growth in quantum many-body systems[39, 73, 94, 114, 117, 149, 153, 155, 158, 187, 189, 214, 237]. The main focus has been on the half-system (or a single

region) EE. To gain more resolution of the many-body entanglement structure, we extend our scope to all possible bipartitions of the system (including multiple disconnected entanglement regions). The question we would like to address is that given $S^{(2)}(A)$ at initial time over all possible subsystems A, what will be the equation of motion governing the evolution for all of them jointly in later time?

However, EE's over all regions contain a large amount of data, because the number of possible bipartitions 2^L grows exponentially in system size L. We need a conceptually concise way to organize these entropy data, in order to make progress in describing their dynamics. In Ref. [228], it was proposed that all these EE's can be organized into "entanglement features", which admit compact representations in terms of Boltzmann weights of Ising models. The key idea is to label each entanglement region A by a set of Ising variables $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \cdots)$, such that $\sigma_i = \downarrow$ (or $\uparrow$) corresponds to $i \in A$ (or $i \in \bar{A}$) for each site i. Then the EE $S^{(2)}(A) \equiv S^{(2)}[\boldsymbol{\sigma}]$ can be treated as a free energy associated to the Ising configuration $\boldsymbol{\sigma}$, and the *entanglement feature* (EF) refers to the corresponding Boltzmann weight $W[\boldsymbol{\sigma}] = e^{-S^{(2)}[\boldsymbol{\sigma}]} = \text{Tr} \rho_A^2$, which is simply the purity for the 2nd Rényi case. Its time evolution can be related to the Loschmidt echo on the duplicated system,[94] which could be of experimental relevance. In this work, we further develop the Ising formulation by encoding the EF as a fictitious spin state $|W\rangle = \sum_{\boldsymbol{\sigma}} W[\boldsymbol{\sigma}]|\boldsymbol{\sigma}\rangle$, which we called the EF state. This rewriting packs the exponentially many entanglement data into a single EF state (as a many-body wave function). This conceptual simplification enables us to formulate the entanglement dynamics in a concise form of imaginary-time Hamiltonian evolution of the EF state

$$\partial_t |W\rangle = -\hat{H}_{\text{EF}}|W\rangle, \tag{2.1}$$

which can be further analyzed using powerful tools that have been developed in quantum many-body physics. Our development is along the line of mapping entanglement dynamics to statistical mechanical problems, as discussed in a few recent works [93, 114, 148, 153, 155, 158, 207, 238]. Given the equivalence between statistical mechanics and imaginary-time quantum mechanics, it

is not surprising that the entanglement dynamics could admit a quantum mechanical formulation as Eq. (2.1).

Treating the EF $W[\boldsymbol{\sigma}]$ as an (unnormalized) probability distribution of entanglement regions $\boldsymbol{\sigma}$, the proposed dynamic equation in Eq. (2.1) could be interpreted as a Markov equation. The assumption behind this equation is that the future EF of a many-body state can be entirely determined based on the current EF without the need to know about the past EF or about other information beyond the EF. Unfortunately, this assumption does *not* hold in general! In fact, the entanglement dynamics is generally non-Markovian, meaning that knowing the present EE's even for all possible regions is still insufficient to determine their evolution in the future,[1] so we should not expect Eq. (2.1) to work in general. In this work, we point out a specific yet rich enough class of quantum dynamics, called the *locally scrambled* quantum dynamics, whose entanglement dynamics can be described by Eq. (2.1) (or some discrete version of it). Quantum dynamics can always be formulated as a unitary evolution $U = \prod_t U_t$ that can be chopped up into products of simpler unitaries U_t at each time slice t following a time ordering. A quantum dynamics is said to be locally scrambled, if for every time step, the unitary U_t is drawn from a random unitary ensemble that is invariant under local (on-site) basis transformations, and U_t at different time t are sampled independently. Such dynamics can be constructed by inserting local scramblers (product of on-site Haar random unitaries) between every time step, as if the system constantly forget about the choice of local basis from one time step to another. It can be used to model those quantum many-body systems with fast and random dynamics on each site, such that the quantum information is scrambled on each site quickly and sufficiently during each step of the time evolution. One famous example in this class is the Haar random unitary circuit[155, 158, 214, 238]. We will provide more examples of locally scrambled quantum dynamics in this work.

The reason that the future EE can be uniquely determined by the present EE under the

---

[1]One can easily construct examples like $|\uparrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$, which are both product states, but their evolution under the same Hamiltonian $H = \boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ will result in states of different entanglements.

locally scrambled quantum dynamics is related to the fact that the EE is a local-basis-independent quantity. As the local scramblers constantly remove the local-basis-dependent information in the quantum many-body state, only the local-basis-independent information can survive in time to govern the future evolution. Such local-basis-independent information can be captured by EE's over all possible entanglement regions, which are summarized as the EF of the quantum many-body state. In this work, we develop the theoretical framework to derive the dynamic equation governing the evolution of the EF under locally scrambled quantum dynamics. We establish a systematic approach to construct the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$ based on the entanglement properties of the physical Hamiltonian or unitary operators that describe the quantum dynamics. We also carry out numerical simulations to justify the assumptions made in the theoretical development, and demonstrate the prediction power of the EF approach.

The paper is organized as follows. In Sec. 2.2, we first develop the theoretical framework for the EF and its application to the locally scrambled quantum dynamics. We start with the definition of EF for both quantum many-body state and quantum unitary circuits in Sec. 2.2.1. We then promote these notions to their quantum mechanical versions, introducing the EF states and EF operators in Sec. 2.2.2. With this setup, in Sec. 2.2.3, we prove an important relation between the the state and the unitary EF's, thereby mapping the unitary evolution of the quantum state to the dissipative evolution of the EF state under the corresponding entanglement dynamics in Sec. 2.2.4. Taking the continuum limit, we obtain the Schrödinger equation for EF state and derived the most generic form of the EF Hamiltonian in Sec. 2.2.5. We analyze the spectral properties of the EF Hamiltonian and their consequences on the universal behavior of entanglement dynamics in Sec. 2.2.6. We investigate the excitation spectrum of the EF Hamiltonian and obtain the quasiparticle dispersion in Sec. 2.2.7, which allows us to predict the long-time saturation behavior of the entanglement. We will provide numerical evidences in Sec. 2.3 to demonstrate the validity of the EF approach. We first introduce two models of locally scrambled quantum dynamics in Sec. 2.3.1, namely the locally scrambled quantum circuit and the locally scrambled Hamiltonian dynamics, which are further discussed in details in Sec. 2.3.2 and

28

Sec. 2.3.3 separately. We sum up in Sec. 2.4 making connections to related topics and potential future development.

2.2 Theoretical Framework

2.2.1 Definition of Entanglement Features

Let us consider a quantum many-body system consisting of L qudits, where each qudit (d-dimensional quantum system) has a d-dimensional physical Hilbert space, such that the total Hilbert space dimension is d^L. To define the 2nd Rényi entropy, we will need to duplicate the system and evaluate the expectation value of swap operators within a subsystem A of interest. There are altogether 2^L possible choices of a subsystem A, as each qudit can independent decide to be included in A or not. To label the 2^L different bipartitions of the system, we introduce a set of classical Ising variables $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \cdots, \sigma_L)$, such that the Ising variable σ_i determines if the ith qudit belongs to region A or its complement $\bar{A}$, following

$$\sigma_i = \begin{cases} \uparrow & i \in \bar{A}, \\ \downarrow & i \in A. \end{cases} \tag{2.2}$$

These Ising variables do not correspond to any degrees of freedom of the underlying quantum many-body system. Instead, they represent the identity or swap operator supported on the duplicated system, which are used to define the 2nd Rényi entropy. To be more specific, we define a permutation operator $\mathscr{X}_{\sigma_i}$ acting on the duplicated Hilbert space of the ith qudit,

$$\mathscr{X}_{\sigma_i} = \begin{cases} \|_i \equiv \sum_{\alpha,\beta=1}^d |\alpha\beta\rangle_i \langle\alpha\beta|_i & \text{if } \sigma_i = \uparrow, \\ \chi_i \equiv \sum_{\alpha,\beta=1}^d |\alpha\beta\rangle_i \langle\beta\alpha|_i & \text{if } \sigma_i = \downarrow, \end{cases} \tag{2.3}$$

which is assigned to the identity operator $\|_i$ or the swap operator χ_i depending on the Ising variable σ_i. Assembling these permutation operators together, we define $\mathscr{X}_{\boldsymbol{\sigma}} = \bigotimes_{i=1}^L \mathscr{X}_{\sigma_i}$ for the duplicated L-qudit system, which implements swap operations in the region A specified by

29

the Ising configuration $\boldsymbol{\sigma}$.
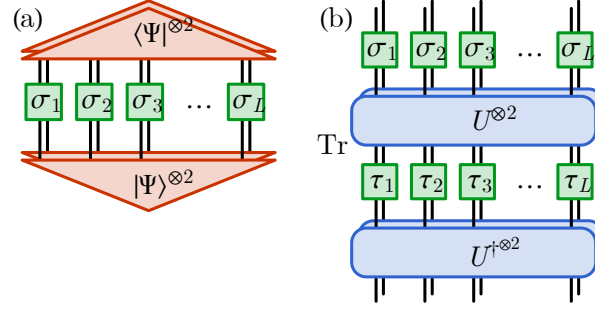
Figure 2.1. Diagrammatic representation of (a) the state EF $W_{|\Psi\rangle}[\boldsymbol{\sigma}]$ and (b) the unitary EF $W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}]$. The Tr operator contracts the dangling bottom legs with the corresponding dangling top legs.

With these notation setup, we can define the *entanglement feature* (EF) of quantum many-body states and time-evolution unitary circuits[228, 230]. The EF of a many-body pure state $|\Psi\rangle$ is defined as

$$W_{|\Psi\rangle}[\boldsymbol{\sigma}] \equiv e^{-S^{(2)}[\boldsymbol{\sigma}]} = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}}(|\Psi\rangle\langle\Psi|)^{\otimes 2} \right), \tag{2.4}$$

which resembles Boltzmann weights for Ising configurations $\boldsymbol{\sigma}$ labeling different entanglement regions. In terms of the tensor network representation, the state EF can be depicted as Fig. 2.1(a). Not only for quantum states, the EF can also be defined for unitary circuits under the state-operator correspondence [96, 124, 163]. The EF of a unitary circuit U is defined as

$$W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} U^{\otimes 2} \mathscr{X}_{\boldsymbol{\tau}} U^{\dagger \otimes 2} \right), \tag{2.5}$$

which depends on two sets of Ising configurations $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ that separately specifies the entanglement regions on the past (input) and the future (output) sides of the unitary circuit, as illustrated in Fig. 2.1(b). The state EF $W_{|\Psi\rangle}[\boldsymbol{\sigma}]$ provides a comprehensive description of the entanglement properties of the pure state $|\Psi\rangle$, which contains the information about EE, mutual information

and multipartite information among different subsystems. Similarly, the unitary EF $W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ characterizes the entanglement properties of the unitary circuit U, including the EE and mutual information between past and future degrees of freedoms, which are also closely related to the operator-averaged out-of-time ordered correlator (OTOC)[64, 96, 133] under the quantum dynamics U.

It worth mention that entanglement features are invariant under *local basis transformations*. A generic local basis transformation takes the form of $V = \bigotimes_{i=1}^{L} V_i$ with V_i being a unitary operator acting on the ith qudit. It is easy to see that both the state EF and the unitary EF are independent of the choice of local basis, i.e.

$$W_{V|\Psi\rangle} = W_{|\Psi\rangle}, \quad W_{V^\dagger U V} = W_U. \tag{2.6}$$

In this way, the EF forgets about the local basis dependent information in quantum states or unitary circuits, and only captures the entanglement properties that are universal to local basis choices.

2.2.2 Operator Formalism of Entanglement Features

To make our notation more concise, let us introduce a set of Ising basis $|\boldsymbol{\sigma}\rangle$, then we can pack $W_{|\Psi\rangle}$ to an *entanglement feature state* (EF state) $|W_\Psi\rangle$ as

$$|W_\Psi\rangle = \sum_{\boldsymbol{\sigma}} W_{|\Psi\rangle}[\boldsymbol{\sigma}]|\boldsymbol{\sigma}\rangle, \tag{2.7}$$

and W_U to an *entanglement feature operator* (EF operator) $\hat{W}_U$ as

$$\hat{W}_U = \sum_{\boldsymbol{\sigma}, \boldsymbol{\tau}} |\boldsymbol{\sigma}\rangle W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}]\langle\boldsymbol{\tau}|. \tag{2.8}$$

The Ising basis $|\boldsymbol{\sigma}\rangle$ span a 2^L-dimensional Hilbert space of L qubits, called the *entanglement feature Hilbert space* (EF Hilbert space). It should not be confused with the d^L-dimensional

physical Hilbert space of the underlying quantum many-body system. Each Ising basis state $|\boldsymbol{\sigma}\rangle$ in the EF Hilbert space simply corresponds to a bipartition of the L physical qudits following Eq. (2.2).

Given the EF state $|W_\Psi\rangle$, the EE $S^{(2)}[\boldsymbol{\sigma}]$ over all regions can be retrieved from the inner product of $|W_\Psi\rangle$ with the corresponding Ising basis state

$$e^{-S^{(2)}[\boldsymbol{\sigma}]} = W_{|\Psi\rangle}[\boldsymbol{\sigma}] = \langle \boldsymbol{\sigma} | W_\Psi \rangle. \tag{2.9}$$

In particular, a *product state* $|\Psi_{\text{prod}}\rangle = \bigotimes_i |\psi_i\rangle$ has zero EE in any region ($\forall \boldsymbol{\sigma} : S^{(2)}[\boldsymbol{\sigma}] = 0$), so its EF state is therefore a equal weight superposition of all Ising configurations,

$$|W_{\text{prod}}\rangle = \sum_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle \quad \text{(product state)}, \tag{2.10}$$

which corresponds to the (ideal) paramagnetic state of Ising spins. On the other hand, a *Page state*[172] $|\Psi_{\text{Page}}\rangle$ exhibits the maximal volume-law EE, whose EF state is given by

$$|W_{\text{Page}}\rangle = \sum_{\boldsymbol{\sigma}} \frac{\cosh(\eta \sum_{i=1}^{L} \sigma_i)}{\cosh(\eta L)} |\boldsymbol{\sigma}\rangle \quad \text{(Page state)}, \tag{2.11}$$

where $\eta = \frac{1}{2}\log d$ and we have adopted $\sigma_i = \pm 1$ in the formula to represent $\uparrow, \downarrow$ spins. This result follows from the definition. Its detailed derivation can be found in Appendix A.1. The state $|W_{\text{Page}}\rangle$ contains extensive ferromagnetic correlations among Ising spins. In this picture, the process of quantum state thermalization corresponds to the process of building up ferromagnetic correlations in the EF state (until saturation to the Page state).

Let us also provide some examples for the EF of unitary gates which will be useful later. The EF of a single-qudit identity operator is straight forward to calculate based on the definition

in Eq. (2.5),

$$\hat{W}_{\mathbb{1}} = d^2(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) + d(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|),$$
$$= d(d + X),$$
(2.12)

where X denotes the Pauli-X operator acting on the qudit site (acting in the EF Hilbert space, not in the qudit Hilbert space), and d is the qudit dimension. A more non-trivial example is the EF of a two-qudit Haar random unitary gate U_{ij} (averaged over Haar ensemble) that acts on qudits i and j,

$$\hat{W}_{\text{Haar}} = d^2(d + X_i)(d + X_j) - \frac{d^2(d^2 - 1)}{d^2 + 1}\frac{1 - Z_i Z_j}{2}(d^2 - X_i X_j),$$
(2.13)

where X_i and Z_i are Pauli-X and Z operators acting on site i. The derivation can be found in Appendix A.2.
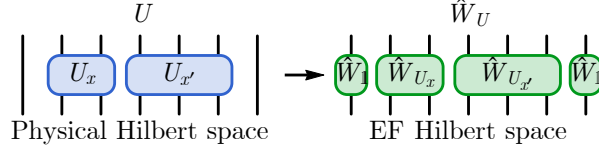
Figure 2.2. The mapping from the unitary operator in the physical Hilbert space to the corresponding EF operator in the EF Hilbert space. Locality is preserved under the mapping, enabling us to factorize the operators in the same manner on both sides.

Unitary gates are the building blocks to construct more complicated unitary circuits. One nice property of the EF operator is that it preserves the locality in space, meaning that if a unitary U operator can be factorized to smaller unitaries U_x over the space x, its corresponding EF operator $\hat{W}_U$ is also factorized in the same manner

$$U = \bigotimes_x U_x \quad \Rightarrow \quad \hat{W}_U = \bigotimes_x \hat{W}_{U_x},$$
(2.14)

as exemplified in Fig. 2.2. This property allows us to assemble the local EF operators together.

For example, the EF operator $\hat{W}_{\mathbb{1}}$ of the identity operator for a L qudit system be obtained by assembling the single-qudit result in Eq. (2.12) together

$$\hat{W}_{\mathbb{1}} = \prod_{i=1}^{L} d(d + X_i) = (\coth \delta \operatorname{csch} \delta)^L \prod_{i=1}^{L} e^{\delta X_i}, \qquad (2.15)$$

where we have introduced the constant δ

$$\delta \equiv \operatorname{arccoth} d = \frac{1}{2} \log \frac{d+1}{d-1} \qquad (2.16)$$

to exponentiate the operator. The exponential form allows us to take the operator inverse easily, such that

$$\hat{W}_{\mathbb{1}}^{-1} = \prod_{i=1}^{L} \frac{1 - d^{-1} X_i}{d^2 - 1} = (\tanh \delta \sinh \delta)^L \prod_{i=1}^{L} e^{-\delta X_i}. \qquad (2.17)$$

These results will be useful in later discussions. In the following, we will show how the evolution of the EF state can be inferred from the EF operator of the unitary circuit.

2.2.3 Relation between State and Unitary Entanglement Features

Suppose U describes a unitary circuit that evolves an initial quantum many-body state $|\Psi\rangle$ to the final state $U|\Psi\rangle$. This quantum dynamics will induce a corresponding entanglement dynamics, under which the EF of the initial state $W_{|\Psi\rangle}$ evolves to that of the final state $W_{U|\Psi\rangle}$. Can we predict the final state EF $W_{U|\Psi\rangle}$ based on our knowledge about the initial state EF $W_{|\Psi\rangle}$ and the EF W_U of the unitary evolution?

In general, this problem is not tractable. Because U and $|\Psi\rangle$ contain many "non-universal" features that are specific to the choice of local basis, such features may affect the final state entanglement, but they are not captured by the EF, as the EF is invariant under local basis transformations. Therefore, the final state EF can not be inferred from the initial state EF and the unitary EF in general. However, instead of dealing with a specific unitary circuit U, we consider an ensemble of unitary circuits $U' = V^\dagger U V$ related to U by local basis transformations

V, denoted by

$$\mathscr{E}_U = \left\{ V^\dagger U V \middle| V = \bigotimes_{i=1}^{L} V_i, V_i \in \text{Haar} \right\}, \tag{2.18}$$

where each V_i is independently drawn from the Haar random unitary ensemble defined on the ith qudit. We will call $\mathscr{E}_U$ the *locally scrambled* unitary ensemble associated with U. According to Eq. (2.6), one immediately see that all unitary operators $U' \in \mathscr{E}_U$ in the ensemble share the *same* entanglement feature as that of U, i.e. $W_{U'} = W_U$. Rather than asking about the EF of a specific final state $U|\Psi\rangle$, if we are allowed to consider the ensemble average of the EF over all final states $U'|\Psi\rangle$ with $U' \in \mathscr{E}_U$, the final state EF $W_{U'|\Psi\rangle}$ will indeed be constructable from the initial state EF $W_{|\Psi\rangle}$ and the unitary EF $W_{U'} = W_U$ on the average level. Using the operator formalism, the relation can be written in a concise form as

$$\mathop{\mathbb{E}}_{U' \in \mathscr{E}_U} |W_{U'\Psi}\rangle = \hat{W}_U \hat{W}_{\mathbb{1}}^{-1} |W_\Psi\rangle, \tag{2.19}$$

where $\hat{W}_{\mathbb{1}}$ is the EF operator for the identity evolution $\mathbb{1}$ and $\hat{W}_{\mathbb{1}}^{-1}$ is its inverse, which was given in Eq. (2.17) explicitly. One can derive Eq. (2.19) using tensor network diagrams, see Appendix A.3 for details. To simplify the notation, we may suppress spelling out the ensemble average $\mathbb{E}_{U' \in \mathscr{E}_U}$ explicitly in later discussions, with the understanding that in this work any unitary operator appeared in the subscript of the EF operator will be implicitly averaged over local basis transformations. Eq. (2.19) establishes an important relation between the state and the unitary EF's, which enables us to compute the evolution of the state EF induced by the underlying quantum dynamics, given the EF of the corresponding unitary evolution U. A special case of Eq. (2.19) has been discussed in Ref. [133, 228], where the initial state is restricted to product states.

As a side remark, we would like to provide some justifications for the use of locally scrambled unitary ensembles $\mathscr{E}_U$. Technically speaking, working with these ensembles enables us to predict the future evolution of EE's purely based on their current data, because the local-

basis-dependent features of a quantum state are removed by local scrambling and the remaining local-basis-independent features are captured by the EF[2]. This setup allows us to make progress in understanding the entanglement dynamics with a tractable theoretical limit. Physically speaking, we can imagine systems with separating time scales between the on-site and the inter-site quantum dynamics. Suppose the on-site dynamics is fast and random, then the quantum information would be sufficiently scrambled on every site, before it can spread out to other sites at a longer time scale. So the overall unitary evolution will constantly be interrupted by the insertion of local scramblers $V_i \in$ Haar, making the evolution effectively local-basis-independent. In fact, many well explored random unitary ensembles in the field of entanglement dynamics are local-basis-independent (or "locally scrambled" in our language), including random unitary dynamics[105, 155, 169], random Hamiltonian dynamics[145, 191, 213, 228] and random Floquet dynamics[34, 36]. This strategy has also been adopted in the discussion of operator dynamics[73, 114, 117, 158, 214] and random tensor networks[93, 184, 207, 230]. Historically, the study of these models has advanced our understanding about the universal behavior of entanglement dynamics, so we would like to carry on this line of research.

2.2.4 Markovian Entanglement Dynamics

As long as we know how to construct the EF operator $\hat{W}_U$ for any unitary evolution U of interest, we can apply the operator formalism in Eq. (2.19) to compute the entanglement dynamics. However, calculating the EF for a large and deep unitary circuit is a difficult many-body problem, hence the relation Eq. (2.19) is still hard to apply. But if all unitary gates in the unitary circuit are independently drawn from locally scrambled unitary ensembles, they will be decoupled in time, such that we can apply the EF operator iteratively to drive the evolution of the EF state.

To be more concrete, let us consider the case where the full unitary evolution can be

---

[2]Strictly speaking, all the local-basis-independent features are capture by the full set of Rényi entropy to all Rényi orders. But here we only focus on a subset described by the 2nd Rényi entropy.
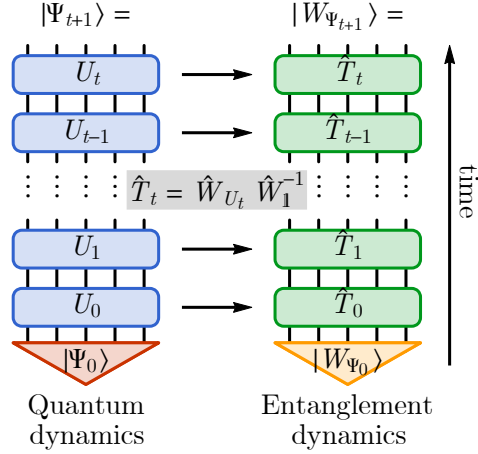
Figure 2.3. Quantum dynamics induces entanglement dynamics, assuming each unitary U_t is drawn from local basis invariant ensemble independently. The operator entanglement property of U_t determines the transfer matrix $\hat{T}_t$ that evolves the EF state via Eq. (2.25), and the EF state $|W_{\Psi_t}\rangle$ encodes the entanglement properties of the quantum state $|\Psi_t\rangle$.

broken up into discrete time steps (or layers), and each single-step unitary evolution at time t is described by U_t, as illustrated on the left of Fig. 2.3. Then the quantum many-body state $|\Psi_t\rangle$ evolves from step to step following

$$|\Psi_{t+1}\rangle = U_t|\Psi_t\rangle. \tag{2.20}$$

Suppose U_t at different time t are independently drawn from random unitary ensembles (not necessary Haar random) which are invariant under local basis transformation, the full unitary evolution

$$U = \prod_t U_t = U_t U_{t-1} \cdots U_1 U_0 \tag{2.21}$$

will form a random unitary circuit that defines a *locally scrambled quantum dynamics*. If we spell out the local basis transformations V_t that has been made at each time step, i.e. $U_t = V_t^\dagger U_t' V_t$,

$$U = V_t^\dagger\, U_t'\, V_t V_{t-1}^\dagger\, U_{t-1}'\, V_{t-1} \cdots, \tag{2.22}$$

we can see that the neighboring transformations $V_t V_{t-1}^\dagger$ can merge into a single layer of local scramblers. Therefore a locally scrambled quantum dynamics can also be viewed as repeatedly applying the on-site scrambling $V_t V_{t-1}^\dagger$ followed by the inter-site unitary U_t'. In this way, the quantum many-body state is always sufficiently scrambled on each qudit and the scrambling is uncorrelated in time, such that the information about local basis choice does not pass on from step to step. Separating each step of the unitary evolution by local scramblers is our key assumption about the quantum dynamics, which enables us to proceed.

The entanglement dynamics induced by the locally scrambled quantum dynamics is Markovian, and admits a simple transfer matrix description. To see this, we evaluate the final state EF averaging over all locally scrambled unitary ensembles at different steps

$$|W_{\Psi_{t+1}}\rangle = \mathop{\mathbb{E}}_{U_t} \mathop{\mathbb{E}}_{U_{t-1}} \cdots |W_{U_t U_{t-1} \cdots \Psi_0}\rangle. \tag{2.23}$$

Applying Eq. (2.19), we arrive at the recurrent equation for the ensemble averaged EF state

$$|W_{\Psi_{t+1}}\rangle = \hat{T}_t |W_{\Psi_t}\rangle, \tag{2.24}$$

where we have introduced the transfer matrix

$$\hat{T}_t = \hat{W}_{U_t} \hat{W}_{\mathbb{1}}^{-1} \tag{2.25}$$

to evolve the EF state $|W_{\Psi_t}\rangle$ according to the EF of the single-step unitary U_t. As summarized in Fig. 2.3, Eq. (2.25) is the key equation that bridges the quantum dynamics and entanglement dynamics, allowing us to predict the evolution of entanglement properties of a quantum state based on the entanglement properties of the unitary operator applied at each time step. If we further assume *locality* of the quantum dynamics such that $U_t = \bigotimes_x U_{t,x}$ can be decomposed into products of non-overlapping local unitary gates $U_{t,x}$ (each gate only acts on a few qudits and its spatial position is labeled by x), the EF operator $\hat{W}_U$ can be factorized in the same manner

following Eq. (2.14)

$$\hat{W}_{U_t} = \bigotimes_x \hat{W}_{U_{t,x}},$$ (2.26)

where $\hat{W}_{U_{t,x}}$ is the EF operator for each local unitary gate, which can be easily computed (as it only involves a few qudits). Along this line, the transfer matrix $\hat{T}_t$ can be constructed purely based on our knowledge about the EF of each unitary gate involved in the quantum dynamics.

Using Eq. (2.24), we can evolve the EF of any initial quantum state in time, given the locally scrambled quantum dynamics. The time evolution of the (2nd Rényi) EE can be read out from the EF by

$$S^{(2)}[\boldsymbol{\sigma}](t) = -\log\langle\boldsymbol{\sigma}|W_{\Psi_t}\rangle,$$ (2.27)

following Eq. (2.9). Strictly speaking, there is a subtle issue about exchanging the order of the logarithm with all the ensemble average in Eq. (2.23). We are typically more interested in the ensemble average of the EE other than the EF. So the correct average for the EF should be the geometric mean $\exp(\mathbb{E}\log W[\boldsymbol{\sigma}])$, but we are replacing it by the algebraic mean $\mathbb{E}W[\boldsymbol{\sigma}]$ in Eq. (2.23), which always over estimates the EF and hence underestimates the EE. So the EE obtained in Eq. (2.27) only serves as a lower bound of the ensemble averaged EE. We may treat this lower bound as an approximation, but we can not claim that it is always a good approximation, because there are known scenarios where this approximation is problematic. For example, near the entanglement transition[13, 111, 207] where critical fluctuation is important, this approximate treatment gives wrong answers about the universality class and critical exponents. There have been more rigorous treatments developed in Ref. [199, 207] using replica tricks, but we will not pursuit that direction in this paper. For thermalizing dynamics and volume-law states, we believe that the lower bound estimation in Eq. (2.27) will provide a decent approximation, because the EF of thermalizing state contains strong ferromagnetic correlation to suppress the spin fluctuation, which allows us to replace the geometric mean by the algebraic mean as the fluctuation is small. We will rely on numerical simulations in Sec. 2.3 to justify this assumption.

To conclude, the EF formalism provides a concise description for the entanglement

39

dynamics, when the underlying quantum dynamics is locally scrambled. However, there are also several limitations of locally scrambled quantum dynamics. First of all, the dynamics is not translation invariant in time, because the local scrambles at each step must be sampled independently. As a result, energy is not conserved under such dynamics. Secondly, global symmetry[117, 173] can not be implemented in the current scheme, because symmetry representations on each site will all be scrambled together, such that the symmetry can not be preserved. Finally, in lack of the local-basis-specific information, we can not discuss the operator dynamics for specific local operators[175] (but we can discuss operator averaged behaviors). To go beyond the local scrambling assumption, one idea could be to gradually introduce the correlation of unitary gates in time. But we will leave that for future study. We believe that our discussion of the locally scrambled quantum dynamics will set a cornerstone for future developments.

2.2.5 Entanglement Feature Hamiltonian

In the previous section, we have derived the dynamic equation Eq. (2.24) for EF states under discrete time dynamics. We can also consider the continuum limit of the dynamics, where we refine the time step and take U_t to be close to identity (up to local basis transformation).

For example, we can consider generating U_t by a local Hamiltonian for a short amount of "time" $\varepsilon \ll 1$ with the local basis scrambled

$$U_t = V_t^\dagger e^{-i\varepsilon H} V_t, \tag{2.28}$$

where $V_t = \bigotimes_{i=1}^{L} V_{t,i}$ is a layer of local scramblers and each scrambler $V_{t,i}$ is an on-site unitary operator independently drawn from Haar random ensemble. The full unitary evolution $U = \prod_t U_t$ is given by the time-order product. The onsite scrambling does not generate entanglement (among different sites). The entanglement generation and propagation all depend on the inter-site couplings in the Hamiltonian H. As ε is small, the entanglement dynamics will be slow (smooth) enough that admits a continuum time description. We will study this model in more details later,

but the goal here is to first establish a Hamiltonian formulation for the evolution of EF state in the continuum limit.

When U_t is close to an identity operator (up to local basis transformations), its EF operator $\hat{W}_{U_t}$ will approach $\hat{W}_{\mathbb{1}}$, hence the transfer matrix $\hat{T}_t = \hat{W}_{U_t}\hat{W}_{\mathbb{1}}^{-1}$ will also be close to the identity operator $\hat{\mathbb{1}}$ (in the EF Hilbert space). It turns out that the difference between $\hat{T}_t$ and $\hat{\mathbb{1}}$ is of the order ε^2 (not ε as one may expect). A general argument for this property is as follows. Given U_t in Eq. (2.28), its EF is described by

$$W_{U_t}[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} e^{-i\varepsilon\mathbb{H}} \mathscr{X}_{\boldsymbol{\tau}} e^{i\varepsilon\mathbb{H}}), \tag{2.29}$$

with $\mathbb{H} = H \otimes \mathbb{1} + \mathbb{1} \otimes H$. It can be shown that $W_{U_t}[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ must be even in ε, because it is real by definition but ε comes with the imaginary unit in Eq. (2.29), thus the odd-power expansions of $W_{U_t}[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ in ε could only be imaginary, and must therefore vanish altogether. So the operators $\hat{W}_{U_t}$ and $\hat{T}_t$ are even in ε, hence the leading order deviation of $\hat{T}_t$ from $\hat{\mathbb{1}}$ is of the order ε^2.

Given this, we expand $\hat{T}_t$ around the identity operator $\hat{\mathbb{1}}$ and define the *entanglement feature Hamiltonian* (EF Hamiltonian)

$$\hat{H}_{\mathrm{EF}} = \frac{1}{\varepsilon^2}(\hat{\mathbb{1}} - \hat{T}_t) = \frac{1}{\varepsilon^2}(\hat{\mathbb{1}} - \hat{W}_{U_t}\hat{W}_{\mathbb{1}}^{-1}), \tag{2.30}$$

such that the recurrent equation Eq. (2.24) transforms to an imaginary-time Schrödinger equation in the continuum limit of $\varepsilon \ll 1$,

$$\partial_t |W_{\Psi_t}\rangle = -\hat{H}_{\mathrm{EF}}|W_{\Psi_t}\rangle. \tag{2.31}$$

The differentiation $\partial_t|W_{\Psi_t}\rangle$ should be considered as the limit of $(|W_{\Psi_{t+\varepsilon^2}}\rangle - |W_{\Psi_t}\rangle)/\varepsilon^2$, where ε^2 serves as the infinitesimal time step. In general, $\hat{H}_{\mathrm{EF}}$ can be time-dependent, but let us omit the explicit time dependence for simplicity. The locality of the EF operator $\hat{W}_{U_t}$ as discussed in Eq. (2.26) translates to the locality of the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$, which allow us to write

$\hat{H}_{\text{EF}} = \sum_x \hat{H}_x$ as sum of local terms. In principle, the specific form of these local terms $\hat{H}_x$ can be derived from the terms in the quantum many-body Hamiltonian H that drives the quantum dynamics, which we will demonstrate later in Sec. 2.3.3. However, even if we have no specific knowledge about H, we can already learn a lot about $\hat{H}_{\text{EF}}$ based on the general properties of entanglement dynamics. In the following, we will show how the physical constraint of entanglement dynamics can pin down the general form of the EF Hamiltonian.

Let us consider the two-local EF Hamiltonian, meaning that the local terms $\hat{H}_x$ span over two sites at most. We find that the most general two-local EF Hamiltonian should take the following form

$$\hat{H}_{\text{EF}} = \sum_{i,j} g_{ij} \frac{1 - Z_i Z_j}{2} e^{-\beta_{ij} X_i X_j - \delta(X_i + X_j)}, \tag{2.32}$$

where $g_{ij} \geq 0$ and $\beta_{ij} \in \mathbb{R}$ are model parameters and the constant δ is fixed by the qudit dimension d via $\coth \delta = d$. Here X_i, Z_i are Pauli operators acting on the ith Ising spin (that labels the entanglement region). Each local term in the Hamiltonian consists of a term $e^{-\beta_{ij} X_i X_j - \delta(X_i + X_j)}$ that fluctuates Ising spins, followed by a ferromagnetic projection operator $(1 - Z_i Z_j)/2$. Although we call $\hat{H}_{\text{EF}}$ a Hamiltonian, it is not a Hermitian operator as expected in conventional quantum mechanics, because fluctuation term and the projection term do not commute. As a result, the left- and the right-eigenstates of $\hat{H}_{\text{EF}}$ could be different. The coupling strength g_{ij} describes the entangling power of the quantum dynamics, i.e. the velocity that the entanglement builds up between sites i and j if initialized from a product state.

The postulated form of $\hat{H}_{\text{EF}}$ in Eq. (2.32) is constrained by the following physical requirements (or assumptions).

- Pure state remains pure under quantum dynamics (i.e. a $\mathbb{Z}_2$ Ising symmetry),

$$[\hat{H}_{\text{EF}}, \prod_i X_i] = 0. \tag{2.33}$$

An important entanglement property of pure states is that the EE of a region A should

be the same as that of its complement $\bar{A}$, therefore the pure state EF must be invariant under Ising symmetry, i.e. $W_{|\Psi\rangle}[\boldsymbol{\sigma}] = W_{|\Psi\rangle}[-\boldsymbol{\sigma}]$, which can be equivalently written as $|W_\Psi\rangle = \prod_i X_i |W_\Psi\rangle$. Since any quantum dynamics (described by a unitary evolution) will preserve the purity of the quantum state, the entanglement dynamics should also respect this Ising symmetry, such that the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$ must commute with the symmetry operator $\prod_i X_i$ as asserted in Eq. (2.33).

- EE must vanish for empty entanglement regions,

$$\langle \uparrow\uparrow\uparrow \cdots | \hat{H}_{\mathrm{EF}} = 0. \tag{2.34}$$

By empty entanglement region, we mean $A = \emptyset$ is an empty set, which correspond to the Ising configuration $\boldsymbol{\sigma} = \uparrow\uparrow\uparrow \cdots \equiv \Uparrow$ (i.e. $\forall i : \sigma_i = +1$). Hereinafter we use the symbol $\Uparrow$ to denote the all-up configuration to simplify the notation. When the entanglement region is empty, the EE must be zero, i.e. $S^{(2)}[\Uparrow] = 0$. This requires $\langle \Uparrow | W_\Psi \rangle = W_{|\Psi\rangle}[\Uparrow] = e^{-S^{(2)}[\Uparrow]} = 1$ to remain at constant under any entanglement dynamics. Now suppose $|W_{\Psi_t}\rangle$ is time dependent under the entanglement dynamics. Taking the time derivative on both sides of $\langle \Uparrow | W_{\Psi_t} \rangle = 1$ and apply the dynamic equation Eq. (2.31), we can see that $\langle \Uparrow | \partial_t | W_{\Psi_t} \rangle = -\langle \Uparrow | \hat{H}_{\mathrm{EF}} | W_{\Psi_t} \rangle = 0$ must hold for any EF state $|W_{\Psi_t}\rangle$, therefore we must require $\langle \Uparrow | \hat{H}_{\mathrm{EF}} = 0$ as claimed in Eq. (2.34).

- Statistical time-reversal symmetry of random unitary ensembles,

$$\hat{W}_{\mathbb{1}} \hat{H}_{\mathrm{EF}}^{\mathsf{T}} = \hat{H}_{\mathrm{EF}} \hat{W}_{\mathbb{1}}. \tag{2.35}$$

We assume that the random unitary gates in the circuit are statistically invariant under time-reversal, meaning that U_t and $U_t^\dagger$ will appear with equal probability in the unitary ensemble. Then according to the definition of unitary EF in Eq. (2.5), the time-reversal symmetry implies to $W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}] = W_U[\boldsymbol{\tau}, \boldsymbol{\sigma}]$, i.e. $\hat{W}_U^{\mathsf{T}} = \hat{W}_U$. As a special case, we also

have $\hat{W}_\mathbb{1}^\mathsf{T} = \hat{W}_\mathbb{1}$ by definition. Transposing both sides of $\hat{W}_U \hat{W}_\mathbb{1}^{-1} = \hat{\mathbb{1}} - \varepsilon^2 \hat{H}_\text{EF}$, we obtain $\hat{W}_\mathbb{1}^{-1} \hat{W}_U = \hat{\mathbb{1}} - \varepsilon^2 \hat{H}_\text{EF}^\mathsf{T}$. Therefore $\hat{H}_\text{EF}^\mathsf{T}$ and $\hat{H}_\text{EF}$ must be related by $\hat{W}_\mathbb{1} \hat{H}_\text{EF}^\mathsf{T} = \hat{H}_\text{EF} \hat{W}_\mathbb{1}$ as stated in Eq. (2.35). One known scenario that the statistical time-reversal symmetry is broken is that the unitary operators cyclically permute the qudit along one direction, which describes a quantum dynamics that has dynamic anomaly.[132, 179] We conjecture that the statistical time-reversal symmetry effectively restricts the quantum dynamics to be anomaly free.

With these conditions, we can start from a generic two-local Hamiltonian $\hat{H}_\text{EF} = \sum_{i,j} \hat{H}_{ij}$ and derive the generic form of Eq. (2.32). First of all, the Ising symmetry in Eq. (2.33) restricts $\hat{H}_{ij}$ to be a linear combination of the following operators $\hat{H}_{ij} = x_1 + x_2 X_j + x_3 X_i + x_4 X_i X_j + x_5 Y_i Y_j + x_6 Y_i Z_j + x_7 Z_i Y_j + x_8 Z_i Z_j$, which contains all the two-local operators that commute with $X_i X_j$. Then the left-null-state requirement in Eq. (2.34) further requires $x_1 = -x_8, x_2 = ix_7, x_3 = ix_6, x_4 = x_5$, which reduce $\hat{H}_{ij}$ to $(1 - Z_i Z_j)(x_1 + x_2 X_j + x_3 X_i + x_4 X_i X_j)$. Finally, the statistical time-reversal symmetry in Eq. (2.35) requires

$$x_2 = x_3 = -\frac{d(x_1 + x_4)}{d^2 + 1}, \tag{2.36}$$

leaving only two independent parameters x_1 and x_4. This relation can be resolved by introducing another two parameters g and β to parametrize $x_1 + x_2 X_j + x_3 X_i + x_4 X_i X_j = \frac{g}{2} e^{-\beta X_i X_j - \delta(X_i + X_j)}$ with $\coth \delta = d$ fixed, such that

$$
\begin{aligned}
x_1 &= \frac{g(d^2 \cosh\beta - \sinh\beta)}{2(d^2 - 1)}, \\
x_2 = x_3 &= -\frac{gde^{-\beta}}{2(d^2 - 1)}, \\
x_4 &= \frac{g(\cosh\beta - d^2 \sinh\beta)}{2(d^2 - 1)},
\end{aligned}
\tag{2.37}
$$

automatically satisfies Eq. (2.36). The resulting local term reads $\hat{H}_{ij} = g\frac{1 - Z_i Z_j}{2} e^{-\beta X_i X_j - \delta(X_i + X_j)}$,

which matches the form of Eq. (2.32).

2.2.6 Universal Behaviors of Entanglement Dynamics

The generic form of the EF Hamiltonian $\hat{H}_{\text{EF}}$ in Eq. (2.32) is already useful to illustrate several universal behaviors about the entanglement dynamics. Suppose the EF Hamiltonian admits the following spectral decomposition

$$\hat{H}_{\text{EF}} = \sum_a |R_a\rangle \lambda_a \langle L_a|, \tag{2.38}$$

where $|R_a\rangle$ and $\langle L_a|$ are respectively the right- and left-eigenstate of the eigenvalue λ_a. The right-eigenstate is related to the corresponding left-eigenstate by $|R_a\rangle \propto (\langle L_a|\hat{W}_{\mathbb{1}})^{\mathsf{T}}$, which follows from Eq. (2.35). Then the Schrödinger equation for EF state Eq. (2.31) can be formally solved as

$$|W_{\Psi_t}\rangle = \sum_a e^{-\lambda_a t} |R_a\rangle \langle L_a|W_{\Psi_0}\rangle. \tag{2.39}$$

The dynamics of the EE can be inferred from Eq. (2.27) as

$$
\begin{aligned}
S^{(2)}[\boldsymbol{\sigma}](t) &= -\log\langle\boldsymbol{\sigma}|W_{\Psi_t}\rangle \\
&= -\log\sum_a e^{-\lambda_a t}\langle\boldsymbol{\sigma}|R_a\rangle\langle L_a|W_{\Psi_0}\rangle.
\end{aligned}
\tag{2.40}
$$

Independent of the choice of model parameters g_{ij}, β_{ij}, the EF Hamiltonian $\hat{H}_{\text{EF}}$ has the following spectral properties:

- $\hat{H}_{\text{EF}}$ is positive semi-definite (all its eigenvalues $\lambda_a \geq 0$ are real and non-negative),

- $\hat{H}_{\text{EF}}$ always has (at least) a zero eigenvalue $\lambda_0 = 0$ in the $\mathbb{Z}_2$ (Ising parity) even sector, whose left- and right-eigenstates are

$$
\begin{aligned}
\langle L_0| &= \frac{\langle\Uparrow| + \langle\Downarrow|}{2}, \\
|R_0\rangle &= |W_{\text{Page}}\rangle.
\end{aligned}
\tag{2.41}
$$

45

The left zero mode $\langle L_0 |$ is the Ising symmetric superposition of the all-up and the all-down states. The right zero mode $|R_0\rangle$ is the Page EF state given in Eq. (2.11).

The proof can be found in Appendix A.4. With these results, we can obtain several universal behaviors of entanglement dynamics with local scrambling in the short-time and long-time limit.

In the short-time limit ($t \to 0$), expanding the solution of EF state in Eq. (2.40) to first order in t, we can show that the EE grows *linearly* in time,

$$S^{(2)}[\boldsymbol{\sigma}](t) = S^{(2)}[\boldsymbol{\sigma}](0) + v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}] \times t + \mathcal{O}(t^2), \tag{2.42}$$

where the linear-time coefficient $v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}]$ is the entanglement growth rate, which is related to the entanglement velocity introduced in Ref.[90, 144]

$$v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}] = \partial_t S^{(2)}[\boldsymbol{\sigma}](0) = \frac{\langle \boldsymbol{\sigma} | \hat{H}_{\mathrm{EF}} | W_{\Psi_0} \rangle}{\langle \boldsymbol{\sigma} | W_{\Psi_0} \rangle}. \tag{2.43}$$

The entanglement velocity $v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}]$ characterizes how fast the EE grows in a given entanglement region specified by $\boldsymbol{\sigma}$. It is proportional to the matrix element of the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$, as can be seen in Eq. (2.43), because $\hat{H}_{\mathrm{EF}}$ is the time-evolution generator that drives the entanglement dynamics. In particular, if the initial state is a generic product state, i.e. $|W_{\Psi_0}\rangle = |W_{\mathrm{prod}}\rangle = \sum_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle$ as given in Eq. (2.10), the entanglement velocity $v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}]$ admits an explicit formula

$$v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}] = \sum_{\langle ij \rangle} \tilde{g}_{ij} \frac{1 - \sigma_i \sigma_j}{2}, \tag{2.44}$$

where $\tilde{g}_{ij} = g_{ij} e^{-\beta_{ij} - 2\delta} \geq 0$ is the effective coupling. Eq. (2.44) describes how the entanglement velocity $v_{\mathrm{E}}^{(2)}$ depends on the choice of the entanglement region $\boldsymbol{\sigma}$. It is obvious that the entanglement velocity $v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}] \geq 0$ is non-negative for all choices of entanglement regions, because the EE can only *grow* from an unentangled product state. If $\tilde{g}_{ij} = \tilde{g}$ is uniform through out the system, $v_{\mathrm{E}}^{(2)}[\boldsymbol{\sigma}]$ will simply be proportional to the number of domain walls in the Ising configuration $\boldsymbol{\sigma}$,

which is also the area $|\partial A|$ of the entanglement region A. Therefore the entanglement velocity follows the *area-law* scaling,

$$v_{\mathrm{E}}^{(2)} = \tilde{g}|\partial A|, \tag{2.45}$$

which can be expected from the locality of the entanglement dynamics in our setup.

In the long-time limit ($t \to \infty$), the EF state is dominated by the zero mode (assuming the zero mode is unique) and all the other modes decays exponentially with time. The positive semi-definite property of the EF Hamiltonian, i.e. $\lambda_a \geq 0$, ensures that all modes (except the zero mode) will decay exponentially in time. As $t \to \infty$, Eq. (2.39) reduces to

$$|W_{\Psi_\infty}\rangle = |R_0\rangle\langle L_0|W_{\Psi_0}\rangle, \tag{2.46}$$

with the left and right zero modes given by Eq. (2.41). Given that the EE vanishes in trivial regions, $\langle\Uparrow|W_\Psi\rangle = \langle\Downarrow|W_\Psi\rangle = 1$, so $\langle L_0|W_\Psi\rangle = 1$ for any EF state $|W_\Psi\rangle$. Then Eq. (2.46) results in

$$|W_{\Psi_\infty}\rangle = |R_0\rangle = |W_{\mathrm{Page}}\rangle, \tag{2.47}$$

meaning that the EF always converge to that of the Page state in the long-time limit regardless what the initial state is. All states are doomed to thermalize under the quantum dynamics with local scrambling. The Page state will be their final destiny, whose EE reads

$$S^{(2)}[\boldsymbol{\sigma}] = -\log \frac{\cosh(\eta \sum_{i=1}^{L} \sigma_i)}{\cosh(\eta L)}, \tag{2.48}$$

which follows from Eq. (2.11). For $|A| \ll L$, the EE exhibits the *volume-law* scaling

$$S^{(2)}(A) = 2\eta|A|, \tag{2.49}$$

with the volume law coefficient given by $2\eta = \log d$. It worth mention that the above conclusion is based on the assumption that the zero mode is unique. If there are other degenerated zero

modes (other than $|W_{\text{Page}}\rangle$), the final state may not converge to the Page state and the system can evade thermalization. We will discuss such a possibility later with a more concrete model in Sec. 2.3.2.

2.2.7 Excitation Spectrum of the Entanglement Feature Hamiltonian

Having discussed the ground state property of the EF Hamiltonian $\hat{H}_{\text{EF}}$, let us turn to the low-lying excited states of $\hat{H}_{\text{EF}}$. According to Eq. (2.39), every eigenmode with finite eigenenergy λ_a will decay exponentially in time as $e^{-\lambda_a t}$. Eventually, only the ground state with zero eigenenergy ($\lambda_0 = 0$) would survive, and the system thermalizes to Page states. Hence the low-energy excitation spectrum determines how the EE approaches to its thermal limit in the late-time regime. Here we will focus on the spectrum of two kinds of excitations, namely the two-domain-wall excitation and the single spin-flip excitation, which dominate the low-energy excitations. We obtain the analytical expression of their dispersion relations, from which we can estimate the excitation gap and determine the relaxation time. In Sec. 2.3.3, we further compare the analytically estimated relaxation time with the numerical fitted one, and find good consistency.

For simplicity, we assume the parameters g_{ij}, β_{ij} in the EF Hamiltonian $\hat{H}_{\text{EF}}$ are spatially homogeneous (i.e $g_{ij} = g, \beta_{ij} = \beta$). For the parameter β, any unitary evolution generated from Hamiltonian $e^{-i\varepsilon H}$ would have nonzero β only at the order of $\mathcal{O}(\varepsilon^2)$ in small ε limit (see Appendix A.7 for details). Hence, we will take $\beta = 0$ in the following. More general results for $\beta \neq 0$ can be found in Appendix A.5 and Appendix A.6. To first gain some intuitions about the excitation spectrum, we start with the exact diagonalization (ED) of EF Hamiltonian. The result is shown in Fig. 2.4. Apart from the eigenenergy λ_a, every state $|R_a\rangle$ is also label by its crystal momentum k_a, which is defined through its translation eigenvalue as $\mathsf{T}|R_a\rangle = e^{ik_a}|R_a\rangle$, where the translation operator T is defined by its action on the Ising basis $\mathsf{T}|\sigma_1\sigma_2\cdots\sigma_L\rangle = |\sigma_L\sigma_1\cdots\sigma_{L-1}\rangle$. One can see that above the ground state at $\lambda_0 = 0$ and $k_0 = 0$, there is a continuum of excited states.
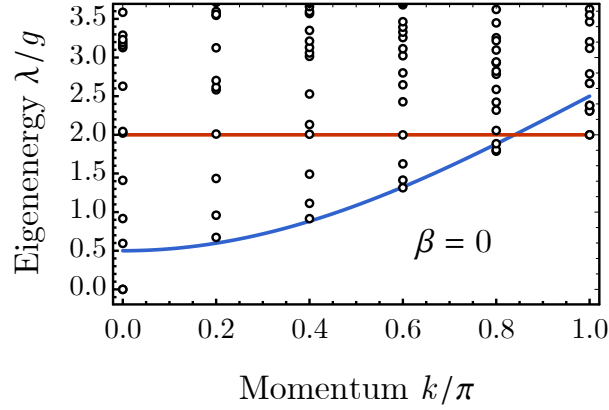
Figure 2.4. We perform exact diagonalization for the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$ with $\beta = 0, L = 10$. Each small circle represents an eigenstate label by its eigenenergy λ and its crystal momentum k. The blue curve is the analytical result of two-domain-wall ansatz Eq. (2.52). The red curve is the analytical result of single spin-flip ansatz Eq. (2.54).

To better understand these excited states, we look into their wave function. We realize that the excitation can be classified based on the number of domain walls in the *left*-eigenstate. For instance, $\langle \uparrow \cdots \uparrow \downarrow \cdots \downarrow \uparrow \cdots \uparrow |$ is an example of two-domain-wall states. As mentioned in Eq. (2.41), the left ground state $\langle L_0 | = (\langle \Uparrow | + \langle \Downarrow |)/2$ contains no domain wall and hence no excitation. Other excited left-eigenstate will be a superposition of states of different domain-wall number. Note that the corresponding right eigenstate can be obtained from $|R\rangle = (\langle L | \hat{W}_{\mathbb{1}})^{\mathsf{T}}$. Fig. 2.5 shows the weights of different domain-wall states in the lowest-energy excited state of various momenta. The ED result indicates that the lowest-energy excited state mainly consists of two-domain-wall states, so we will focus on them in the following.

Based on the numerical observation, we approximate low energy excitation by the two-domain-wall (2DW) ansatz state as follows,

$$\langle k | \propto \sum_{i_1, i_2} e^{ik \frac{i_1 + i_2}{2}} \phi^*_{i_2 - i_1} \langle i_1, i_2 |, \tag{2.50}$$

where $\langle i_1, i_2 | = \langle \Uparrow | \prod_{i=i_1}^{i_2 - 1} X_i$ is a two-domain-wall state with domain walls located at i_1 and i_2. k labels the center of mass momentum of the pair of domain walls. $\phi_{\Delta i}$ is a variational wave

49
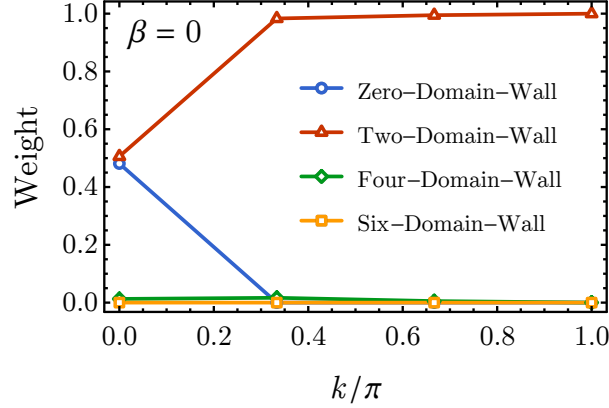
Figure 2.5. We perform exact diagonalization for the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$ with $g = 1, \beta = 0, L = 6$. The weight is defined as follows: the left excited state $\langle L|$ can be expressed as the linear combination of two-, four- and six-domain-wall states with the coefficient (c_2, c_4, c_6). The weight of individual type is equal to $|c_n|^2$. For $k = 0$, zero-domain wall states take half of the weight in the lowest-energy excited state. However, they have no contribution in dispersion relation since their eigenenergy is zero.

function that describes the relative motion between the domain walls. We can then evaluate the energy expectation value $\lambda(k)$ on the ansatz state $\langle k|$,

$$\lambda_{\mathrm{2DW}}(k) = \frac{\langle k|\hat{H}_{\mathrm{EF}}\hat{W}_{\mathbb{1}}|k\rangle}{\langle k|\hat{W}_{\mathbb{1}}|k\rangle}, \tag{2.51}$$

where $\hat{W}_{\mathbb{1}}|k\rangle$ is understood as the corresponding right-state of the ansatz left-state $\langle k|$. Two assumptions are made to derive the analytical expression of the dispersion relation. The first assumption is that these domain walls have no interaction with each other and thus $\phi_{\Delta i}$ can be approximated by plane waves. The second assumption is the thermodynamic limit $L \to \infty$, which would simplify the calculation but suppress the contribution from short two-domain-wall states (see Appendix A.5 for details). Based on these assumptions, the dispersion relation for $\beta = 0$ can be derived as,

$$\lambda_{\mathrm{2DW}}(k) = 2g\left(1 + \frac{1}{d^2}\right) - \frac{4g}{d}\cos\frac{k}{2} + \mathcal{O}(d^{-3}). \tag{2.52}$$

50

The band minimum is at $k = 0$, which defines the excitation gap

$$\Delta = \min_k \lambda(k) = 2g\left(1 - \frac{1}{d}\right)^2 + \mathcal{O}(d^{-3}). \tag{2.53}$$

It turns out that the gap remains open (i.e. $\Delta > 0$) for any finite $g > 0$.
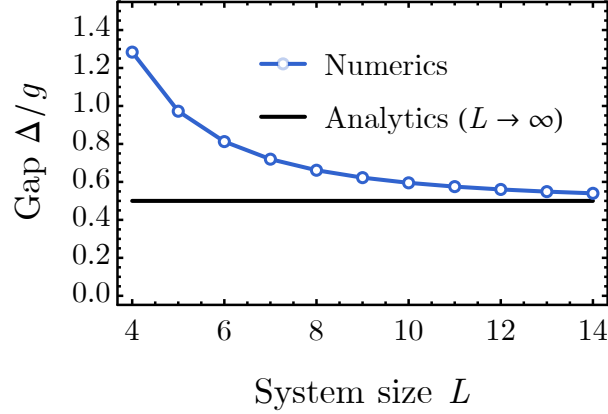
Figure 2.6. Comparison of the excitation gap between the finite-size ED result and the analytical result of two-domain-wall ansatz in the thermodynamic limit for the qudit dimension $d = 2$. The analytical result $\Delta = g/2$ is given by Eq. (2.53).

The comparison between ED result (black circles) and our analytical expression (blue curve) is shown in Fig. 2.4. The lower-edge of the excitation spectrum is pretty well captured by the two-domain-wall ansatz. The comparison also reveals a finite-size-effect in the spectrum. In Fig. 2.6, we show how the gap at $k = 0$ (from ED) approaches to the analytic result of Eq. (2.53) with increasing system size L. We also observe a systematic deviation of our analytical result from the excitation edge near $k = \pi$. The reason is that the eigenstate around $k = \pi$ is dominated by single-site excitations, where the domain-walls are next to each other such that their interaction can not be ignored. To capture the interaction effect, we switch to another ansatz state, which describes the motion of a tightly-bound domain-wall pair, or equivalently a single spin-flip (SSF) excitation (see Appendix A.6 for details). The dispersion of the SSF excitation reads

$$\lambda_{\text{SSF}}(k) = 2g, \tag{2.54}$$

which turns out to be independent of the qudit dimension d and the momentum k. This dispersion relation basically passes a series of points in Fig. 2.4 and only becomes the lowest excited state around $k = \pi$.

2.3 Applications and Numerics

2.3.1 Models of Locally Scrambled Quantum Dynamics

In the following, we will apply the entanglement feature formalism to several scenarios of locally scrambled quantum dynamics. We will consider two types of models: random circuit models with discrete time as in Fig. 2.7(a), and Hamiltonian generated evolutions with local scramblers in the limit of continuous time as in Fig. 2.7(b). For the discrete time models, namely *locally scrambled random circuits*, we will adopt the transfer matrix method to study the entanglement dynamics. For the continuous time models, namely *locally scrambled Hamiltonian dynamics*, we will apply the EF Hamiltonian approach.
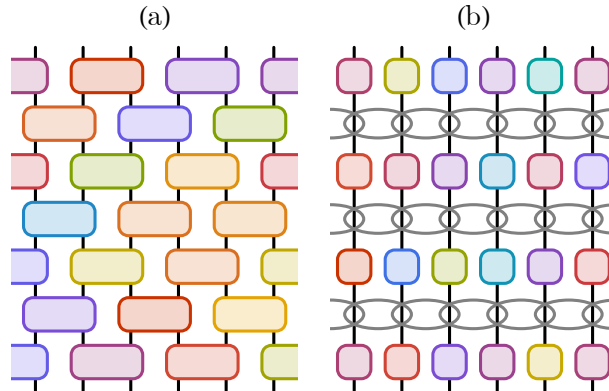
Figure 2.7. (a) Locally scrambled random circuit. The gates are drawn independently in both space and time (as indicated by different colors). (b) Locally scrambled Hamiltonian dynamics. The unitary operators generated by the local Hamiltonian are overlapping gray ovals in each layer. The on-site scramblers are uncorrelated in both space and time (as indicated by different colors).

The random circuit we consider will be of the "brick wall" structure as shown in Fig. 2.7(a). The entire unitary circuit $U = \prod_t U_t$ is constructed by stacking layers of unitary gates.

Each layer U_t is described by

$$U_t = \begin{cases} \bigotimes_x U_{t;2x-1,2x} & t \in \text{odd}, \\ \bigotimes_x U_{t;2x,2x+1} & t \in \text{even}, \end{cases} \qquad (2.55)$$

where $U_{t;ij}$ denotes the two-qudit unitary gate acting on sites i and j at time t. Each gate $U_{t;ij}$ is independently sampled from a locally scrambled unitary ensemble, so the quantum circuit U will be dubbed as a locally scrambled random circuit. In fact, any gate can be made locally scrambled by symmetrizing over local basis transformations as constructed in Eq. (2.18). The construction here is more general than the Haar random unitary circuit[155], as the unitary gate here does not need to be Haar random. As the quantum state evolves by $|\Psi_{t+1}\rangle = U_t|\Psi_t\rangle$, the corresponding EF state evolves by $|W_{\Psi_{t+1}}\rangle = \hat{T}_t|W_{\Psi_t}\rangle$. The transfer matrix $\hat{T}_t$ follows the same structure as U_t,

$$\hat{T}_t = \begin{cases} \bigotimes_x \hat{T}_{2x-1,2x} & t \in \text{odd}, \\ \bigotimes_x \hat{T}_{2x,2x+1} & t \in \text{even}. \end{cases} \qquad (2.56)$$

According to Eq. (2.25), $\hat{T}_{ij}$ is fully determined by the EF of $U_{t;ij}$ via

$$\hat{T}_{ij} = \hat{W}_{U_{t;ij}} \hat{W}_{\mathbb{1}_{ij}}^{-1}. \qquad (2.57)$$

Here we have assumed that $U_{t;ij}$ are drawn from identical unitary ensembles, such that $\hat{T}_{ij}$ is time-independent (despite of the time-dependence in $U_{t;ij}$). In the following, we will provide examples of the locally scrambled two-qudit unitary ensemble. We will use the transfer matrix approach to calculate the entanglement dynamics. The result will be compared with exact numerics by explicitly constructing the random circuit and average the final state EE over random realizations.

Another type of locally scrambled quantum dynamics that we will consider is generated by a local Hamiltonian $H = \sum_{\langle ij \rangle} H_{ij}$, which is a sum of local terms H_{ij} defined on nearest neighboring bonds $\langle ij \rangle$ along a 1D chain. Each step of the unitary evolution U_t is independently

drawn from the locally scrambled unitary ensemble $\mathscr{E}_{e^{-i\varepsilon H}}$ generated by the Hamiltonian H,

$$\mathscr{E}_{e^{-i\varepsilon H}} = \{V^\dagger e^{-i\varepsilon H} V | V = \bigotimes_{i=1}^{L} V_i, V_i \in \text{Haar}\}, \tag{2.58}$$

which may be simply denoted by $U_t = V_t^\dagger e^{-i\varepsilon H} V_t$, as in Eq. (2.28). Combining the adjacent local scramblers following Eq. (2.22), the unitary evolution can be considered as repeatedly applying a short-time unitary evolution $e^{-i\varepsilon H}$ followed by a layer of local scramblers, as illustrated in Fig. 2.7(b). Such dynamics will be called the locally scrambled Hamiltonian dynamics. It is similar to the Brownian random circuit model[130] in that each step of the evolution is driven by a different random Hamiltonian, but our construction is more general in that the random Hamiltonian ensemble only needs to be invariant under local basis transformations other than the full basis transformation of the many-body Hilbert space. For small ε, we can take the continuous time approach to calculate the entanglement dynamics by solving the imaginary-time Schrödinger equation $\partial_t |W_{\Psi_t}\rangle = -\hat{H}_{\text{EF}} |W_{\Psi_t}\rangle$ in Eq. (2.31). It worth mentioning that the locally scrambled quantum dynamics we considered here should be distinguished from Trotterizing a Hamiltonian dynamics. Here, the short-time evolutions $e^{-i\varepsilon H}$ are interrupted by local scramblers, such that they do not combine to a coherent long-time evolution generated by the same Hamiltonian H. The local scramblers destroy the original notion of time. In the quantum dynamics, $e^{-i\varepsilon H}$ advances the quantum state by ε in time, but after the insertion of layers of local scramblers, the entanglement dynamics only progress by ε^2, which is much slower. This phenomenon is analogous to the quantum Zeno effect due to the insertion of measurement. We conjecture that the local scramblers play a similar role as random local measurement in implementing random local basis transformations, such that the quantum dynamics is no longer coherent.

2.3.2 Locally Scrambled Random Circuits

Let us first consider the locally scrambled random circuit as in Fig. 2.7(a). The building blocks of the random circuit are two-qudit unitary gates. Each gate is independently drawn

from local basis independent random ensembles. The EF of a two-qudit unitary operator U_{ij} is completely characterized by two parameters: the cross channel mutual information $I_{ij}^{\times}$ and the tripartite information $I_{ij}^{\triangledown}$. Let us label the input and output channels of the two-qudit unitary by A, B, C, D as shown in Fig. 2.8(a), then $I_{ij}^{\times}$ and $I_{ij}^{\triangledown}$ are defined as follows

$$
\begin{aligned}
I_{ij}^{\times} &= I^{(2)}(A:D) = I^{(2)}(B:C), \\
I_{ij}^{\triangledown} &= I^{(2)}(A:C) + I^{(2)}(A:D) - I^{(2)}(A:CD).
\end{aligned}
\tag{2.59}
$$

The mutual information, such as $I^{(2)}(A:D) = S_A^{(2)} + S_D^{(2)} - S_{AD}^{(2)}$, is understood by treating the unitary gate as a quantum state by bending the input and output legs to the same side, and calculating the operator EE following the definition in Ref. [96, 163].
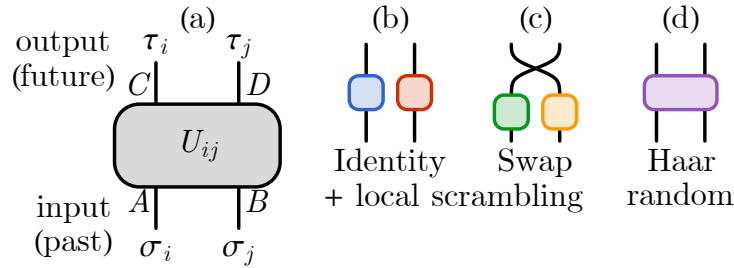
Figure 2.8. (a) A generic two-qudit gate acting on qudits i and j. The input channels are labeled by A and B, and the output channels are labeled by C and D. The EF of the gate will be labeled by the Ising configuration $\boldsymbol{\sigma} = (\sigma_i, \sigma_j)$ on the input side and $\boldsymbol{\tau} = (\tau_i, \tau_j)$ on the output side. (b-d) Examples of local basis independent ensembles of two-qudit gates: (b) identity gate with local scrambling, (c) swap gate with local scrambling, (d) Haar random unitary gate acting on both qudit (local basis automatically scrambled).

In terms of these information measures $I_{ij}^{\times}$ and $I_{ij}^{\triangledown}$ of the unitary gate U_{ij}, the EF operator

$\hat{W}_{U_{ij}}$ is given by

$$
\begin{aligned}
\hat{W}_{U_{ij}} &= d^2(d+X_i)(d+X_j) \\
&\quad - \frac{1-Z_iZ_j}{2}(A_{ij} - B_{ij}X_iX_j), \\
A_{ij} &= d^4(1 - e^{I_{ij}^{\triangledown} - I_{ij}^{\times}}), \\
B_{ij} &= d^2(e^{I_{ij}^{\times}} - 1).
\end{aligned}
\tag{2.60}
$$

The cross channel mutual information $I_{ij}^{\times} \geq 0$ is non-negative by the subadditivity[10] of entropy. It describes the entanglement propagation, as it measures the amount of information transferred between site i and j. The tripartite information $I_{ij}^{\triangledown}$ must be negative for unitary gates[96], and therefore $I_{ij}^{\times} - I_{ij}^{\triangledown} \geq 0$ holds. The negative tripartite information $(-I_{ij}^{\triangledown})$ is proposed[96] to be a description of information scrambling, since it measures the amount of information about A that is encoded in C and D jointly but can not be told by local measurements exclusively performed on C or D.

To gain more intuition about $I_{ij}^{\times}$ and $I_{ij}^{\triangledown}$, let us provide a few examples of local basis independent ensembles of two-qudit gates, as pictured in Fig. 2.8(b-d).

- Identity gate with local scrambling, i.e. two on-site Haar random unitary gates direct product together, as Fig. 2.8(b). In this rather trivial case, we have

$$
I_{ij}^{\times} = I_{ij}^{\triangledown} = 0,
\tag{2.61}
$$

such that the EF operator in Eq. (2.60) reduces to $\hat{W}_{\mathbb{1}} = d^2(d+X_i)(d+X_j)$, consistent with the previous result in Eq. (2.15) by direct evaluation.

- Swap gate with local scrambling, i.e. two on-site Haar random unitary gates followed by an inter-site swap operator, as Fig. 2.8(c). In this case,

$$
I_{ij}^{\times} = 2\log d, \quad I_{ij}^{\triangledown} = 0,
\tag{2.62}
$$

such that the EF operator in Eq. (2.60) reduces to

$$\hat{W}_{\text{swap}} = d^2(d+X_i)(d+X_j)$$
$$- d^2(d^2-1)\frac{1-Z_iZ_j}{2}(1-X_iX_j). \tag{2.63}$$

The swap gate can generate and propagate quantum entanglement due to the non-vanishing cross channel information $I_{ij}^{\times}$. But there is no information scrambling happening *between* the qudits (despite of the sufficient on-site scrambling), because the qubits are simply interchanged by the swap gate, such that local operators do not spread out other than being moved around in the space. The zero scrambling power of the swap gate is reflected in the zero tripartite information $I_{ij}^{\triangledown}$.

- Haar random unitary gate acting on the two qudits, as Fig. 2.8(d). In this case,

$$I_{ij}^{\times} = \log\frac{2d^2}{d^2+1}, \quad I_{ij}^{\triangledown} = \log\frac{4d^2}{(d^2+1)^2}, \tag{2.64}$$

such that the EF operator in Eq. (2.60) reduces to $\hat{W}_{\text{Haar}}$ given in Eq. (2.13), see Appendix A.2 for derivation. The Haar random unitary gate not only propagates quantum entanglement, but also scrambles the quantum information efficiently, as it has a negative tripartite information $I_{ij}^{\triangledown}$ (as long as $d > 1$).

The above are examples of locally scrambled random unitary ensembles. Unitary gates drawn from such ensembles serve as the building block of locally scrambled random circuits. The entanglement dynamics of locally scrambled random circuits can be universally described by the transfer matrix approach as has been discussed in Sec. 2.3.1. On the level of EF, the formulation is exact: the evolution of the average state EF can be precisely calculated from $|W_{\Psi_{t+1}}\rangle = \hat{W}_{U_t}\hat{W}_{\mathbb{1}}^{-1}|W_{\Psi_t}\rangle$ given the EF of the unitary. However, when applying the result to predict the EE, we rely on the assumption that the average EE can be approximated by the negative log of average EF following Eq. (2.27), where we effectively switch the order between

the ensemble average and the logarithm. One major goal of the following is to provide numerical evidences to check this assumption in various different cases. It turns out that the negative log of EF generally provides a good estimate of the averaged EE, which makes our EF formulation useful in describing the entanglement dynamics for a broad class of random unitary circuits.
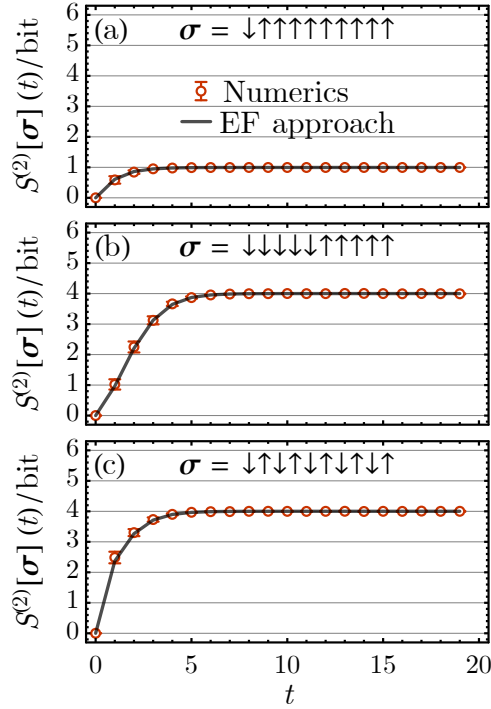
Figure 2.9. The finial state EE of the Haar random circuit on a 10-site system for different choices of the entanglement regions: (a) single site, (b) half-system, (c) alternating[97]. The qudit dimension is $d = 2$ and the entropy is measured in unit of bit $(= \log 2)$.

Our first example is the standard Haar random unitary circuit, where each two-qudit gate is drawn from Haar random unitary ensemble independently. The model has be extensively studied in the literature,[117, 155, 158, 214] and the statistical mechanical model description has been developed by Zhou and Nahum in their pioneering work Ref. [238]. We revisit this model to show that our formalism is equivalent to the Zhou-Nahum approach and can reproduce the known behaviors. Let us first calculate the transfer matrix $\hat{T}_{ij}$ of a single Haar random unitary

gate U_{ij} from its EF. Based on Eq. (2.13) and Eq. (2.17), we obtain

$$\hat{T}_{ij} = \hat{W}_{\text{Haar}}\hat{W}_{\mathbb{1}}^{-1} = \left(1 + \frac{d(X_i + X_j)}{d^2 + 1}\right)\frac{1 + Z_i Z_j}{2}. \tag{2.65}$$

Using the Ising basis $|\sigma_i \sigma_j\rangle$, Eq. (2.65) can be expressed as

$$\begin{aligned}
\hat{T}_{ij} &= |\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{d}{d^2 + 1}(|\uparrow\downarrow\rangle\langle\uparrow\uparrow| + |\downarrow\uparrow\rangle\langle\uparrow\uparrow|) \\
&+ |\downarrow\downarrow\rangle\langle\downarrow\downarrow| + \frac{d}{d^2 + 1}(|\downarrow\uparrow\rangle\langle\downarrow\downarrow| + |\uparrow\downarrow\rangle\langle\downarrow\downarrow|),
\end{aligned} \tag{2.66}$$

which is equivalent to the triangle weights $\triangledown = 1$ and $\triangledown\!\!\!\!\cdot = \cdot\!\!\!\!\triangledown = d/(d^2 + 1)$ that defines the Ising model in Ref. [238]. An equivalent form of the transfer matrix Eq. (2.65) was previously obtained in Ref. [242]. Plugging Eq. (2.65) to Eq. (2.56), we obtain the transfer matrix $\hat{T}_t$ that describes the EF state evolution under the quantum dynamics of the Haar random circuit. We assume the initial state is a product state, s.t. $|W_0\rangle = |W_{\text{prod}}\rangle$. We evolve the EF state by Eq. (2.24). We can then compute the EE following Eq. (2.27) and compare the result with the numerical simulation. In the simulation, we applied randomly sampled unitary gates to an initial product state and measure the final state EE, then perform the ensemble average of the EE. As shown in Fig. 2.9, the EF approach provides pretty good prediction of the EE that matches the numerical result.
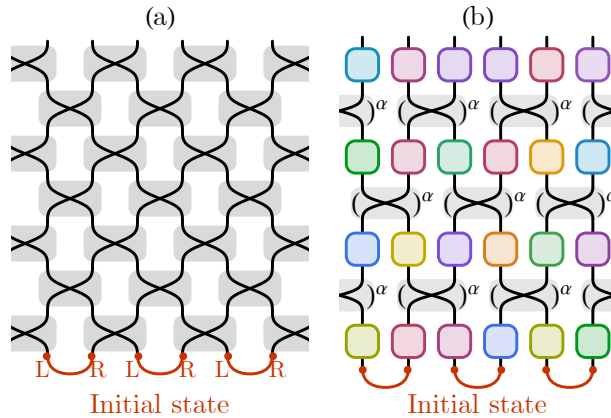
Figure 2.10. (a) Swap gate circuit. Gray blocks mark out the swap gates. (b) Locally scrambled fractional swap gate circuit. Each swap gate is powered by the fraction $0 < \alpha < 1$.

Now let us turn to a new example of locally scrambled random circuits, namely the swap gate circuit, which is designed to mimic the entanglement dynamics in integrable conformal field theories (CFT) where entanglement spreads with the propagation of quasi-particles.[29, 124, 163] The circuit takes the architecture of the brick wall circuit in Fig. 2.7(a) with gates drawn from the locally scrambled swap gate ensemble in Fig. 2.8(c), the resulting circuit is equivalent to an interweaving network as shown in Fig. 2.10(a). The local scramblers in different layers can commute through the swap gates and combine to a single scrambling layer acting on the initial state, which can further be dropped as long as the initial state ensemble is already local basis invariant. For this model, we use a different initial state other than the product state. As illustrated in Fig. 2.10(a), the initial state is chosen to be a product of Einstein-Podolsky-Rosen (EPR) pairs arranged along a one-dimensional chain, whose EF can be described by

$$|W_0\rangle = \prod_x \left(1 + \frac{1}{d}X_{2x-1} + \frac{1}{d}X_{2x} + X_{2x-1}X_{2x}\right)|\Uparrow\rangle. \tag{2.67}$$

For each EPR pair, the qudit labeled by L (or R) will travel to the left (or right) in the swap gate circuit, which mimics the behaviors of left (or right) moving quasi-particles in an integrable CFT. In this way, entanglement spread out along the chain as EPR pairs stretch out, following the steps depicted in Fig. 2.11. On a finite-sized chain with periodic boundary condition, we expect to observe the half-system entanglement entropy to first grow and then decrease in time, and continue to oscillate like this. This recurrent behavior can be perfectly produced by the EF formulation, because, based on Eq. (2.63), the transfer matrix for a single swap gate turns out to be

$$\hat{T}_{ij} = \hat{W}_{\text{swap}}\hat{W}_{\mathbb{1}}^{-1} = \frac{1}{2}(1 + X_iX_j + Y_iY_j + Z_iZ_j), \tag{2.68}$$

which is precisely the swap operator for Ising spins. In this way, the permutation of entangled qudits under the quantum dynamics is equivalently modeled by the permutation of correlated Ising spins in the EF formulation.
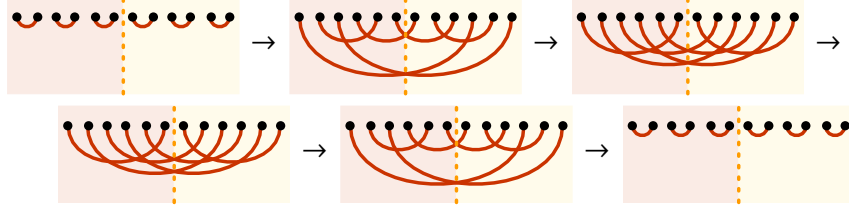
Figure 2.11. Evolution of EPR pairs under the swap gate circuit on a 12-site chain with periodic boundary condition. The entanglement entropy between the left- and right-half system is proportional to the EPR pairs across the cut (indicated by dotted vertical line).

The recurrent (periodically oscillating) behavior of the half-system EE is demonstrated in Fig. 2.12(a), where the EF approach matches the numerical simulation perfectly. The periodic recurrence of the low-entanglement state in the swap gate circuit seems to contradict with our previous conclusion in Sec. 2.2.6 that locally scrambled quantum dynamics generally thermalize. The swap gate circuit evades thermalization because its corresponding EF transfer matrix admits more than one leading eigenstate. Let $\hat{T} = \bigotimes_x \hat{T}_{2x-1,2x} \bigotimes_x \hat{T}_{2x,2x+1}$ be two steps (one period) of the transfer matrix that translates the L (or R) sublattice to the left (or right) by one unit-cell. On a chain of $2n$ sites, the operator $\hat{T}$ has $n^{-1}\sum_{d|n} \varphi(d)4^{n/d}$ fold degenerated eigenstates of eigenvalue 1, with $\varphi(d)$ being the Euler totient function and d running over all divisors of n. These eigenstates can be constructed by taking any Ising basis state and symmetrizing over the cyclic group generated by $\hat{T}$. Their degeneracy can be counted by mapping the problem to the number of n-bead necklaces with four colors,[72] where the four colors correspond to the four choices of $\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$ configurations in each unit-cell. Therefore the Page state is not the unique state that can survive in the long-time limit, and thermalization is not the ultimate fate.

The swap gate circuit model can be generalized by introducing the *fractional swap gate* that interpolates between the identity gate and the swap gate. The fractional swap gate can be written as a fractional power α of the swap gate with $0 < \alpha < 1$

$$\mathsf{SWAP}^\alpha = \frac{1+e^{i\alpha\pi}}{2}\, || + \frac{1-e^{i\alpha\pi}}{2}\, \mathsf{X}. \tag{2.69}$$

61
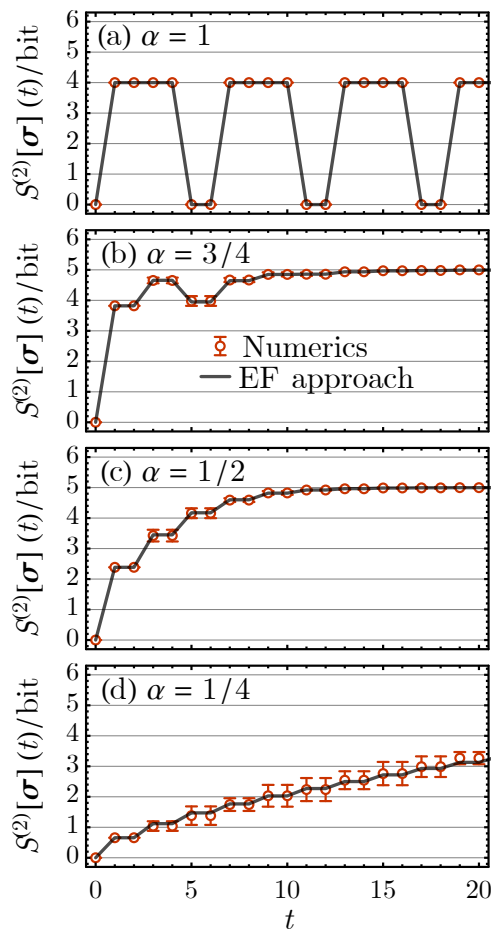
Figure 2.12. Half-system entanglement entropy of the locally scrambled fractional swap gate circuit on a 12-site system with different fraction α: (a) $\alpha = 1$, (b) $\alpha = 3/4$, (c) $\alpha = 1/2$, (d) $\alpha = 1/4$. The model is realized on a 12-site chain with periodic boundary condition. The entanglement region is chosen to be the first 6 sites. The qudit dimension is $d = 2$ and the entropy is measured in unit of bit ($= \log 2$).

The fractional swap gate reduces to the identity gate (or the swap gate) at $\alpha = 0$ (or $\alpha = 1$). But unlike both identity and swap gates which do not scramble quantum information between the two qudits, the fractional swap gate does has finite scrambling power. We can construct a locally scrambled fractional swap gate circuit by starting from the architecture of the random circuit in Fig. 2.7(a) and sampling every gate independently from local basis invariant fractional swap gate ensemble, as illustrated in Fig. 2.10(b). The EF operator of the fractional swap gate follows the general form of Eq. (2.60) with parameters A_{ij} and B_{ij} given by

$$
\begin{aligned}
A_{ij} &= d^2(d^2 - 1)\frac{3 + \cos \alpha \pi}{2} \sin^2 \frac{\alpha \pi}{2}, \\
B_{ij} &= d^2(d^2 - 1) \sin^4 \frac{\alpha \pi}{2}.
\end{aligned}
\tag{2.70}
$$

Based on this result, the corresponding transfer matrix $\hat{T}_{ij}$ can be constructed by Eq. (2.57) and the evolution of the EF state can be calculated following the transfer matrix approach described in Eq. (2.56). In Fig. 2.12(b-d), we compare the EE calculated based on the EF approach with the ensemble averaged EE from numerical simulation. They match perfectly for different values of α. Because the fractional swap gate has finite scrambling power, the recurrence behavior no longer persist and the system can now thermalize. The entanglement dynamics is somewhat between that of the swap gate circuit and the Haar random circuit, in that the EE grows mostly linearly in time with small oscillations, until the EE eventually saturates to the thermal limit. As α becomes small, the system will take longer time (more steps) to thermalize. As shown in Fig. 2.12(d), the oscillation of EE is suppressed and its growth curve is more smooth. In the $\alpha \to 0$ limit, the entanglement dynamics approaches the continuum limit that can be described by the EF Hamiltonian, which is the topic of the following discussion.

2.3.3 Locally Scrambled Hamiltonian Dynamics

Now we turn to the locally scrambled Hamiltonian dynamics as illustrated in Fig. 2.7(b). We consider the local Hamiltonian $H = \sum_{\langle ij \rangle} H_{ij}$ and assume that H_{ij} on every bond is drawn

from a local-basis-independent ensemble of two-qudit Hermitian operators. Equivalently, we can choose H to be a fixed Hamiltonian and construct a locally scrambled unitary ensemble $\mathscr{E}_{e^{-i\varepsilon H}}$ by applying local basis transformations following Eq. (2.58). The quantum dynamics is described by the unitary

$$U = \prod_t \left(V_t e^{-i\varepsilon H} \right), \tag{2.71}$$

where V_t describe the layer of local scramblers at time t, as illustrated in Fig. 2.7(b). The corresponding entanglement dynamics is described by the imaginary-time Schrödinger equation Eq. (2.31), where the EF Hamiltonian takes the form of

$$\hat{H}_{\mathrm{EF}} = \sum_{\langle ij \rangle} g_{ij} \frac{1 - Z_i Z_j}{2} e^{-\beta_{ij} X_i X_j - \delta(X_i + X_j)}. \tag{2.72}$$

It turns out that the parameters $\beta_{ij} \sim \mathcal{O}(\varepsilon^2)$ always vanish in the $\varepsilon \to 0$ limit. The parameters g_{ij} are the only non-trivial parameters to the leading order of ε, which are determined by the local terms H_{ij} in the Hamiltonian

$$\begin{aligned}
g_{ij} = \frac{2}{d^2(d^2 - 1)} \Big(&(\operatorname{Tr} H_{ij})^2 + d^2 \operatorname{Tr}(H_{ij}^2) \\
&- d \operatorname{Tr}_j (\operatorname{Tr}_i H_{ij})^2 - d \operatorname{Tr}_i (\operatorname{Tr}_j H_{ij})^2 \Big).
\end{aligned} \tag{2.73}$$

The detailed derivation of these results can be found in Appendix A.7.

One well-studied example of the locally scrambled Hamiltonian dynamics is the Brownian random circuit,[130] where each step of the time evolution is generated by a random Hamiltonian drawn from the Gaussian unitary ensemble (GUE). The Hamiltonian can be written as a random U(d) spin model,

$$H_t = \sum_{\langle ij \rangle} J_{t,ij}^{ab} T_i^a T_j^b, \tag{2.74}$$

where T_i^a (for $a = 1, 2, \cdots, d^2$) are U(d) generators on site i with $\operatorname{Tr} T_i^{a\dagger} T_i^b = \delta^{ab}$. The coupling $J_{t,ij}^{ab}$ are independently drawn for each time t and indices i, j, a, b from the Gaussian distribution

with zero mean and d^{-2} variance. The quantum dynamics is described by $U = \prod_t e^{-i\varepsilon H_t}$. The operator growth dynamics and the spectral form factor of the Brownian random circuit has been investigated in Ref. [41, 69, 224, 234] recently, where differential equations governing the evolution of operator weight distribution were derived. Our approach also applies to the Brownian circuit model and results in similar differential equations for the evolution of EF state, whose EF Hamiltonian reads (see Appendix A.7 for derivation)

$$\hat{H}_{\text{EF}} = \frac{2(d^2 - 1)}{d^2} \sum_{\langle ij \rangle} \frac{1 - Z_i Z_j}{2} e^{-\delta(X_i + X_j)}. \tag{2.75}$$

We will not discuss this model in further details, given the extensive study of Brownian circuits in the literature. Instead, we will consider a new type of locally scrambled Hamiltonian dynamics.

We start with a fixed Hamiltonian on the one dimensional chain of qudits

$$H = -\sum_{\langle ij \rangle} T_i T_j, \tag{2.76}$$

where T_i is one particular traceless Hermitian operator on site i that squares to identity (i.e. $\text{Tr}\, T_i = 0$ and $T_i^2 = 1$). For the qubit case ($d = 2$), Eq. (2.76) reduces to an Ising model. Note that there is no randomness in the Hamiltonian H. The randomness will be introduced by the local scramblers, when we use H to generate the locally scrambled Hamiltonian dynamics following Eq. (2.71). The entanglement dynamics will be described by the following EF Hamiltonian

$$\hat{H}_{\text{EF}} = \frac{2d^2}{d^2 - 1} \sum_{\langle ij \rangle} \frac{1 - Z_i Z_j}{2} e^{-\delta(X_i + X_j)}, \tag{2.77}$$

which takes the same form as Eq. (2.75) but with a different parameter g. We can test the EF approach with numerical simulation on a 12-qubit system with the choice of $\varepsilon = 0.01$. We start with a product state $|W_{\Psi_0}\rangle = |W_{\text{prod}}\rangle$, evolve the EF state by Eq. (2.31) and calculate the EE from Eq. (2.27). The result is shown in Fig. 2.13. We can see that the averaged EE

obtained from numerics matches well with the result of the EF approach over difference choices of the entanglement regions. These numerical evidences suggest that exchanging the order between taking ensemble average and taking logarithm does not seems to matter much, so the evolution equation we established for the EF in this work can provide reliable descriptions for the entanglement dynamics under locally scrambled quantum dynamics. Comparing Fig. 2.13 with Fig. 2.9, one can see that the entanglement dynamics of the locally scrambled Hamiltonian dynamics closely resembles that of the Haar random unitary circuit. Thus the former can be considered as a continuum limit of the later.
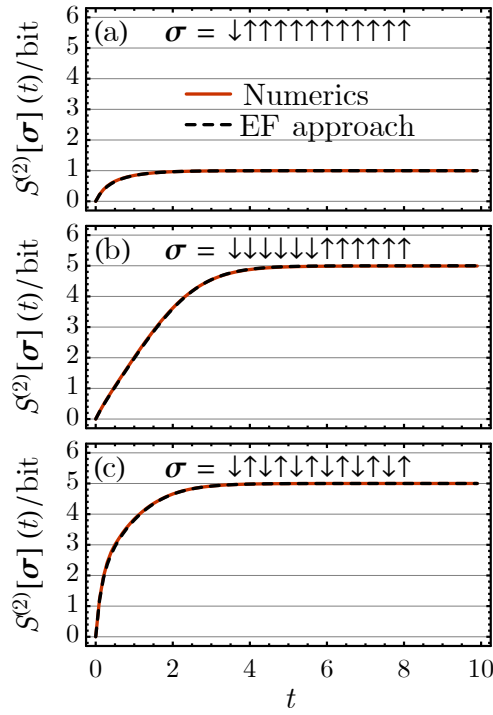
Figure 2.13. The finial state EE of the locally scrambled Hamiltonian dynamics on a 12-site system for different choices of the entanglement regions: (a) single site, (b) half-system, (c) alternating[97]. The qudit dimension is $d = 2$ and the entropy is measured in unit of bit ($= \log 2$).

We also notice that, in agreement with the imaginary time EF Schrödinger equation, the EE always approaches to its final thermalized value exponentially with the same relaxation time

τ independent of the choice of the entanglement region,

$$S^{(2)}[\boldsymbol{\sigma}](t) \to S^{(2)}[\boldsymbol{\sigma}](\infty) - A[\boldsymbol{\sigma}]e^{-t/\tau}. \tag{2.78}$$

The relaxation time τ is intrinsically related to the excitation gap Δ of the EF Hamiltonian $\hat{H}_{\mathrm{EF}}$, which can be estimated by Eq. (2.53) in the thermodynamic limit,

$$\tau^{-1} = \Delta = 2g\left(1 - \frac{1}{d}\right)^2 = \frac{4}{3}, \tag{2.79}$$

where the coupling g, according to Eq. (2.77), is given by $g = 2d^2/(d^2 - 1) = 8/3$ for qubits ($d = 2$). To check this prediction, we fit the numerical simulation data using Eq. (2.78) in the late-time regime to extract the excitation gap Δ. As shown in Fig. 2.14, the EE approaches to the thermal value with the same rate (within error bars) regardless of the different choice of entanglement regions. The numerically fitted gap is around $\Delta = 1.48$, which is close to the thermodynamic-limit analytic prediction $\Delta = 4/3 = 1.33$. The small discrepancy mainly arises from the finite-size effect. If we use the finite-size gap formula $\Delta = 0.56g$ based on the ED result in Fig. 2.6 at the system size $L = 12$, we will obtain a better prediction of the gap $\Delta = 1.49$, which matches the simulation result perfectly.

2.4 Summary and Discussions

In this work, we introduced the concept of locally scrambled quantum dynamics, where each step of the unitary evolution is randomized by local scramblers (on-site Haar random unitary gates). Surrounding each unitary gate in a quantum circuit by local scramblers effectively blocks the local-basis-specific quantum information from propagating in the circuit and decouples the gates from each other under ensemble average. In this way, the average EF of the entire circuit can be constructed piece-by-piece from the EF of each gate, which makes the entanglement dynamics Markovian and enables us to write down the evolution equation for the EF of quantum
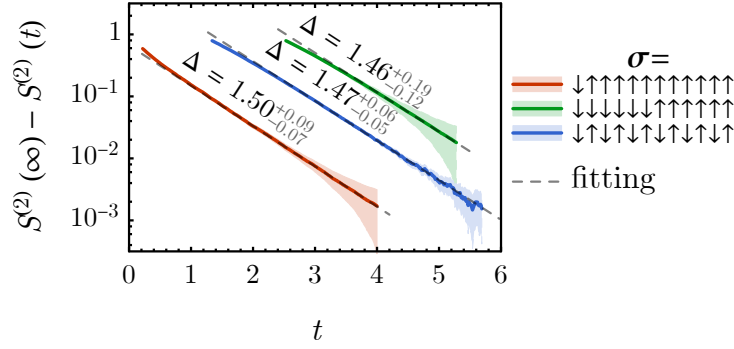
Figure 2.14. The difference between the EE and its final saturation value, i.e. $A[\boldsymbol{\sigma}]e^{-t/\tau}$, plot in the logarithmic scale vs time t. Different colors correspond to different choices of entanglement region (labeled by $\boldsymbol{\sigma}$). The shaded region indicates the error interval. The excitation gap $\Delta = \tau^{-1}$ is extracted by fitting the decay rate.

states. The framework provides us the freedom to design the EF for each gate, such that we can go beyond the conventional Haar random gates and build the random circuit with more general random gates as long as their ensemble is local-basis-independent. This enables us to define and explore the continuum limit of locally scrambled quantum dynamics, under which the evolution of the EF state will be governed by an EF Hamiltonian. We obtained the general form of the EF Hamiltonian on symmetry ground and discussed the implication of its spectral properties on the entanglement dynamics. When the EF Hamiltonian is *gapped*, the excitation modes in the EF state will decay exponentially in time $W_{|\Psi\rangle} = e^{-S} \sim e^{-t/\tau}$, which corresponds to a linear growth of EE in time, i.e. $S \sim t/\tau$, as the system thermalizes. What has not been much discussed previously is the possibility that the EF Hamiltonian can become *gapless* under fine-tuning, then the EF will decay in a power-law manner $W_{|\Psi\rangle} = e^{-S} \sim t^{-\alpha}$, which corresponds to a logarithmic growth of EE, i.e. $S \sim \alpha \log t$. Such scenario could happen at entanglement transitions,[13, 111, 207] where the entanglement scaling of the long-time final state switches from volume-law to area-law. The transition can be interpreted as an order to disorder phase transition of the EF Hamiltonian. One interesting future direction is to explore different models of the EF Hamiltonian and to map out the phase diagram using analytical and numerical tools developed in quantum many-body physics.

Although we focused on the entanglement dynamics of pure states in this work, the EF formulation can be easily generalized to describe mixed state or operator entanglement. Following Eq. (2.4), it is straight forward to define the EF $W_O[\boldsymbol{\sigma}]$ for any many-body operator O (including the density matrix ρ as a special case),

$$W_O[\boldsymbol{\sigma}] = \text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} O^{\otimes 2}), \tag{2.80}$$

and quantize the EF to a state $|W_O\rangle = \sum_{[\boldsymbol{\sigma}]} W_O[\boldsymbol{\sigma}] |\boldsymbol{\sigma}\rangle$. Suppose the operator evolves in time under a locally scrambled quantum dynamics $O' = UOU^\dagger$, the average EF will still be described by the same set of equation $|W_{O'}\rangle = \hat{W}_U \hat{W}_{\mathbb{1}}^{-1} |W_O\rangle$ as Eq. (2.19). Based on this, all the dynamic equation that we developed in this work applies directly, such that we do not need to derive a new set of equations for operator dynamics. The EF state $|W_O\rangle$ encodes the operator EE[237] over all possible regions, which can be used to construct various quantities characterizing the operator size. To name a few, let us first assume O to be a traceless Hermitian operator normalized to $\text{Tr}\, O^2 = d^L$. We can decompose the operator $O = \sum_{[a]} O_{[a]} T^{[a]}$ in the operator basis $T^{[a]} = \prod_i T_i^{a_i}$ (where T_i^a denotes the SU(d) generator on the ith qudit), and define the operator weight $p_{[a]} = O_{[a]}^2$.[117, 158, 224, 234] The fraction of the operator in a subsystem A then reads $p_A = \sum_{[a] \in A} p_{[a]}$,[3] which can be extracted from the EF state $|W_O\rangle$ by taking its inner product with a state $|P_A\rangle$ that labels the subsystem A:

$$p_A = \langle P_A | W_O \rangle, \quad |P_A\rangle = \frac{1}{d^{2L}} \prod_{i \in A} (dX_i - 1) |\Uparrow\rangle. \tag{2.81}$$

The fraction p_A can be further used to characterize the average operator size $\ell_O = \sum_A p_A |A|$. The evolution equation for p_A under Brownian dynamics is recently discussed in Ref. [224, 234]. Another way to probe $|W_O\rangle$ is to consider the variance of the expectation value of O on random mixed states ρ, i.e. $\text{var}_\rho \langle O \rangle_\rho = \mathbb{E}_\rho (\text{Tr}\, \rho O)^2$. Suppose ρ is drawn from a local-basis-independent

---

[3]The notation $\sum_{[a] \in A} = \prod_{i \in A} \sum_{a_i=1}^{d^2-1}$ denotes the summation over the operator configuration $[a]$ which is non-trivial in region A.

ensemble characterized by its EF state $|W_\rho\rangle$, then the variance of operator expectation value is given by

$$\mathrm{var}_\rho \langle O \rangle_\rho = \langle W_\rho | \hat{W}_{\mathbb{1}}^{-1} | W_O \rangle. \tag{2.82}$$

It was recently pointed out by Ref. [185] that $\mathrm{var}_\rho \langle O \rangle_\rho = \sum_A p_A (d+1)^{-|A|}$ can be expressed in terms of p_A, if ρ is uniformly sampled from the ensemble of pure product states. The fact that $\mathrm{var}_\rho \langle O \rangle_\rho$ and p_A are related to each other is less surprising in the EF formulation, because they are simply two different ways to probe the same EF state $|W_O\rangle$. The evolution equation of $|W_O\rangle$ under locally scrambled quantum dynamics is identical to that of $|W_\Psi\rangle$, from which the evolution equations of p_A, ℓ_O or $\mathrm{var}_\rho \langle O \rangle_\rho$ follow automatically. In this way, the EF formulation developed in our work provides a unified framework to discuss various aspects of the operator dynamics.

Another immediate generalization of the framework is to extend the unitary evolution to generic quantum channels allowing measurements to take place. The recent observation of measurement-induced entanglement transition in random unitary circuits [38, 138, 199] has attracted much research interest.[13, 44, 111, 140, 203, 205] In these models, the quantum circuit is doped with local measurements (which can be either weak measurements or projective measurements happened with probability), and the final state EE is studied conditioned on the measurement outcome. If each measurements basis is randomly chosen each time, or if the local measurement take place only after the local basis has been sufficiently scrambled by the unitary evolution, the whole quantum channel still falls in the scope of locally scrambled quantum dynamics, which can be described by the EF approach developed in this work. In this case, each measurement, described by the Kraus operator M, is also a local-basis-independent component in the quantum circuit, and has its own EF similar to Eq. (2.5)

$$W_M[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} M^{\otimes 2} \mathscr{X}_{\boldsymbol{\tau}} M^{\dagger \otimes 2} \right), \tag{2.83}$$

from which the EF operator

$$\hat{W}_M = \sum_{\boldsymbol{\sigma},\boldsymbol{\tau}} |\boldsymbol{\sigma}\rangle W_M[\boldsymbol{\sigma},\boldsymbol{\tau}]\langle\boldsymbol{\tau}| \tag{2.84}$$

can be constructed. The EF state will evolve under measurement by $|W_{\Psi'}\rangle = \hat{W}_M \hat{W}_{\mathbb{1}}^{-1}|W_\Psi\rangle$, such that the approaches developed in this work seamlessly apply. The EF provides a finer resolution of the entanglement structure of a quantum many-body state beyond the single region scaling of EE, which turns out to be useful in diagnosing the error correction capacity[44] in the volume-law states prepared by the measurement-doped quantum circuits. We will leave this topic to future works[61]. More generally, the EF formulation can be further generalized to locally scrambled tensor networks, which does not even need to have a preferential time direction. As long as each tensor in the tensor network is independently drawn from local-basis-independent ensembles, the entanglement structured of the random tensor network can be described by the EF approach. The freedom to design the EF for each separate tensor in the tensor networks opens up a large space of models to explore in the future.

There are also a few more challenging future directions that worth further investigation. The first direction is to generalize the 2nd Rényi EF to arbitrary Rényi index. As a consequence, the Ising variable on each site will be promoted to a permutation group element $\sigma_i \in S_n$. Such generalization will also allow us to access other measures of entanglement, such as Rényi negativity[124, 146, 223], as the moment of the partial transposed density matrix ρ^{T_A}[7–9, 31, 32, 47] can be expressed in terms of the nth Rényi EF,

$$\mathrm{Tr}(\rho^{\mathsf{T}_A})^n = W_\rho[\boldsymbol{g}], \quad g_i = \begin{cases} (n\cdots 21) & i \in A, \\ (12\cdots n) & i \in \bar{A}. \end{cases} \tag{2.85}$$

The nth Rényi generalization of EF states $|W_\Psi\rangle$ and EF operators $\hat{W}_U$ can still be defined, but it will be more difficult to perform explicit calculations given that the number of group elements $n!$ grows quickly with n. Perhaps the most subtle issue is how to take the $n \to 1$ replica limit systematically, which has been identified[13, 111, 207] as an important step to understand the

nature of entanglement transitions. The second direction is to include global symmetries and conservation laws[117, 173] into the discussion. This amounts to refining the generic local scramblers to symmetry-preserving local scramblers, which only performs basis transformations within each irreducible representations of the symmetry group. The formulation to describe the interference between the entanglement dynamics and the flow of symmetry representations in the quantum circuits still need to be developed. The third direction is to go beyond the locally scrambled quantum dynamics and to gradually introduce correlations among random gates in the spacetime. Can the current EF formulation serves as a good starting point to construct phenomenological descriptions for weakly correlated random gates? Can we eventually approach the limit of coherent quantum evolution for Hamiltonian or Floquet dynamics? There are many interesting open question awaiting us to explore.

<div style="text-align:center">— ◆ —</div>

Chapter 2, in full, is a reprint of the material as it appears in Markovian entanglement dynamics under locally scrambled quantum evolution 2020. Kuo, Wei-Ting; Akhtar, Ahmed; You, Yi-Zhuang; Arovas, Dan, Physical Review B, 2020. The dissertation author was the second author of this paper.

Chapter 3

A Two-Parameter Ansatz For Entanglement Feature States

3.1 Introduction

The entanglement dynamics in quantum many-body systems has attracted much research interest[19, 34, 36, 39, 94, 114, 115, 117, 119, 148, 149, 153, 155, 158, 187, 189, 213, 214, 237, 239, 245]. For an isolated quantum many-body system described by a pure state $|\Psi\rangle$, the bipartite quantum entanglement can be quantified by the entanglement entropy $S_{\Psi}(A)$, which characterizes the amount of entanglement between a region (subsystem) A and its complement. While much progress has been made in understanding the entanglement entropy growth [114, 149, 155, 189, 214], most work is restricted to studying single-region entanglement, namely, when A is a single contiguous region. However, very little is known about *multi-region* entanglement, where the entangling region A can consist of several disjoint subregions. The multi-region entanglement has been used to reveal distinctive entanglement structures within volume-law states[62, 212]. In this work, we will focus on the multi-region entanglement and analyze its effects on the entanglement dynamics.

Consider a system with N qudits, where each qudit corresponds to a d-dimensional Hilbert space. Then the total number of choices for the entangling region A is 2^N. To organize all the 2^N corresponding entanglement entropies systematically, Ref. [128, 228] introduced the *entanglement feature* state as a fictitious many-body state $|W_{\Psi}\rangle = \sum_A W_{\Psi}(A)|A\rangle$ that stores the

entanglement features $W_\Psi(A) \equiv e^{-S_\Psi(A)}$ as coefficients of the state vector. In this way, all regions (both single- and multi-regions) are treated on equal footing. The entanglement feature state $|W_\Psi\rangle$ characterizes all the bipartite entanglement of the physical state $|\Psi\rangle$. Ref. [62] suggested that the entanglement feature state $|W_\Psi\rangle$ can be compressed in terms of the matrix product state (MPS) [170, 194, 209], which allows us to use much less parameters (polynomial in N) to approximately parametrize exponentially many entanglement entropies in all 2^N regions. In this work, we further develop this idea and propose a two-parameter MPS ansatz for $|W_\Psi\rangle$, which can capture both the area-law and volume-law scaling for the single-region entanglement, while providing a systematic modeling for the multi-region entanglement at the same time. We interpret the physical meaning of the MPS parameters and use them to define a two-dimensional phase space for different entanglement structures.

The entanglement feature formalism provides a unified approach to study the entanglement dynamics in a large class of models. As proven in Ref. [128], the time evolution of the entanglement feature state is Markovian for any locally-scrambled quantum dynamics, which allows us to predict the entanglement dynamics for all regions at once by solving an (imaginary-time) Schrödinger equation $-\partial_t |W_\Psi\rangle = H_{\text{EF}}|W_\Psi\rangle$ (or its discrete version). Many of the quantum dynamics studied in the literature are locally-scrambled, including random unitary circuits[153, 155, 214, 239], random Hamiltonian dynamics[213, 228], and quantum Brownian dynamics[69, 130, 224, 234]. They all share the locally-scrambled property that every step of the time-evolution is drawn from a random unitary ensemble which is invariant under local (on-site) basis transformations (as if the locally basis are separately scrambled in each step). Using the entanglement feature formalism, we can explore how the corresponding MPS parameter of $|W_\Psi\rangle$ evolves, as the quantum system thermalizes from an initial product state. We can also study the effect of multi-region entanglement on the entanglement dynamics. We found that the dynamics of single-region and multi-region entanglement are generally coupled together. The only known exception is the random unitary circuit, where the dynamical equation is closed within the single-region sector. For generic locally-scrambled quantum dynamics, we derived

the multi-region correction to the single-region entanglement dynamics.

Within the entanglement feature formalism, we can also study the operator spreading[73, 114, 117, 149, 158, 163, 175, 185, 214, 234, 237], which is closely related to the operator entanglement of the unitary time-evolution operator. One well-studied measure of the operator spreading is the out-of-time ordered correlator (OTOC) [64, 67, 96, 137, 181, 187, 214]. We calculate the operator-averaged OTOC for locally-scrambled quantum dynamics using the entanglement feature Hamiltonian. Then from the asymptotic behavior of the OTOC, we determine the butterfly velocity v_B, which characterizes the effective Leib-Robinson velocity and the rate of operator entanglement growth[53, 90, 96, 116, 224]. On the other hand, we can also calculate the entanglement velocity v_E from the entanglement dynamics at different volume-law coefficient s. By comparing v_B and v_E, we check and confirm the previous conjectures about the velocity bounds $v_E \leq (\ln d - |s|)v_B$, as proposed in Ref. [53, 114].

This paper is organized as follows: in 3.2, we outline the *excitation spectrum* of the multi-region entanglement continuum, such as its boundaries and the gap between single-region and multi-region entanglements. This gap Δ is finite in the thermodynamic limit and contributes to the single-region entanglement dynamics. In this section, we also propose our MPS ansatz for describing the entanglement of generic locally scrambled states as they evolve from product state, through an area law phase eventually thermalizing to a volume law configuration. Despite that the *physical* state's entanglement rapidly grows as it evolves, the corresponding entanglement feature state, which describes the full single and multi-region entanglement structure, is well described by our $D = 2$ MPS ansatz. The ansatz enables to calculate exactly many features of this multi-region continuum, including its boundaries and gap, which then illuminates how to interpret our MPS parameters α, θ, whose dynamics are explored in the following section. In 3.3, we briefly summarize the meaning of locally scrambled quantum dynamics and the EF Hamiltonian ansatz parameters g, β. To study Haar random and swap circuits using an EF Hamiltonian, we define a continuum limit for these models by analyzing their *entanglement cut* (domain wall) dynamics in the EF Hilbert space. Having defined a continuum limit for these models, we

may then focus exclusively on entanglement dynamics generated by the EF Hamiltonian to study locally scrambled circuits. We find that entanglement velocity, and therefore single-region entanglement dynamics, depend on Δ which is unique to the multi-region continuum, and can in fact dominate the dynamics. In 3.4, we derive the butterfly velocity from analyzing the infinite temperature operator averaged OTOC, and show large system size numerics generated by large-D MPS that agree with our result. Lastly, we compare our butterfly velocity with known bounds on entanglement velocity proposed in Ref. [53, 114], and find that they agree.

3.2 Matrix Product State Ansatz for Multi-Region Entanglement

3.2.1 Multi-Region Entanglement

Quantum many-body system can exhibit rich and complex entanglement structures. We consider an isolated many-body system described by a pure state $|\Psi\rangle$, and focus on the bipartite entanglement quantified by the von Neumann or Rényi entanglement entropies. In this work, we will restrict to the 2nd Rényi entanglement entropy, which will admit the simplest formulation in the entanglement feature approach. Given the many-body state $|\Psi\rangle$, the 2nd Rényi entanglement entropy over a region A is defined as

$$S_\Psi(A) = -\ln \mathrm{Tr}_A \, \rho_A^2, \tag{3.1}$$

where $\rho_A = \mathrm{Tr}_{\bar{A}} |\Psi\rangle\langle\Psi|$ is the reduced density matrix of the subsystem A. $S_\Psi(A)$ characterize the amount of entanglement between the entangling region A and its complement $\bar{A}$. Suppose the many-body system is a chain of qudits arranged along a one-dimensional lattice. The region A can be taken to be any subset of the qudits, and does not need to be a continuous segment. However, most current studies of the entanglement dynamics have focused on the growth of entanglement entropy in a single region or the half system. What about the multi-region entanglement entropies? How do they evolve under non-equilibrium quantum dynamics? Do

they mutually affect each other during the evolution? These are the problems that we will explore in this work.
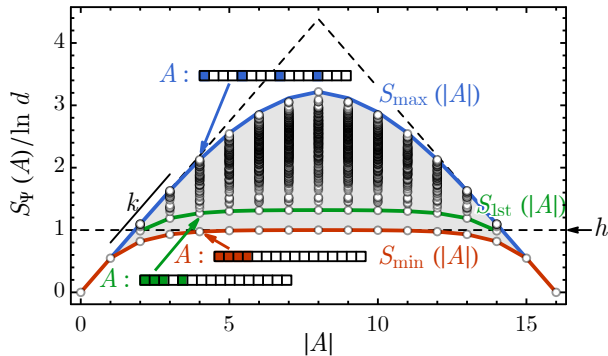
Figure 3.1. Entanglement entropy $S_\Psi(A)$ by the region size $|A|$ for a typical many-body state on a one-dimensional lattice of $N = 16$ sites. The $S_\Psi(A)$ data are produced by the $D = 2$ MPS model at $(\alpha, \theta) = (\frac{1}{2}, \frac{\pi}{4})$ at qudit dimension $d = 2$.

To study these problems, we first need to organize the multi-region entanglement entropy $S_\Psi(A)$ systematically. If A is a single-region, we can parameterize A by its region size $|A|$ (i.e. the number of qudits in A), because the entanglement entropy $S_\Psi(A)$ will only depend on $|A|$ in the presence of translation symmetry. However for multi-regions, the region size $|A|$ is insufficient to parametrize the region A, since $S_\Psi(A)$ will also depend on how A is segmented along the one-dimensional chain. Fig. 3.1 displays the scatter plot of the entanglement entropy $S_\Psi(A)$ with respect to the region size $|A|$ for a typical quantum many-body state. Similar phenomenology derived from a random-matrix framework is discussed in [204]. Data points of $S_\Psi(A)$ distributes in a dome-shaped continuum bounded between the lower edge $S_{\min}(|A|)$ (in red) and the upper edge $S_{\max}(|A|)$ (in blue). Given the region size $|A|$, the different values of $S_\Psi(A)$ originates from different segmentations of the entangling region A. Because $S_\Psi(A)$ always increases with the number of entanglement cuts in A when $|A|$ is fixed, the lower-bound $S_{\min}(|A|)$ should be given by the single-region entanglement entropy which has the least number of entanglement cuts, while the upper-bound $S_{\max}(|A|)$ should correspond to the entangling region A being disjoint sites separated from each other as far as possible which has the most number of entanglement

cuts. All the multi-region entanglement entropies are distributed between the curves $S_{\min}(|A|)$ and $S_{\max}(|A|)$, forming a dense continuum in the thermodynamic limit $N \to \infty$, which might be dubbed as the *multi-region continuum*.

If we imagine the entanglement entropy $S_\Psi(A)$ as a kind of "energy" associated with each region A, Fig. 3.1 can be viewed as an "excitation spectrum" of entangling regions. It is the whole spectrum that fully characterize the entanglement structure of the underlying quantum many-body state $|\Psi\rangle$. $S_{\min}(|A|)$ describes how the "ground state energy" varies with the region size $|A|$, which is mostly discussed in the literature. But we are also curious about the "excited states" in the spectrum. For example, $S_{1\text{st}}(|A|)$ describes how the "1st excited state energy" varies with $|A|$, as the green curve in Fig. 3.1. The entangling region A that contributes to $S_{1\text{st}}(|A|)$ always contains two subregions of sizes $(|A| - 1)$ and 1, separated by one site in between, which provides an example of the multi-region entanglement. One may further define the "excitation gap" between $S_{1\text{st}}(|A|)$ and $S_{\min}(|A|)$ as

$$\Delta(|A|) = S_{1\text{st}}(|A|) - S_{\min}(|A|). \tag{3.2}$$

Ref. [62] found that the *entropy gap* $\Delta(|A|)$ plays an important role in quantifying the error-correcting capacity in sub-thermal volume-law states. This motivates us to further investigate the multi-region entanglement.

However, it is still challenging to organize all the 2^N entropies. To meet this challenge, we took the "entropy-energy correspondence" farther to define the "Boltzmann weight" for each region A

$$W_\Psi(A) = e^{-S_\Psi(A)}, \tag{3.3}$$

which was first introduced as the *entanglement feature* in Ref. [228, 230] (see Chapter 2). In the case of 2nd Rényi entropy, the entanglement feature $W_\Psi(A) = \text{Tr}_A \rho_A^2$ is simply the purity of subsystem A. If we attempt to arrange the entanglement features as components of a single

vector, the number of components (2^N) will be the same as that of the state vector of a N-spin system. This motivates us to organize the 2^N entanglement features into a fictitious many-body state, called the entanglement feature state [128],

$$|W_\Psi\rangle = \sum_{[\sigma]} W_\Psi[\sigma]|[\sigma]\rangle, \tag{3.4}$$

where the 2^N different entangling regions A are equivalently represented as the 2^N Ising configurations $[\sigma] = (\sigma_1, \sigma_2, \cdots, \sigma_N)$ with one-to-one correspondence

$$\sigma_i = \begin{cases} +1 \ (\uparrow) & i \notin A, \\ -1 \ (\downarrow) & i \in A. \end{cases} \tag{3.5}$$

In this way, all regions (no matter single-regions or multi-regions) are treated on equal footing. The entanglement entropy of any region A can be taken back from the entanglement feature state $S_\Psi(A) = -\ln\langle A|W_\Psi\rangle$, where $|A\rangle$ is the Ising basis state with down-spins in region A and up-spins in the complement of A.

3.2.2 Matrix Product State Representation

At this point, Eq. (3.4) does not seem to really simplify the problem, other than encoding the 2^N entanglements into a many-body state. However, an important observation is that the entanglement feature $W_\Psi[\sigma]$ is always positive. According to Ref. [80], many-body states with positive wavefunctions in a local basis should typically exhibit a constant-law scaling of entanglement entropy, which is only violated in fine-tune cases. Since entanglement feature states $|W_\Psi\rangle$ always have positive components, they should have low entanglement and should admit efficient matrix product state (MPS) [209] representations. Here we would like clarify that the entanglement feature state $|W_\Psi\rangle$ was originally introduced to describe the entanglement property of the physical quantum state $|\Psi\rangle$, but $|W_\Psi\rangle$ itself as a many-body state also has its own entanglement properties. Given the sign-free nature of $|W_\Psi\rangle$, the entanglement of $|W_\Psi\rangle$ should

79

typically follow a constant-law, regardless of the entanglement properties of corresponding physical state $|\Psi\rangle$, although we have not been able to strictly prove that $|W_\Psi\rangle$ must be constant-law entangled in every cases. Nevertheless, as we will see, even if the underlying physical state $|\Psi\rangle$ is a maximally entangled Page state (i.e. a random state in the many-body Hilbert space), the corresponding entanglement feature state $|W_{\text{Page}}\rangle$ still remains constant-law entangled and can be precisely written as a MPS with bond dimension $D = 2$. Among all currently known examples, the $|W_\Psi\rangle$ will be most entangled at the entanglement transition [13, 44, 111, 138, 140, 199, 202, 207, 223], where it exhibits a logarithmic-law entanglement. Since it requires fine-tuning to hit the entanglement transition, we argue that for most cases, the entanglement feature state should be MPS-representable.

Therefore, assuming translation symmetry, we propose the following MPS ansatz for the entanglement feature state $|W_\Psi\rangle$

$$W_\Psi[\sigma] \propto \text{Tr}\left(\prod_i M_\Psi^{\sigma_i} \right), \tag{3.6}$$

where $M_\Psi^{\sigma_i}$ is a $D \times D$ matrix that depends on $\sigma_i = \pm 1$. We take the periodic boundary condition along the chain, such that the product of matrices can be simply traced over without specifying additional boundary conditions. The right-hand-side of Eq. (3.6) is not properly normalized yet. The normalization constant should be determined by $\langle\emptyset|W_\Psi\rangle = 1$, because the entanglement entropy $S_\Psi(\emptyset) = 0$ for empty region ($A = \emptyset$) must be zero. Based on the MPS representation, we can apply efficient numerical algorithms, such as the time-evolving block decimation (TEBD), to simulate the entanglement dynamics for large systems. Such numerical approach has been explored in Ref. [62] recently.

Here we would like to further investigate along the analytic direction. We will construct the minimal MPS model for the entanglement feature state to capture the major features of the multi-region entanglement in Fig. 3.1. In particular, we will consider the MPS ansatz with bond dimension $D = 2$. We could also consider a larger bond dimension for a stronger representation

power, but for the purpose of analytical treatment here, we would like to keep the bond dimension as small as possible, such that we can possibly interpret the MPS parameters in the end. The efficacy of the $D = 2$ MPS is numerically verified in Appendix B.1, which shows that $D = 2$ MPS is already successful in capturing all the multi-region entanglement over the entire thermalization process.

For $D = 2$, M_Ψ^σ will be a σ-dependent 2×2 matrix of the following form

$$M_{(\alpha,\theta)}^\sigma = \cosh \alpha \, I + \sinh \alpha (\sin \theta \, X + \sigma \cos \theta \, Z), \qquad (3.7)$$

where I, X, Z denote the identity, Pauli-x, and Pauli-z matrices respectively. The ansatz is only controlled by two real parameters $\alpha \geq 0$ and $0 \leq \theta \leq \pi/2$. The form in Eq. (3.7) can be determined based on the following considerations:

(i) For pure state $|\Psi\rangle$, the entanglement entropy in region A should be the same as that in the complement region $\bar{A}$, i.e. $S_\Psi(A) = S_\Psi(\bar{A})$. This implies $W_\Psi[\sigma] = W_\Psi[-\sigma]$ for the entanglement feature, i.e. the entanglement feature state $|W_\Psi\rangle$ must respect the $\mathbb{Z}_2$ symmetry ($\sigma \to -\sigma$). But before imposing the $\mathbb{Z}_2$ symmetry on the MPS ansatz, we notice that the matrix M_Ψ^σ carries a gauge freedom, since $|W_\Psi\rangle$ is invariant under the following gauge transformation

$$M_\Psi^\sigma \to G M_\Psi^\sigma G^{-1}, \qquad (3.8)$$

induced by any $G \in \mathrm{GL}(2, \mathbb{C})$. Therefore the $\mathbb{Z}_2$ symmetry action on M_Ψ^σ will be followed by a corresponding gauge transformation in general [126, 180, 193]. We can choose the gauge transformation to be $G = X$ [62], then the $\mathbb{Z}_2$ symmetry requires $M_\Psi^\sigma = X M_\Psi^{-\sigma} X$, which can be resolved by $M_\Psi^\sigma = (c_0 I + c_1 X) + \sigma(i c_2 Y + c_3 Z)$. The coefficients $c_{0,1,2,3} \in \mathbb{R}$ should all be real to ensure that the resulting entanglement features are real.

(ii) We can always rescale M_Ψ^σ by an overall factor, such as $M_\Psi^\sigma \to c M_\Psi^\sigma$. The factors will be absorbed into the normalization constant in Eq. (3.6), which can always be fixed by

$\langle \emptyset | W_\Psi \rangle = 1$ in the end. So we are free to set $c_0 = 1$ by rescaling.

(iii) We can use the gauge freedom to eliminate c_2 (as a gauge fixing) by performing the gauge transformation of $G = aI + bX$, where a, b can be any solution of $a^2 + b^2 - 2abc_3/c_2 = 0$.

(iv) The remaining parameters c_1 and c_3 can be parametrized by an positive amplitude $c \geq 0$ and an angle θ following $c_3 + ic_1 = c e^{i\theta}$, such that $M_\Psi^\sigma = I + c(\sin\theta X + \sigma\cos\theta Z)$. Using the gauge transformation of $G = Z$ or $G = X$, we can flip the sign of the coefficient in front of X or Z independently. Thus we can make both $\sin\theta$ and $\cos\theta$ positive. So we only need to consider $0 \leq \theta \leq \pi/2$.

(v) Now the eigenvalues of M_Ψ^σ are $1 \pm c$. To ensure that the entanglement features $W_\Psi[\sigma]$ are positive, we must at least require both eigenvalues to be positive. If, for example, $c > 1$ and $\theta = 0$, we can show that the single-site entanglement feature will become negative. Thus we should have $0 \leq c < 1$, so that c can be rewritten as $c = \tanh\alpha$.

Thus we end up with the final form in Eq. (3.7) (up to additional rescaling by $\cosh\alpha$).

3.2.3 Edges of Multi-Region Continuum

The MPS ansatz Eq. (3.7) provides a minimal model for all the 2^N entanglement entropies in terms of two real parameters (α, θ),

$$S_{(\alpha,\theta)}[\sigma] = -\ln \text{Tr}\left(\prod_i M_{(\alpha,\theta)}^{\sigma_i}\right) + S_0, \tag{3.9}$$

where the background entropy $S_0 = \ln(2\cosh\alpha N)$ is attached to ensure the entanglement entropy $S_{(\alpha,\theta)}(\emptyset) = 0$ vanishes for empty region (which corresponds to properly normalize the entanglement feature state). Fig. 3.1 is actually generated by Eq. (3.9) at $(\alpha, \theta) = (\frac{1}{2}, \frac{\pi}{4})$, which demonstrated that the minimal model can describe the multi-region continuum. In particular, we can determine both its lower edge $S_{\min}(|A|)$ and its upper edge $S_{\max}(|A|)$, see Appendix B.2 for

derivation. The lower edge is given by the single-region configuration,

$$
\begin{aligned}
S_{\min}(|A|) &= -\ln \text{Tr}(M^{\downarrow})^{|A|}(M^{\uparrow})^{N-|A|} + S_0 \\
&= -\ln\left(\sin^2\theta + \cos^2\theta \, \tfrac{\cosh\alpha(N-2|A|)}{\cosh\alpha N}\right).
\end{aligned}
\tag{3.10}
$$

The upper edge corresponds to the entangling region A of equally spaced single sites. Given the region size $|A|$, the region A will be a disjoint union of $|A|$ sites separated from their neighbors by $N/|A| - 1$ sites (see the blue inset of Fig. 3.1). Ignoring the subtlety of the possible incommensurability between $|A|$ and N, the upper edge for $|A| < N/2$ should read

$$
\begin{aligned}
S_{\max}(|A|) &= -\ln \text{Tr}\left(M^{\downarrow}(M^{\uparrow})^{N/|A|-1}\right)^{|A|} + S_0 \\
&= -\ln \frac{\cosh\eta|A|}{\cosh\alpha N},
\end{aligned}
\tag{3.11}
$$

where η also depends on $|A|$ and is given by $\cosh\eta = \sin^2\theta \cosh\frac{\alpha N}{|A|} + \cos^2\theta \cosh\frac{\alpha(N-2|A|)}{|A|}$. For $|A| > N/2$, we simply take its $\mathbb{Z}_2$ reflection $S_{\max}(|A|) = S_{\max}(N-|A|)$.

When $N \to \infty$, Eq. (3.10) and Eq. (3.11) reduce to (see Appendix B.2)

$$
\begin{aligned}
S_{\min}(|A|) &= -\ln(\sin^2\theta + \cos^2\theta \, e^{-2\alpha|A|}), \\
S_{\max}(|A|) &= -\ln(\sin^2\theta + \cos^2\theta \, e^{-2\alpha})|A|.
\end{aligned}
\tag{3.12}
$$

Usually when we discuss the area-law v.s. volume-law entanglement, we are talking about the scaling of the lower edge $S_{\min}$ with respect to $|A|$. Here we could also talk about the scaling of the upper edge $S_{\max}$, which, according to Eq. (3.12), is always volume-law (even for the area-law state). Such behavior is rather trivial to understand, as the "area" $|\partial A|$ is proportional to the volume $|A|$ in this case. So let us focus on the lower edge, which generally exhibits an area-law scaling (unless $\theta = 0$), as $S_{\min}$ gradually saturates to $-2\ln\sin\theta$ for large $|A|$. Only when $\theta = 0$, the lower edge also exhibits a volume-law scaling and coincides with the upper edge, i.e. $S_{\min}(|A|) = S_{\max}(|A|) = 2\alpha|A|$, such that the multi-region continuum is squeezed to

vanish in this case. However, we should mention that the vanishing multi-region continuum for all volume-law states is an artifact of the $D = 2$ MPS model. When a non-vanishing multi-region continuum appears on top of the volume-law lower edge $S_{\min}$, it implies that the volume-law state has a non-trivial quantum error correcting capacity, which has been discussed in Ref. [44, 62]. As a minimal model, the $D = 2$ MPS ansatz has too few parameters to capture the error-correcting properties in the volume-law state. If we extend the MPS ansatz to higher bond dimension $D > 2$, the error-correcting volume-law state can be described as well (as has been shown in Ref. [62]).

3.2.4 Physical Meaning of MPS Parameters

To describe the overall shape of the multi-region continuum, we define the following two characteristic quantities (as illustrated in Fig. 3.1): the area-law plateau height of the lower edge $S_{\min}(|A|)$

$$h \equiv \frac{S_{\min}(N/2)}{\ln d} = -\log_d \sin^2 \theta, \tag{3.13}$$

and the volume-law coefficient of the upper edge $S_{\max}(|A|)$

$$
\begin{aligned}
k &\equiv \lim_{|A| \to 0} \frac{\partial_{|A|} S_{\max}(|A|)}{\ln d} \\
&= -\log_d(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}),
\end{aligned}
\tag{3.14}
$$

where $\log_d$ denotes the logarithm of base d (and d is the qudit dimension). The other way round, the MPS parameters (α, θ) can also be expressed in terms of the height h and the slope k as

$$
\begin{aligned}
\alpha &= \frac{1}{2} \ln \frac{d^h - 1}{d^{h-k} - 1}, \\
\theta &= \arcsin d^{-h/2}.
\end{aligned}
\tag{3.15}
$$

Now the physical meaning of the MPS parameters becomes clear. The θ parameter is directly determined by the area-law plateau height h. For product states, $h = 0$, hence $\theta = \pi/2$. For generic area-law states, $h \geq 0$ is positive, hence $0 < \theta \leq \pi/2$. For volume-law state, $h \to \infty$ (the

area-law plateau never appears), hence $\theta = 0$. We summarize these cases in Fig. 3.2(a). Once θ is fixed, the remaining parameter α will be set by the slope k of the upper edge. In particular, for the volume-law state ($\theta = 0$), α is directly related to the volume-law coefficient k by $\alpha = \frac{1}{2}k\ln d$, which is half of the entanglement entropy of a single qudit.
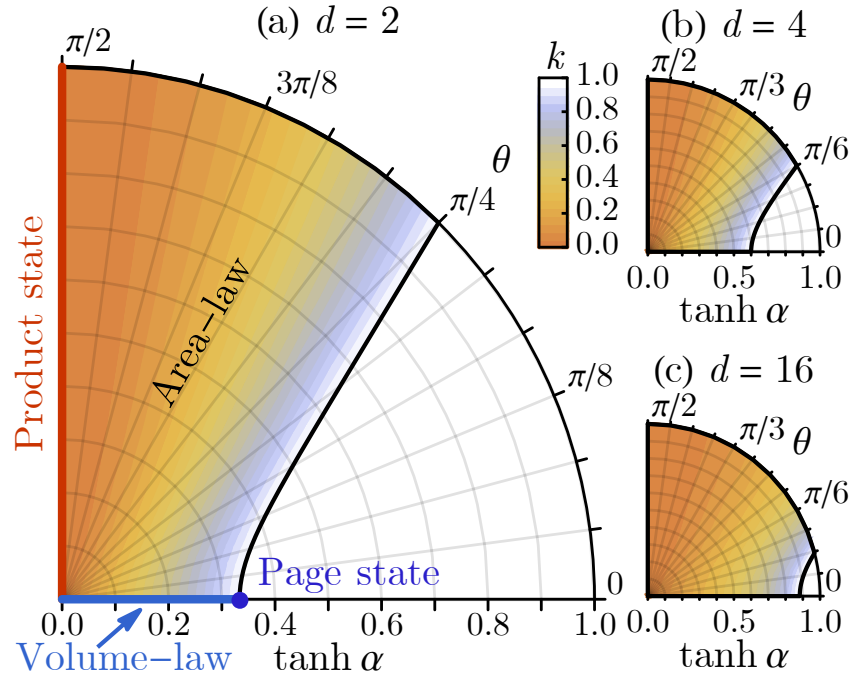
Figure 3.2. Feasible domain of MPS parameters (α, θ) at different qudit dimensions: (a) $d = 2$, (b) $d = 4$, (c) $d = 16$. The background color indicates the upper edge volume-law coefficient k.

As the volume-law coefficient approaches unity ($k = 1$), the system reaches the maximally entangled Page state $(\alpha, \theta) = (\frac{1}{2}\ln d, 0)$, where the MPS ansatz in Eq. (3.7) reduces to $M^\sigma_{\text{Page}} = I + \frac{d-1}{d+1}\sigma Z$, which gives the exact description of the Page state entanglement features $|W_{\text{Page}}\rangle$. Although the underlying physical state $|\Psi\rangle$ is highly entangled, its entanglement features can still be captured by a low-entanglement MPS efficiently, because it generally takes much less entanglement resources to describe the entanglement property of a many-body state $|\Psi\rangle$ than the state itself. Therefore using the MPS approach, we can achieve a huge compression of the many-body entanglement features. Just starting from two characteristic quantities h and k, the

MPS ansatz Eq. (3.9) can provide a comprehensive modeling of entanglement entropies in all the 2^N possible regions (see Fig. 3.1), demonstrating the prediction power of the MPS model. So, within the MPS model, the question of how entanglement entropies evolve in different regions boils down to how MPS parameters (such as α, θ) evolve, which will be investigated in more details soon.

Before discussing the entanglement dynamics, let us first determine the feasible domain of the MPS parameters. A non-trivial constraint comes from the requirement that the volume-law coefficient $k \leq 1$ must not exceed one, because the entanglement entropy of a single qudit can not be greater than $\ln d$. According to Eq. (3.15), $k \leq 1$ implies

$$(d\cos 2\theta + 1)\tanh \alpha \leq d - 1. \tag{3.16}$$

This inequality further restricts the primitive domain of $\alpha \geq 0$ and $0 \leq \theta \leq \pi/2$, leading to the feasible domain shown in Fig. 3.2. The shape of the feasible domain varies with the qudit dimension d, as illustrated in Fig. 3.2(b,c). In the following section, we will study the evolution of the entanglement feature state under locally scrambled quantum dynamics. Using the MPS ansatz developed in this section, we will be able to represent the evolution in the MPS parameter space, which will provide an intuitive picture of how the quantum system thermalizes with time.

3.3 Entanglement Dynamics in Locally Scrambled Quantum Systems

3.3.1 Entanglement Feature Formalism

Let us put aside the MPS model shortly and discuss the dynamics of generic entanglement feature state $|W_\Psi\rangle$. In the entanglement feature formalism, $|W_\Psi\rangle$ captures the entanglement entropies over all regions for a given quantum many-body state $|\Psi\rangle$. As the state $|\Psi\rangle$ evolves in time, so does its corresponding entanglement feature state $|W_\Psi\rangle$. Ref. [128] has proven that the

entanglement dynamics will be governed by an imaginary-time Schrödinger equation

$$-\partial_t |W_\Psi\rangle = H_{\text{EF}} |W_\Psi\rangle, \tag{3.17}$$

if the underlying quantum dynamics is *locally scrambled.* A quantum dynamics $|\Psi\rangle \to U|\Psi\rangle$ is said to be locally scrambled, if each step of the unitary evolution U is drawn from a ensemble whose probability measure $P(U)$ is invariant under local basis transformations, i.e. $P(U) = P(VUV^\dagger)$ for all $V = \prod_i V_i$ with arbitrary $V_i \in \text{U}(d)$ on site-i. Examples of locally scrambled quantum system include random unitary circuits[153, 155, 214, 239], random Hamiltonian dynamics[213, 228], and quantum Brownian dynamics[69, 130, 224, 234]. Because local basis information is fully scrambled at each step of the time-evolution, the entanglement dynamics will be Markovian, which can be described by the entanglement feature Hamiltonian H_{EF} following Eq. (3.17).

As derived in Ref. [128], up to the nearest neighbor coupling on a one-dimensional lattice, H_{EF} should take the following general form (as a many-body spin model)

$$H_{\text{EF}} = g \sum_{\langle ij\rangle} \frac{1 - Z_i Z_j}{2} e^{-\delta(X_i + X_j) - \beta X_i X_j}, \tag{3.18}$$

so as to preserve the $\mathbb{Z}_2$ symmetry and the normalization of $|W_\Psi\rangle$ under the entanglement dynamics and to respect the time-reversal symmetry. Here X_i and Z_i are Pauli operators acting on site-i. The parameter δ is fixed by the qudit dimension d via $\tanh \delta = 1/d$, and H_{EF} is only controlled by two free parameters g and β. The parameter g sets the time-scale and determines how fast the dynamics will happen on the overall scale. The parameter β is tied to the type of the quantum dynamics. A few examples are listed in Tab. 3.1. Unlike conventional spin models, H_{EF} contains a projection operator $\frac{1 - Z_i Z_j}{2}$ on each bond, which imposes dynamic constraints on the motion of domain walls[55, 227]. We will discuss its effect on the entanglement dynamics later.

Table 3.1. Examples of locally scrambled quantum dynamics and their corresponding β value (in terms of $\tanh\beta$).

Type of quantum dynamics	$\tanh\beta$
Quantum Brownian dynamics	0
Random unitary circuit (continuum limit)	$1/d^2$
Fractional swap circuit (continuum limit, $x \to 1$)	1

3.3.2 Locally Scrambled Quantum Dynamics

Let us briefly review some important examples of locally scrambled quantum dynamics, in order to gain some intuition about the parameter β. The first example is the quantum Brownian dynamics (also known as the Brownian random circuit)[130], where the quantum many-body state $|\Psi\rangle$ evolves under a time-dependent random Hamiltonian following $|\Psi\rangle \to e^{-iH_t dt}|\Psi\rangle$, with

$$H_t = \sum_{\langle ij \rangle} J_{t,ij}^{ab} T_i^a T_j^b, \tag{3.19}$$

where T_i^a (for $a = 1, 2, \cdots, d^2$) are U(d) generators on site i with the normalization $\operatorname{Tr} T_i^{a\dagger} T_i^b = \delta^{ab}$, and the couplings $J_{t,ij}^{ab}$ are independently drawn for each time t and indices i, j, a, b from a Gaussian distribution with zero mean and d^{-2} variance. According to Ref. [128], the corresponding entanglement dynamics (the evolution of $|W_\Psi\rangle$) is precisely described by Eq. (3.18), with $g = 2(1 - d^{-2})$ and $\beta = 0$.

The next example is the random unitary circuit[155], depicted in Fig. 3.3(a). The physical quantum state $|\Psi\rangle \to U_t|\Psi\rangle$ evolves by the application of random unitary gates layer-by-layer (U_t denotes the whole layer of gates on an equal-time slice at time t), where each gate is independently drawn from two-qudit Haar random unitary ensemble. As a quantum circuit model, the time is discrete (each layer is a step in time) and the corresponding entanglement
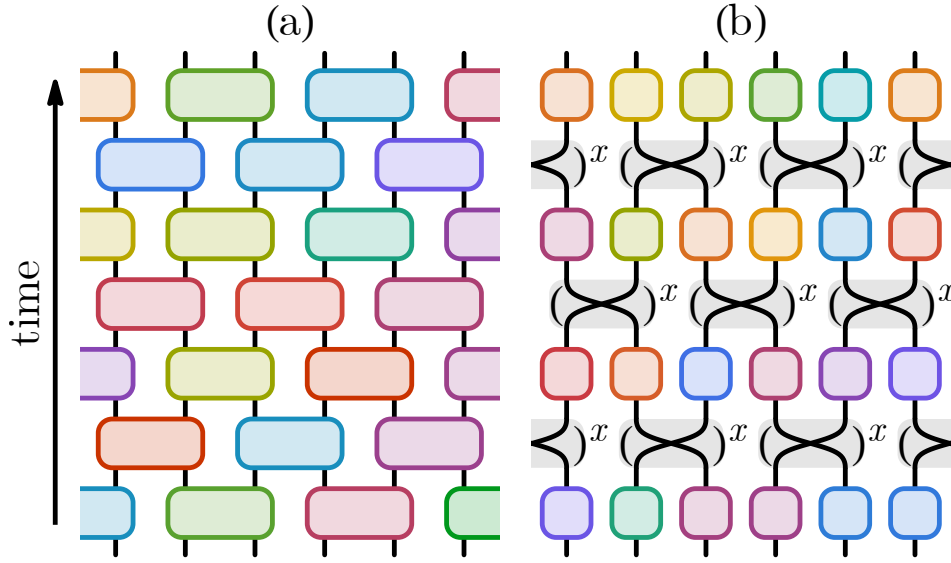
Figure 3.3. (a) Random unitary circuit. (b) Fractional swap circuit. Each color block represents an independent Haar random unitary gate. The fractional swap gate is depicted as swap gate to the fractional power x.

dynamics is described by the following transfer matrix[128]

$$|W_\Psi\rangle \to \prod_{\langle ij\rangle \in \Lambda_\pm} T_{ij}^{\text{Haar}}|W_\Psi\rangle,$$

$$T_{ij}^{\text{Haar}} = 1 - \frac{1-Z_iZ_j}{2}\left(1 - \frac{d}{d^2+1}(X_i + X_j)\right),$$

(3.20)

where Λ_+ (Λ_-) denotes the collection of even (odd) bonds which are chosen alternately follow-

ing the brick-wall pattern of the circuit. Another circuit model is the fractional swap circuit,

as illustrated in Fig. 3.3(b), which was first introduced in Ref. [128]. It is constructed by a

layer of fractional swap gates (i.e. swap gates to the fractional power $x \in [0, 1]$: $\text{SWAP}^x = \frac{1+e^{ix\pi}}{2} + \frac{1-e^{ix\pi}}{2}\text{SWAP}$) followed by a layer of on-site Haar random unitary gates, and so on. The

corresponding entanglement dynamics is described by the following transfer matrix[128]

$$|W_\Psi\rangle \to \prod_{\langle ij\rangle \in \Lambda_\pm} T_{ij}^{\text{SWAP}^x}|W_\Psi\rangle,$$

$$T_{ij}^{\text{SWAP}^x} = 1 - \frac{1-Z_iZ_j}{2}(u - v(X_i + X_j) + wX_iX_j),$$

(3.21)

where $(u,v,w) = (d^2a - b, da - db, a - d^2b)/(d^2 - 1)$ with $a = (2 - \sin^2 \frac{x\pi}{2})\sin^2 \frac{x\pi}{2}$ and $b = \sin^4 \frac{x\pi}{2}$. In the limit of $x \to 1$, the fractional swap gate reduces to the swap gate and $(u,v,w) \to (1,0,-1)$, such that the transfer matrix takes a simpler form

$$T_{ij}^{\mathsf{SWAP}} = 1 - \tfrac{1-Z_iZ_j}{2}(1 - X_iX_j). \tag{3.22}$$

This limit may be called the swap circuit model, which was proposed[128] to mimic the entanglement dynamics in integrable conformal field theories (CFT) where entanglement spreads with the ballistic propagation of quasi-particles[29, 124, 163].

3.3.3 Causal Structure and Continuum Limit

To make connection to the continuous-time entanglement dynamics described by H_{EF}, we would like to take the continuum limit of the transfer matrixes T^{Haar} and T^{SWAP}. In general, it is unclear how to define the continuum limit for discrete circuit models, but for the random unitary circuit and the swap circuit dynamics, we can use their definitive causal structure to pin down the corresponding parameter β in H_{EF}, hence obtaining the continuum version of their entanglement dynamics.
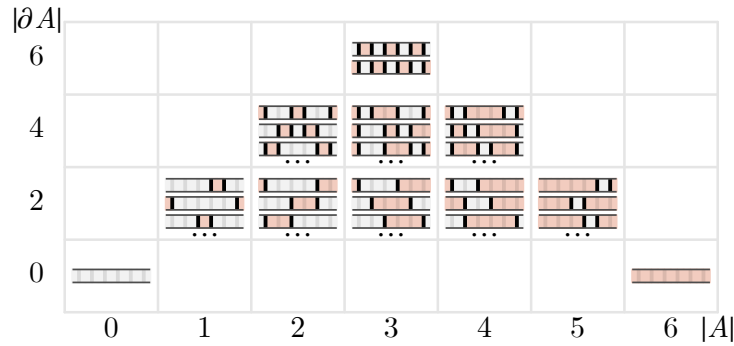
Figure 3.4. Classification of entanglement regions A by their volumes $|A|$ and areas $|\partial A|$ on a $N = 6$ lattice.

To reveal the structure of the entanglement feature Hilbert space spanned by the entanglement region basis states $|A\rangle$, we classify the entanglement region A by its volume $|A|$ (the

number of sites in A) and its area $|\partial A|$ (the number of entanglement cuts, also twice of the number of entanglement regions). In this way, the Hilbert space is partitioned into sectors labeled by $(|A|, |\partial A|)$, as tabulated in Fig. 3.4. We would like to understand how the entanglement feature transfer matrix connects different sectors of the Hilbert space. Since the entanglement dynamics is $\mathbb{Z}_2$ symmetric, we only need to keep track of the $\mathbb{Z}_2$ invariant objects, which are the entanglement cuts (domain walls of Ising spins). Therefore the multi-region entanglement dynamics is essentially a many-body dynamics of entanglement cuts.

Cut dynamics		(a)	(b)	(c)
hopping		✓	✗	v-term
pair creation		✓	✗	v-term
pair annihilation		✗	✗	--
pair hopping		✗	✓	w-term
triple division		✗	✓	w-term
triple fusion		✗	✓	w-term

Figure 3.5. Dynamical constraints on the entanglement cut motion in (a) random unitary circuit T^{Haar}, (b) swap circuit T^{SWAP}. (c) Corresponding terms in H_{EF} that leads to the cut dynamics.

Given the transfer matrix in Eq. (3.20) or Eq. (3.22), the basic moves of entanglement cuts can be read out and summarized in Fig. 3.5. The column (a) in Fig. 3.5 indicates whether a process is allowed (checked) or forbidden (crossed) by the transfer matrix T^{Haar} for the random unitary circuit defined in Eq. (3.20). The entanglement cuts are free to move along the lattice and can be created in pairs. But once created, they can not be annihilated, as the pair annihilation process is forbidden. As a result of this dynamical constraint, the entanglement features in the lower-$|\partial A|$ sectors can only affect those in the higher-$|\partial A|$ sectors (via pair creation), but not the other way round, as shown in Fig. 3.6(a). In particular, the single-region entanglement is not affected by the multi-region entanglement under the entanglement dynamics, which is a unique property of the random unitary circuits.

Similar analysis can be done for the swap circuit, described by the transfer matrix T^{SWAP}

in Eq. (3.22). As shown in Fig. 3.5(b), a single entanglement cut can not move but a pair of them can hop together; moreover, pairs of entanglement cuts can only be created or annihilated in the presence of the third cut via the triple division or fusion process. One can see that these moves conserve the region volume $|A|$, so the entanglement dynamics only happens within each $|A|$ sector separately, as illustrated in Fig. 3.6(b).
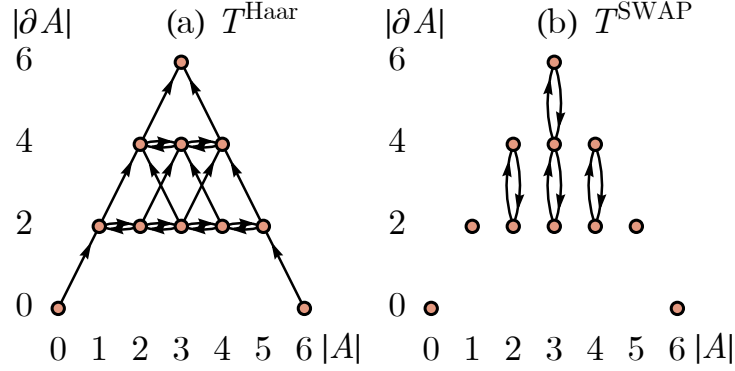
Figure 3.6. Causal structure in the entanglement feature Hilbert space for (a) random unitary circuit, (b) swap circuit. Each note represents a $(|A|, |\partial A|)$-sector. The entanglement features in one sector will only affect those in the other sector along the causal flow (indicated by arrows).

The causal structures in Fig. 3.6 can be formulated as the following algebraic conditions,

$$
\begin{aligned}
P_{|\partial A| \leq 2k} T^{\text{Haar}} P_{|\partial A| > 2k} &= 0 \quad (k = 0, 1, \cdots, N/2), \\
[P_{|A|=n}, T^{\text{SWAP}}] &= 0 \quad (n = 0, 1, \cdots, N),
\end{aligned}
\tag{3.23}
$$

where $P_{\ldots}$ denotes the projection operator that projects to the subspace specified by its subscript. We expect their corresponding continuum limit to respect the same causal structure. We assume that the continuous-time entanglement dynamics will be described by the entanglement feature Hamiltonian in Eq. (3.18), which can be equivalently written as the following expanded form

$$
H_{\text{EF}} = \sum_{\langle ij \rangle} \frac{1 - Z_i Z_j}{2} (u - v(X_i + X_j) + w X_i X_j),
\tag{3.24}
$$

92

with parameters u, v, w related to g, β as

$$
\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{g \cosh \beta}{d^2 - 1} \begin{pmatrix} d^2 - \tanh \beta \\ d - d \tanh \beta \\ 1 - d^2 \tanh \beta \end{pmatrix}. \tag{3.25}
$$

The dynamic process associated to v and w terms are listed in Fig. 3.5(c). For random unitary circuit, the causal structure requires

$$
P_{|\partial A| \leq 2k} H_{\mathrm{EF}} P_{|\partial A| > 2k} = 0 \quad (k = 0, 1, \cdots, N/2), \tag{3.26}
$$

which implies $w = 0$ and hence $\tanh \beta = 1/d^2$. For swap circuit (as the $x \to 1$ limit of the fractional swap circuit), the causal structure requires

$$
[P_{|A|=n}, H_{\mathrm{EF}}] = 0 \quad (n = 0, 1, \cdots, N), \tag{3.27}
$$

which implies $v = 0$ and hence $\tanh \beta = 1$. These parameter correspondences are also obvious by comparing the dynamical constraints between Fig. 3.5(a,b) and Fig. 3.5(c). By comparing the u, v, w parameters in Eq. (3.21) and those in Eq. (3.25), we conjecture that $\tanh \beta = \sin^2 \frac{x\pi}{2} / (2 - \sin^2 \frac{x\pi}{2})$ for the fractional swap circuit at fraction x.

In conclusion, we define the continuum limit of random circuit models according to their definitive causal structures. The results are summarized in Tab. 3.1, which provides us some intuitions about the meaning of β. However, as a free parameter in H_{EF}, β can take any real positive value in general. We are not yet clear what will be (or how to define) the corresponding microscopic circuit models beyond the examples listed in Tab. 3.1, but this might also be an advantage of the entanglement feature formalism, which allows us to investigate the universal behavior of entanglement dynamics at a higher level without much knowledge about underlying model details.

3.3.4 Flow of MPS Parameters

Having clarified the entanglement feature formalism $-\partial_t|W_\Psi\rangle = H_{\text{EF}}|W_\Psi\rangle$ and specified the entanglement feature Hamiltonian H_{EF} in Eq. (3.18), we set out to solve the Schrödinger equation to study the entanglement dynamics. However, as a many-body problem, the Schrödinger equation is hard to solve analytically. Ref. [128] has analyzed the long-time asymptotic behavior of the entanglement dynamics (correspond to the "low-energy" physics of H_{EF}), where it was found that all systems will thermalize to the Page state under locally scrambled quantum dynamics with the thermalization (relaxation) time τ given by $\tau^{-1} = 2g\cosh\beta(1 - 1/d)^2$. Here we would like to pursuit a different direction by representing the entanglement feature state $|W_\Psi\rangle$ as an MPS and exploring the flow of MPS parameters induced by the entanglement dynamics. This is in line with the MPS-based time dependent variational principle (TDVP)[74, 87, 88, 121, 134] developed to simulate quantum many-body dynamics.

The classical dynamics of the MPS parameter $q \equiv (\alpha, \theta)$ (unified as q) follows from (see Appendix B.3 for derivation)

$$-\langle \partial_{q_i} W_\Psi | W_{\mathbb{1}}^{-1} | \partial_{q_j} W_\Psi \rangle \dot{q}_j = \langle \partial_{q_i} W_\Psi | W_{\mathbb{1}}^{-1} H_{\text{EF}} | W_\Psi \rangle, \tag{3.28}$$

where $|W_\Psi\rangle$ is given by the MPS ansatz in Eq. (3.6) and Eq. (3.7), H_{EF} is given by Eq. (3.18). The operator $W_{\mathbb{1}}^{-1} = (\tanh\delta \sinh\delta)^N e^{-\delta \sum_i X_i}$ is a non-trivial metric that maps a state $|W_\Psi\rangle$ to its dual state $\langle W_\Psi | W_{\mathbb{1}}^{-1}$ in the dual Hilbert space (since the entanglement feature Hamiltonian H_{EF} is not Hermitian, the state vector and its dual vector do not coincide)[1]. By solving Eq. (3.28) numerically (see Appendix B.3 for details), we can map out the flow of MPS parameters in the parameter space as shown in Fig. 3.7 for different types of dynamics specified by the β parameter in H_{EF}.

As one can see in Fig. 3.7, for various choice of the β parameter, the MPS parameters all

---

[1] See Ref. [128] for the meaning of the notation $W_{\mathbb{1}}^{-1}$ and why it should correspond to the metric in the entanglement feature Hilbert space.
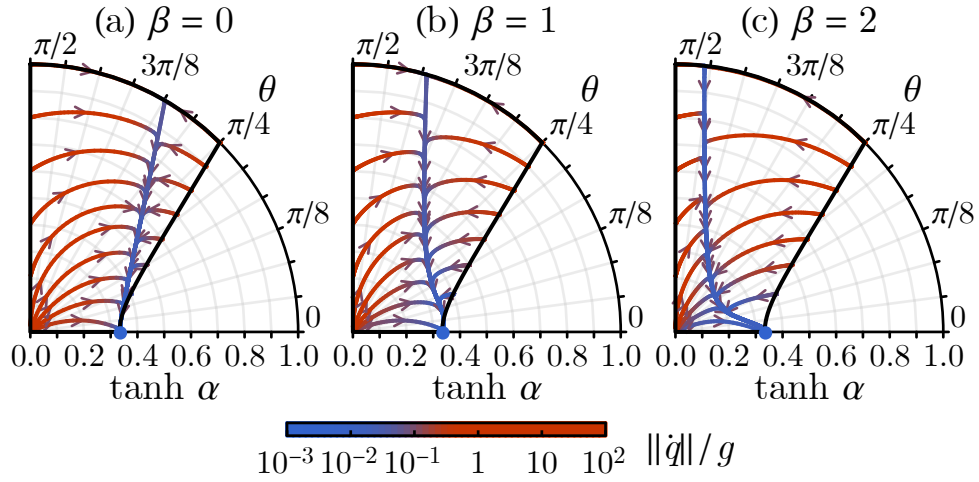
Figure 3.7. Flow of MPS parameters under the entanglement dynamics at (a) $\beta = 0$, (b) $\beta = 1$, (c) $\beta = 2$, in the thermodynamic limit $N \to \infty$ with $d = 2$ qudit dimension. Curve color indicates the flow speed.

flow to the universal fixed point $(\alpha, \theta) = (\frac{1}{2} \ln d, 0)$ which corresponds to the Page state. But the evolution typically divides into two distinct stages. In the early stage (in red), the parameters quickly converge to a mainstream curve. The process is characterized by a large flow speed $\|\dot{q}\|$. Then in the late stage (in blue), the parameters slowly flow along the mainstream towards the Page state fixed point, with a flow speed $\|\dot{q}\|$ orders of magnitude smaller than that of the early stage. The mainstream curve (the central blue curve in Fig. 3.7) exhibits a systematic dependence on β. We recover the multi-region continuum of entanglement entropies from the MPS parameters (α, θ) along the evolution trajectory, as illustrated in Fig. 3.8. Starting from a product state Fig. 3.8(a), where $S_\Psi(A) = 0$ in all regions (such that $S_{\max} = S_{\min} = 0$). Under the early stage evolution, the upper edge $S_{\max}$ quickly evolves to a volume-law curve, while the lower edge $S_{\min}$ remains almost the same near zero, which establishes the multi-region continuum between $S_{\min}$ and $S_{\max}$ as in Fig. 3.8(b). Then the evolution enters the late stage, where the area-law lower edge $S_{\min}$ gradually catches up and finally approaches the (volume-law) Page curve together with $S_{\max}$ as in Fig. 3.8(c). We can thus identify the early (late) stage evolution with the local (global) thermalization.
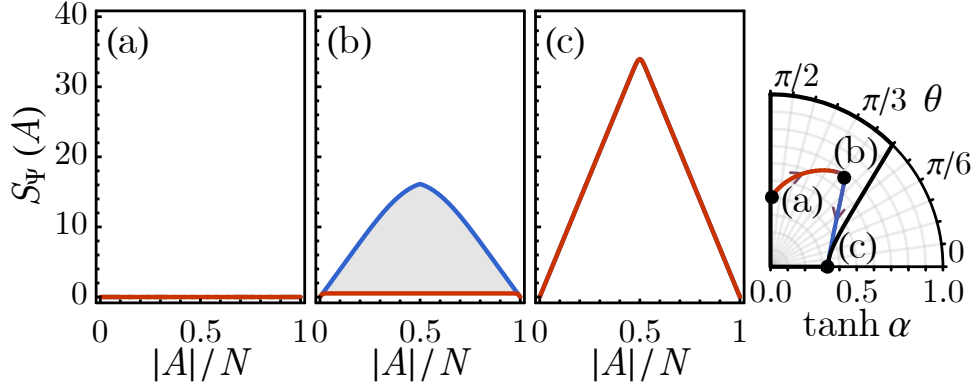
Figure 3.8. Evolution of the entanglement entropy continuum along a typical trajectory in the MPS parameter space through (a) the initial product state, (b) the intermediate area-law state, and (c) the final Page state. The red (blue) curves outlines the bottom $S_{\min}$ and top $S_{\max}$ edges of the continuum, which coincide in (a,c). (a-b) the early stage and (b-c) the late stage evolutions are distinct.

To characterize the two evolution stages more quantitatively, we investigate the behavior of the entropy gap $\Delta(|A|) = S_{1\text{st}}(|A|) - S_{\min}(|A|)$, which was introduced in Eq. (3.2). Using the MPS representation of the entanglement feature state, the "1st excited" entanglement entropy $S_{1\text{st}}(|A|)$ is given by

$$S_{1\text{st}}(|A|) = -\ln \text{Tr}(M^{\downarrow})^{|A|-1} M^{\uparrow} M^{\downarrow} (M^{\uparrow})^{N-|A|-1} + S_0, \tag{3.29}$$

while $S_{\min}(|A|)$ was given in Eq. (3.10). By definition, the entropy gap is positive, i.e. $\Delta(|A|) \geq 0$. It characterizes the minimal deviation of the multi-region entanglement from the single-region entanglement, which reflects the prominence of the multi-region entanglement. In both the short-time $t \to 0$ and long-time $t \to \infty$ limits, the entropy gap vanishes $\Delta(|A|) = 0$ as the whole multi-region continuum collapses. During the time-evolution, as shown in Fig. 3.9, the entropy gap $\Delta(|A|)$ first increases in the early stage as the multi-region continuum gets established, and then decreases in the late stage as $S_{\min}$ catches up with $S_{\max}$ and wipes out the continuum. The decrease of $\Delta(|A|)$ happens gradually from small region to large region as thermalization progresses in the system.
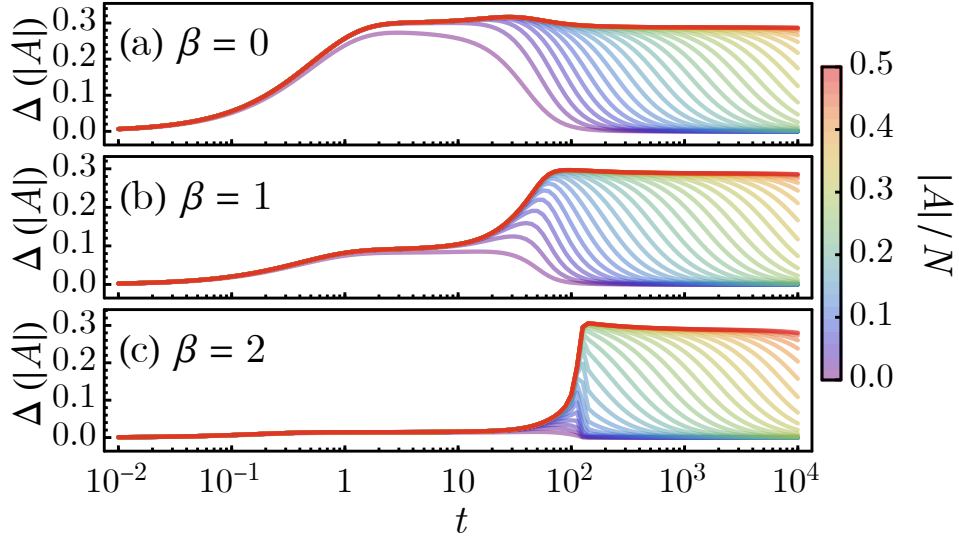
Figure 3.9. Time-evolution of the entropy gap $\Delta(|A|)$ for different dynamics at (a) $\beta = 0$, (b) $\beta = 1$, (c) $\beta = 2$ (assuming $g = 1$). Curves of different colors correspond to different $|A|/N$ ratio.

In conclusion, by studying the evolution of MPS parameters, we effectively parametrize the evolution of all the multi-region entanglement entropies approximately. We find that the two-stage evolution is universal for all types of locally scrambled quantum dynamics, however dynamic details are different for different types of dynamics as specified by different β values.

3.3.5 Effect of Multi-Region Entanglement on Entanglement Velocity

Despite of the concise and effective MPS description of the multi-region entanglement on the overall level, it is still not transparent how the multi-region entanglement affects the entanglement dynamics and in what circumstance is it important to consider the multi-region entanglement. We will analyze these problems from the perspective of single-region entanglement dynamics.

As discussed in Sec. 3.3.3, entanglement entropies from different $|\partial A|$-sectors are causally related under the entanglement dynamics, so it is generally not possible to write down a close-form equation for the evolution of the single-region entanglement entropy. Only in the special case of $\tanh \beta = 1/d^2$ (corresponding to the random unitary circuit), the single-region entangle-

ment is not affected by the multi-region entanglement, such that a close-form equation within the single-region sector ($|\partial A| = 2$) becomes possible. Earlier works[114, 239, 239] that formulated the single-region entanglement dynamics are indeed based on the random unitary circuit model (either explicitly or implicitly). Once we deviate from this special point, the multi-region entanglement will enter the dynamic equation for the single-region entanglement. We will explore its effect in the following.

Our starting point is the entanglement feature formalism, where the entanglement entropy is given by $S_\Psi(A) = -\ln \langle A | W_\Psi \rangle$ with $|A\rangle$ being the basis state labeled by the entanglement region A. From $-\partial_t |W_\Psi\rangle = H_{\mathrm{EF}} |W_\Psi\rangle$, we have

$$\partial_t S_\Psi(A) = \frac{\langle A | H_{\mathrm{EF}} | W_\Psi \rangle}{\langle A | W_\Psi \rangle}. \tag{3.30}$$

Due to the dynamic constrained imposed by the projection operator $\frac{1 - Z_i Z_j}{2}$ in the entanglement feature Hamiltonian H_{EF}, when H_{EF} acting on $\langle A|$ from the right, only those terms across the entanglement cuts of the region A are active. So if the number of entanglement cuts $|\partial A|$ is small, the corresponding dynamics will be simple. The simplest non-trivial case is the dynamics of the single-region entanglement entropy $S_{\min}(|A|)$. Applying the general formula Eq. (3.30) to $S_{\min}(|A|)$, the single-region entanglement growth follows

$$\partial_t S_{\min}(|A|) = 2\Gamma\big(\partial_{|A|} S_{\min}(|A|)\big) + 2\Omega\big(\Delta(|A|)\big),$$
$$\Gamma(s) = g e^{-\beta} \frac{d^2 + 1}{d^2 - 1}\Big(1 - \frac{2d}{d^2 + 1} \cosh s\Big), \tag{3.31}$$
$$\Omega(\Delta) = \frac{g \cosh \beta}{d^2 - 1} (d^2 \tanh \beta - 1)(1 - e^{-\Delta}).$$

The first term $\Gamma(s)$ describes the dynamics within the single-region sector, which depends on the single-region entanglement entropy gradient $s(|A|) \equiv \partial_{|A|} S_{\min}(|A|)$. The second term $\Omega(\Delta)$ describes the contribution of multi-region entanglement, which depends on the entropy gap $\Delta(|A|)$ introduced previously in Eq. (3.2). The factor 2 in front of both terms comes from the two

98

entanglement cuts associated with a single-region A. The effect of the multi-region entanglement on the single-region entanglement dynamics enters explicitly from the $\Omega(\Delta)$ term, which could be of the same order as $\Gamma(s)$ in general. However, only at one special point, i.e. $\tanh\beta = 1/d^2$, does the multi-region effect strictly vanishes $\Omega(\Delta) = 0$, where Eq. (3.31) reduces to

$$\partial_t S_{\min}(|A|) = 2\Gamma\big(\partial_{|A|} S_{\min}(|A|)\big), \tag{3.32}$$

which recovers the result in Ref. [114, 155]. This point corresponds to the random unitary circuit dynamics, where the entanglement dynamics admits the causal structure in Fig. 3.6(a), that the single-region entanglement is not causally affected by the multi-region entanglement, therefore one can arrive at a closed-form equation Eq. (3.32) within the single-region sector. Away from the $\tanh\beta = 1/d^2$ point, Eq. (3.32) is incomplete, we will need to take into account the effect of multi-region entanglement as in Eq. (3.31).

To make broader connections, we study the entanglement velocity v_{E}, which is defined to be the growth rate of the entanglement entropy in half-infinite region on an open chain (with only one entanglement cut in the middle of the chain). It is therefore half of the entropy growth rate $\partial_t S_{\min}$ in Eq. (3.31), i.e. $v_{\mathrm{E}} = \frac{1}{2}\partial_t S_{\min}$. Ref. [53, 114] proposed the entanglement velocity $v_E(s) = \Gamma(s)$ as a function of the entropy gradient s. However, in more general cases, the entanglement velocity $v_{\mathrm{E}}(s,\Delta)$ could also be affected by the multi-region entanglement, characterized by the entropy gap Δ,

$$v_{\mathrm{E}}(s,\Delta) = \Gamma(s) + \Omega(\Delta). \tag{3.33}$$

There is no definitive relation between the entropy gap Δ and the entropy gradient $s = \partial_{|A|} S_{\min}$. However, as we collect data of (s,Δ) pairs over different region sizes $|A|$ at different times t under different dynamics β from different initial conditions, we find that they mostly lies

in a triangle region, as depicted in Fig. 3.10(a) and described by

$$1 - e^{-\Delta} \lesssim (1 - 1/d)^2(1 - |s|/\ln d). \tag{3.34}$$

In fact, most of the data points concentrate along the triangle edges. A few outliers only appear at the tip of the triangle near $s = 0$, which are generated in the early stage local thermalization (which will not show up in the coarse-grained long-time dynamics). The magnitude of the entropy gradient can be interpreted as the entropy density $|s|$, which can not exceed $\ln d$, as each qudit can at most contribute $\ln d$ entanglement entropy. The triangle region shape is only affected by the qudit dimension d as shown in Fig. 3.10(b).
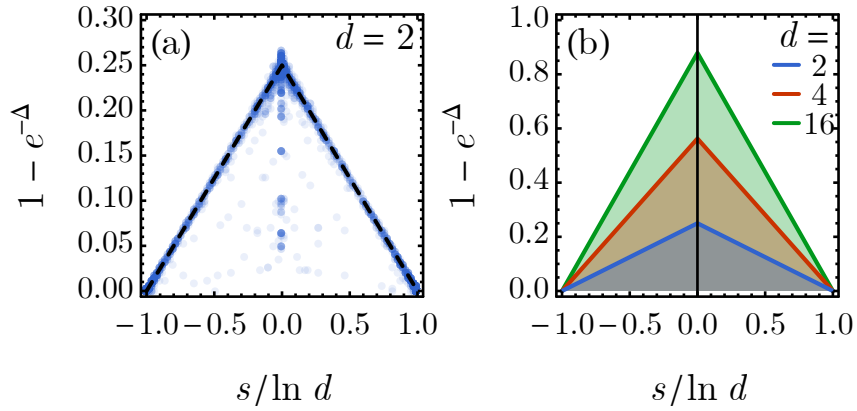
Figure 3.10. (a) Collections of s and Δ (in terms of $1 - e^{-\Delta}$) under locally scrambled quantum dynamics at $d = 2$. Data points mostly distributes within a triangle region. (b) The triangle region for different qudit dimensions d.

Given the result in Eq. (3.33) and the bound in Eq. (3.34) and the definition, we can derive the following bound for the entanglement velocity

$$v_{\mathrm{E}}(s,\Delta) \leq \begin{cases} \Gamma(s) & \tanh\beta \leq 1/d^2, \\ \Gamma(s) + \tilde{\Omega}(s) & \tanh\beta > 1/d^2, \end{cases} \tag{3.35}$$

where $\tilde{\Omega}(s)$ is the upper-bound for $\Omega(\Delta)$ in terms of s,

$$\tilde{\Omega}(s) = \frac{g(d-1)}{d^2(d+1)}(d^2 \sinh\beta - \cosh\beta)\left(1 - \frac{|s|}{\ln d}\right). \tag{3.36}$$

For $\tanh\beta \leq 1/d^2$, $\Omega(\Delta)$ is negative, meaning that the multi-region entanglement tends to slow down the entropy growth in this case, so v_{E} is still bounded by $\Gamma(s)$ from above. For $\tanh\beta > 1/d^2$, $\Omega(\Delta)$ is positive, meaning that the multi-region entanglement will speed up the entropy growth. Especially for large β, $\Omega(\Delta)$ can be much greater than $\Gamma(s)$ and dominates the contribution to v_{E}. In this case, we should bound v_E by $\Gamma(s) + \tilde{\Omega}(s)$, where $\tilde{\Omega}(s)$ provides an upper bound for $\Omega(\Delta)$ according to the observation in Eq. (3.34). In the following, we will further calculate the butterfly velocity v_{B} and compare it with v_{E} to examine the validity of the velocity inequality $v_{\mathrm{E}} \leq (\ln d - |s|)v_{\mathrm{B}}$ proposed in Ref. [53].

3.4 Operator Dynamics in Locally Scrambled Quantum Systems

3.4.1 Out-of-Time-Order Correlator

The entanglement feature formalism not only describes the entanglement dynamics of quantum states, but also applies to the dynamics of operator spreading under locally scrambled quantum systems. Let $U(t)$ describe the unitary time evolution operator by time t. In the Heisenberg picture, a local Hermitian operator O_i on site-i will evolve as $O_i(t) = U(t)O_iU(t)^\dagger$. We are interested in the operator-averaged out-of-time-order correlator (OTOC) at infinite temperature

$$\mathrm{OTOC}(i, j; t) = \mathop{\mathbb{E}}_{O_i, O_j} \mathrm{Tr}\, O_i(t)O_jO_i(t)O_j, \tag{3.37}$$

which provides one way to quantify scrambling by probing how an operator $O_i(t)$ grows with time. For locally scrambled quantum dynamics, the operator $O_i(t)$ is expected to expand ballistically with a butterfly velocity v_{B}, which, in this case, is also the Lieb-Robinson velocity as the OTOC

is calculated in the infinite temperature limit. The butterfly velocity v_B can be extracted from the causal light-cone-like structure of the OTOC in the spacetime. We will calculate it using the entanglement feature formalism as follows.

The operator-averaged OTOC only depends on the operator entanglement of the unitary evolution $U(t)$[133, 137], which can be captured by the entanglement feature operator $W_{U(t)}$, defined as

$$W_{U(t)} = \sum_{A,A'} |A\rangle e^{-S_{U(t)}(A,A')} \langle A'|, \qquad (3.38)$$

where $S_{U(t)}(A,A')$ denotes the entanglement entropy of the operator $U(t)$ (under the operator-state mapping)[96, 163, 183, 237] over region A' on the input (past) side and region A on the output (future) side. As derived in Ref. [228], the OTOC in Eq. (3.37) can be obtained from the entanglement feature operator $W_{U(t)}$ as

$$\text{OTOC}(i,j;t) = d^{-(N+2)} \langle i|W_{U(t)}P|j\rangle, \qquad (3.39)$$

where $|i\rangle$ denotes the Ising basis state $|[\sigma]\rangle$ with a down-spin at site-i and up-spin elsewhere, i.e. $\sigma_i = -1$ and $\sigma_j = +1$ for $j \neq i$, and $P = \prod_i X_i$ is the global spin flip operator.

Now we restrict $U(t)$ to the locally scrambled quantum dynamics, whose corresponding entanglement dynamics is described by the entanglement feature Hamiltonian H_{EF}, then according to Ref. [128], the entanglement feature operator $W_{U(t)}$ will be given by

$$W_{U(t)}W_{\mathbb{1}}^{-1} = e^{-tH_{\text{EF}}}. \qquad (3.40)$$

So the OTOC in Eq. (3.39) can be expanded as

$$\begin{aligned}
\text{OTOC}(i,j;t) &= d^{-(N+2)} \langle i|e^{-tH_{\text{EF}}}W_{\mathbb{1}}P|j\rangle \\
&= \sum_{k=0}^{\infty} \frac{(-t)^k}{k!d^{N+2}} \langle i|H_{\text{EF}}^k W_{\mathbb{1}}P|j\rangle,
\end{aligned} \qquad (3.41)$$

where $W_{\mathbb{1}} = \prod_i (d^2 + dX_i)$ is the entanglement feature operator for the identity operator. For convenience, we define $x \equiv |i - j|$ to be the distance between i and j sites, and the OTOC will only depend on x (and t) given the translation symmetry of H_{EF}. For the $k = 0$ term, $\langle i|W_{\mathbb{1}}P|j\rangle = d^{N+2}$ naturally cancels the denominator. For $k > 0$ terms, the first non-vanishing contribution comes at the $k = x$ order, because it takes at least x steps of local operations in Fig. 3.5 to transform $|j\rangle$ to $|i\rangle$ (by moving the entanglement cuts all the way from j to i). Careful analysis shows that to the leading order in time, we have

$$\text{OTOC}(x,t) = 1 - (1 - d^{-2})\frac{(tg\cosh\beta)^x}{x!} + \mathcal{O}(t^{x+1}). \tag{3.42}$$

Its detailed derivation can be found in Appendix B.4.

3.4.2 Butterfly Velocity and Velocity Inequality

To extract the butterfly velocity v_B, we examine the velocity-dependent OTOC[116] by setting $x = vt$ in $\text{OTOC}(x,t)$. If v happens to match the butterfly velocity, the OTOC will remain constant along the velocity cut in the long time limit (as the OTOC is riding on the kink front of the light-cone). According to Eq. (3.42),

$$\lim_{t\to\infty} \text{OTOC}(vt,t) = \lim_{t\to\infty} 1 - (1 - d^{-2})\left(\frac{g\cosh\beta}{v}\right)^{vt}, \tag{3.43}$$

such that a finite limit of $\lim_{t\to\infty} \text{OTOC}(vt,t)$ (i.e. neither vanishing or diverging) is achieved when and only when $(g/v)\cosh\beta = 1$, thus the butterfly velocity reads

$$v_B = g\cosh\beta. \tag{3.44}$$

We can verify this result by MPS-base numerical evaluation of Eq. (3.39). We start with the MPS representation of the initial state $W_{\mathbb{1}}P|j\rangle$ and apply the evolution operator $e^{-tH_{\text{EF}}} = \prod e^{-\delta t H_{\text{EF}}}$ in its Trotterized form. Following the time evolving block decimation (TEBD)[208,

210, 244] algorithm, we update the MPS tensors by truncated singular value decomposition method. The final MPS state is then overlapped with the $\langle i |$ state to extract the OTOC. The result is shown in Fig. 3.11, which confirms the expression Eq. (3.44) of the butterfly velocity v_B.
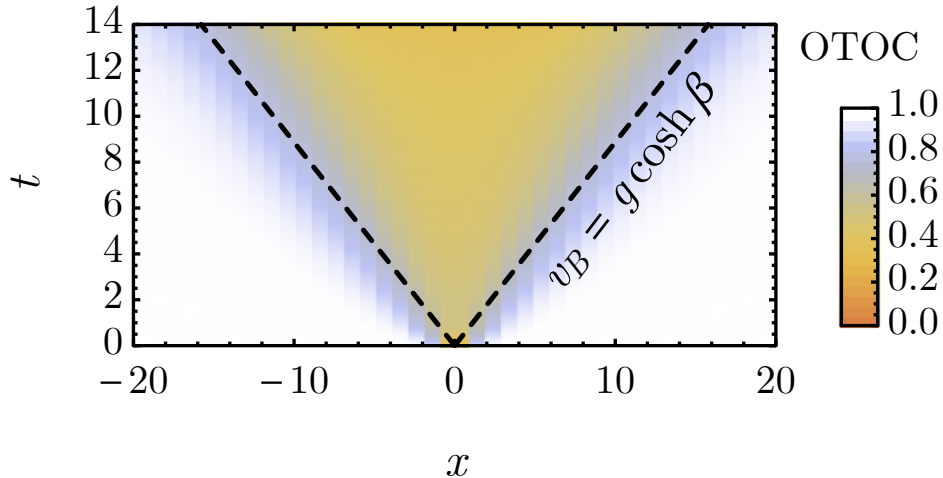
Figure 3.11. OTOC for locally scrambled quantum dynamics at $g = 1$ and $\beta = 0.5$ on a 100-site lattice, calculated by the MPS-based numerical approach with a MPS bond dimension $D = 192$.

We can compare the butterfly velocity v_B and the entanglement velocity $v_E(s, \Delta)$ to test the following velocity inequality

$$v_E(s, \Delta) \leq (\ln d - |s|)v_B, \tag{3.45}$$

which was originally proposed in Ref. [53] in the context of AdS/CFT and in Ref. [114, 148] for the membrane models of entanglement growth. Based on our previous discussion, we have known that the entanglement velocity not only depends on single-region entanglement features like the entropy gradient $s = \partial_{|A|} S_{\min}$, but also depends on multi-region entanglement features like the entropy gap Δ defined in Eq. (3.2). We would like to check if $v_E(s, \Delta)$ can be universally bounded by the Δ-independent right-hand-side $(\ln d - |s|)v_B$ of Eq. (3.45).

Given the entanglement velocity bound in Eq. (3.35), we only need to check a more

restrictive inequality

$$\frac{\tilde{v}_E(s)}{v_B \ln d} \equiv \frac{\Gamma(s) + \tilde{\Omega}(s)}{v_B \ln d} \leq 1 - \frac{|s|}{\ln d}, \tag{3.46}$$

where $\Gamma(s)$ is given in Eq. (3.31) and $\tilde{\Omega}(s)$ is given in Eq. (3.36). It is understood that for $\tanh \beta \leq 1/d^2$, the $\tilde{\Omega}(s)$ term is automatically switched off. $\tilde{v}_E(s) = \Gamma(s) + \tilde{\Omega}(s)$ provides a Δ-independent upper-bound for $v_E(s, \Delta)$, which effectively maxing out the multi-region entanglement effect. If the velocity ratio $\tilde{v}_E/(v_B \ln d)$ satisfies the inequality Eq. (3.46), the velocity inequality in Eq. (3.45) will also hold. As we show in Fig. 3.12, this is indeed the case for all qudit dimensions d at all levels of entropy density $|s|$, under the locally scrambled quantum dynamics of any g and β parameters.
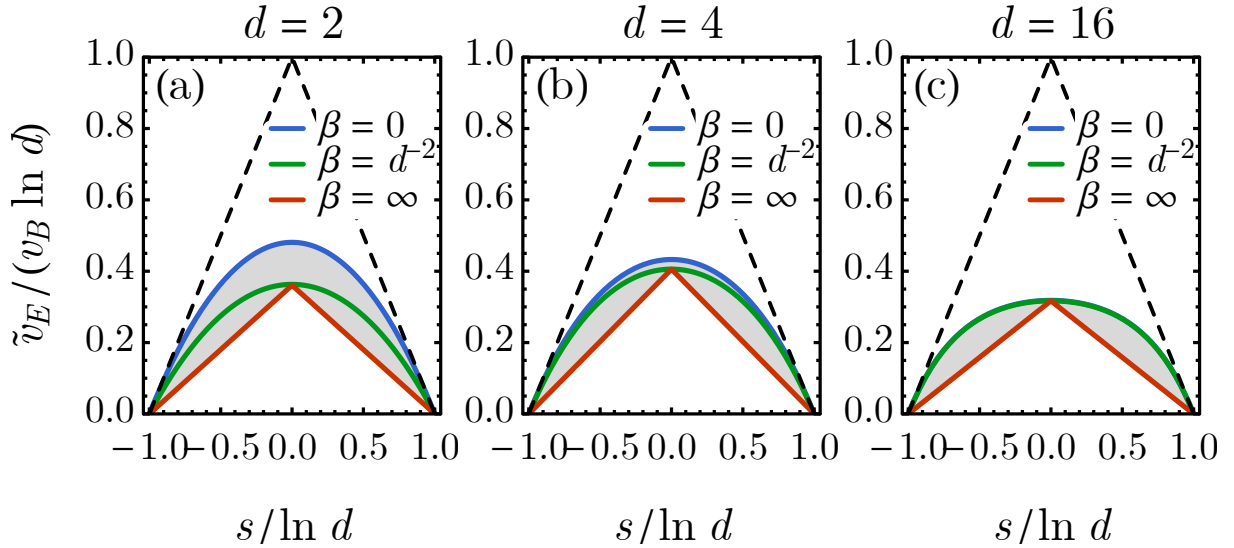
Figure 3.12. The velocity ratio $\tilde{v}_E/(v_B \ln d)$ defined in Eq. (3.46) v.s. the entropy gradient s for different β parameter and different qudit dimensions (a) $d = 2$, (b) $d = 4$, (c) $d = 16$. The dashed line indicates the upper-bound $(1 - |s|/\ln d)$. The gray area is the area swept by the curve as β varies from 0 (blue) to ∞ (red).

To gain a better analytic understanding, we notice that the velocity inequality is tight when $|s| \to \ln d$ and $\beta \to 0$, where the bounding line is tangent to the velocity ratio curve in Fig. 3.12. Given that $\tilde{v}_E(s)$ is a concave function of s, it is actually sufficient to check that its (negative) slope $-\partial_s \tilde{v}_E$ at the $s = \ln d$ corner is smaller than v_B, which allows us to obtain some

simpler analytic results. We can show that

$$
-\partial_s \tilde{v}_E|_{s=\ln d} \leq
\begin{cases}
g e^{-\beta} & \tanh \beta \leq 1/d^2, \\
g e^{-\beta} + \tilde{\Omega}' & \tanh \beta > 1/d^2,
\end{cases}
\tag{3.47}
$$

where the derivative $\tilde{\Omega}'$ is given by

$$
\tilde{\Omega}' = \frac{(d-1)g\cosh\beta}{(d+1)\ln d}(\tanh\beta - 1/d^2).
\tag{3.48}
$$

For $\tanh\beta \leq 1/d^2$, it is obvious that $-\partial_s \tilde{v}_E|_{s=\ln d} \leq g e^{-\beta} \leq g\cosh\beta = v_B$, meaning that the slope is within the bound. For $\tanh\beta > 1/d^2$, we have

$$
\begin{aligned}
-\partial_s \tilde{v}_E|_{s=\ln d} &\leq g e^{-\beta} + \tilde{\Omega}' \\
&= g\cosh\beta \left(1 - \frac{\alpha}{d^2} - (1-\alpha)\tanh\beta\right),
\end{aligned}
\tag{3.49}
$$

where $\alpha = (d-1)/((d+1)\ln d)$ is a d-dependent constant satisfying $0 < \alpha < 1$. Given that $\tanh\beta > 1/d^2$ in this case, Eq. (3.49) can be relaxed to

$$
\begin{aligned}
-\partial_s \tilde{v}_E|_{s=\ln d} &< g\cosh\beta \left(1 - \frac{\alpha}{d^2} - \frac{1-\alpha}{d^2}\right) \\
&= g\cosh\beta(1 - d^{-2}),
\end{aligned}
\tag{3.50}
$$

which is still smaller than $v_B = g\cosh\beta$. Thus we have proven Eq. (3.46) and hence Eq. (3.45) follows. In conclusion, our result shows that the velocity inequality Eq. (3.45) holds for all locally scrambled quantum dynamics (including quantum Brownian dynamics and random unitary circuits).

3.5 Summary

In this work, we study the evolution of bipartite entanglement entropy under locally scrambled quantum dynamics. We point out the importance of the multi-region entanglement in describing the entanglement dynamics. We show that the effect of multi-region entanglement can modify or even dominate the entanglement growth. The common assumption that the entanglement growth rate is only a function of the local entanglement entropy gradient is shown to be incomplete, as it ignores the multi-region entanglement. We identify the explicit contribution from the multi-region entanglement to the entanglement dynamics. We show that the entanglement feature approach reduces to the entanglement membrane approach if the multi-region entanglement ignored, which clarifies the difference and relations between these two approaches.

Our systematic characterization of the multi-region entanglement is based on the recent development of the entanglement feature formalism, which organize the entanglement entropies over all possible bipartitions into a many-body state. We further notice that such a many-body state can be efficiently represented by matrix product states, which could enable efficient numerical simulation of the entanglement dynamics. We propose a two-parameter matrix product state ansatz to capture all the multi-region entanglement entropy. We provide physical interpretations of the ansatz parameters and study their evolution under locally scrambled quantum dynamics. We show that evolution generally consists of an early-stage local thermalization and a late-stage global thermalization with distinct dynamics signatures.

We also gain a deeper understanding of the physical meaning of the parameters in the entanglement feature Hamiltonian by classifying and comparing the causal structure of the entanglement dynamics. Our analysis indicates that different values of the parameter β in the entanglement feature Hamiltonian could correspond to different types of quantum dynamics. Thus different models in the field of non-equilibrium quantum dynamics, such as the quantum Brownian dynamics and the random unitary circuits, are unified within the scope of locally

scrambled quantum dynamics, and can be discussed in a more systematic manner.

We calculate the operator-averaged out-of-time-order correlator in the infinite temperature limit for locally scrambled quantum dynamics, from which we extract the butterfly velocity and establish its dependence on the β parameter. Despite of the multi-region region entanglement effect, we still find that the previously conjectured inequality between the entanglement velocity and the butterfly velocity remains valid for all values of β.

For future works, it is desired to extend the current approach to more general dynamics beyond the locally scrambled quantum dynamics, or to incorporate symmetry into entanglement dynamics. It will also be interesting to explore the entanglement dynamics in higher dimensions, where the entanglement feature state could exhibit topological order.

<hr>

Chapter 3, in full, is a reprint of the material as it appears in Multiregion entanglement in locally scrambled quantum dynamics 2020. Akhtar, Ahmed; You, Yi-Zhuang, Physical Review B, 2020. The dissertation author was the first author and primary investigator of this paper.

Chapter 4

A Useful Application for Efficient Average Purity Dynamics: Classical Shadows in Shallow Circuits

4.1 Introduction

Quantum technology is advancing rapidly. A central task is to characterize and exploit the features of many-qubit quantum states created in the lab [33, 54, 82, 110]. To fully determine the density matrix ρ of a quantum system of n qubits, exponential ($\sim 4^n$) amount of repeated measurements and classical processing is needed [66, 84, 166]. Therefore full quantum state tomography is not scalable to large systems. However, for many purposes (e.g. estimating physical observables on the quantum state), a far less complete description is adequate [1, 3], and the amount of measurement and classical processing can be drastically reduced. Much of the recent progress has been made by exploiting the randomized measurement strategy [58, 165, 167], particularly through the *classical shadow tomography* [40, 48, 85, 89, 99, 101, 102, 122, 135, 192, 196, 233].

Classical shadow tomography converts quantum states to classical data with a superior sample efficiency. It can estimate expectation values of M observables using only $\sim (\log M)\|O\|_{\mathrm{shd}}^2$ independent randomized measurements [102, 174], saturating the theoretical optimal bound on sample efficiency. Nevertheless, the constant coefficient $\|O\|_{\mathrm{shd}}^2$, known
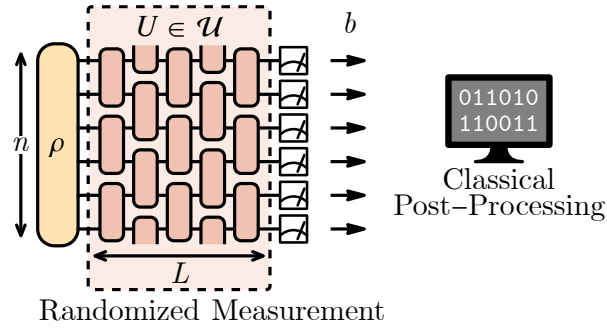
Figure 4.1. Classical shadow tomography in a system of *n* qubits by randomized measurement via a finite-depth local random unitary circuit of *L* layers.

as the *operator shadow norm*, does depend on the type of observable O and the randomized measurement scheme. For example, the single-qubit (local) Pauli measurements are efficient for predicting local observables, while the global Clifford measurements are efficient in estimating certain global properties such as quantum fidelity. However, an efficient and scalable approach to interpolate the local and global limits is still missing in classical shadow tomography.

As illustrated in Fig. 4.1, the randomized measurement protocol in classical shadow tomography is generally realized by first transforming the quantum state ρ of interest by a random unitary circuit U sampled from a random unitary ensemble $\mathscr{U}$, then performing a computational basis measurement and collecting the measurement outcomes b (as a bit-string). Much of the literature has focused on two measurement protocols: the single-qubit Pauli measurements (where $\mathscr{U} = \mathrm{Cl}(2)^{\otimes n}$) and the global Clifford measurements (where $\mathscr{U} = \mathrm{Cl}(2^n)$). In terms of the depth L of the random unitary circuit $U \in \mathscr{U}$, we may think of these two randomized measurement schemes as the $L \to 0$ and $L \to \infty$ limits of randomized measurement protocols defined on more general finite-depth unitary ensembles (see Fig. 4.1). This work aims to explore the intermediate measurement schemes based on finite-depth circuits between these two limits. For simplicity, we focus on the collection of unitary ensembles defined by the brick wall arrangement of two-local random unitary gates as shown in Fig. 4.1, but our approach is directly applicable to other circuit structures as well.

The ability to adjust the circuit depth L (or even the circuit structure) provides great
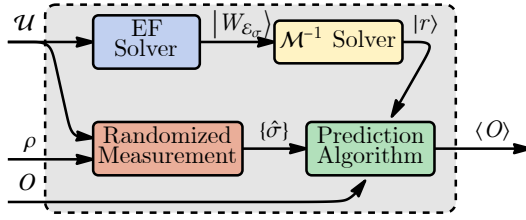
Figure 4.2. We outline the procedure for predicting $\langle O \rangle = \text{Tr}(\rho O)$ using the shallow circuit classical shadow tomography approach. The ingredients are the state ρ, the observable O, and the unitary ensemble $\mathcal{U}$ defining the measurement protocol. The entanglement feature $W_{\mathcal{E}_\sigma}$ is totally determined by the second moment of the unitary ensemble $\mathcal{U}$. The snapshot states $\hat{\sigma}$ is collected from randomized measurements given $\mathcal{U}$ and ρ. The reconstruction map $\mathcal{M}^{-1}$ is determined by the entanglement features only, and is encapsulated by the reconstruction coefficients r. These reconstruction coefficients r, alongside the snapshot states $\hat{\sigma}$ and the observable O are used to make predictions. Importantly, the entire algorithm only involves matrix product states and stabilizer states. Both are efficient representations of quantum states that can be processed on classical computers.

flexibility to modify the measurement scheme adaptively so as to optimize the tomography efficiency for the given target observables. This can be particularly useful when the target observables are not neatly local or low-rank (such that neither Pauli nor Clifford measurement is optimal). The shallow circuit implementation of randomized measurements is also friendly to near-term quantum devices. Motivated by these objectives, Ref. [98] develops the foundation for shallow-circuit classical shadow tomography based on the theory of *locally scrambled* quantum dynamics and the *entanglement feature* formalism [5, 63, 100, 129, 229, 231]. The key finding is that for a large class of randomized measurements, the classical shadow reconstruction map only depends on the entanglement properties of the classical snapshot states $U^\dagger |b\rangle$, which can be computed by solving a quantum entanglement dynamics problem.

However, the reconstruction algorithm demonstrated in Ref. [98] is based on a brute-force computation, which is not scalable to larger systems (the complexity for classical post-processing will be exponential in system size). The largest system size achieved in Ref. [98] was nine qubits. Facing this challenge, in this work, we further develop efficient numerical methods based on *matrix product state* (MPS) [50, 68, 92, 171] for calculating entanglement features and solving reconstruction coefficients. The key idea is to pack the entanglement feature (purities in

different entanglement regions) of the classical snapshot state $U^\dagger |b\rangle$ into a fictitious quantum many-body state and represent this entanglement feature state by an MPS. This enables efficient computation of the entanglement dynamics and finding the classical shadow reconstruction map. Our tomography procedure is outlined in Fig. 4.2. Using this approach, we can perform the classical shadow post-processing on 22 qubits ($n = 22$) for the randomized measurement scheme based on three-layer random Clifford circuits ($L = 3$) with superior sample efficiency. The primary advantages of this approach are its scalability and flexibility, i.e. the classical post-processing can be performed for large system sizes with *polynomial complexity* (given a fixed amount of measurements), while the method applies to a wide range of random Clifford measurement schemes.

The paper is organized as follows. In Sec. 4.2, we review the basic setup of classical shadow tomography and the key concepts of entanglement feature formalism. In Sec. 4.3, we describe the general algorithm/procedure for computing the expectation values of generic observables and analyzing the corresponding shadow norm. In Sec. 4.4, we demonstrate the shallow-circuit classical shadow tomography on the Greenberger-Horne-Zeilinger (GHZ) state and the cluster state. The simulation confirms an unbiased tomographic reconstruction with favorable sample efficiency compared to that of Pauli measurements. In Sec. 4.5, we suggest some future lines of inquiry and summarize the key results.

4.2   Background

4.2.1   General Framework of Classical Shadow Tomograhy

Classical shadow tomography is an efficient protocol for predicting features of a quantum state based on a few measurements [102]. The trick is to use entanglement generated from a random unitary ensemble $\mathcal{U}$ to form approximate classical "shadows" $\hat{\rho}$ of the original state ρ which are then used for predicting the desired observable. Different unitary ensembles produce different collections of shadows, which are suited for predicting different kinds of observables.

The protocol consists of two steps: randomized measurement and classical post process-ing, as illustrated in Fig. 4.1. Starting with the initial state ρ, apply a randomly sampled unitary $U \in \mathscr{U}$ so that $\rho \rightarrow U\rho U^\dagger$. Then perform a projective measurement on the transformed state $U\rho U^\dagger$ in the computational basis (Z-basis on each qubit independently). The resulting bit-string state $|b\rangle$ (labeled by the bit-string b of measurement outcomes), as well as the performed unitary transformation U, will be recorded. These two pieces of information define a *snapshot state* $\hat{\sigma} = U^\dagger|b\rangle\langle b|U$. The randomized measurement protocol can be formulated as a quantum channel $\mathscr{M}$, which maps the initial state ρ to the average snapshot state σ by

$$\sigma = \mathscr{M}(\rho) := \mathop{\mathbb{E}}_{\hat{\sigma} \in \mathscr{E}_{\sigma|\rho}} \hat{\sigma} = 2^n \mathop{\mathbb{E}}_{\hat{\sigma} \in \mathscr{E}_\sigma} \hat{\sigma}\, \mathrm{Tr}(\hat{\sigma}\rho), \tag{4.1}$$

where $\mathscr{E}_\sigma = \{U^\dagger|b\rangle\langle b|U \mid b \in \{0,1\}^{\times n}, U \in \mathscr{U}\}$ denotes the *prior* ensemble of all possible classical snapshots. In the prior snapshot ensemble, the joint probability to sample a (b,U) pair is $P(b,U) = 2^{-n}P(U)$, where $P(U)$ is the probability to sample U and 2^{-n} is the probability to sam-ple b uniformly (independent of ρ). Given the observation of ρ, the *posterior* snapshot ensemble $\mathscr{E}_{\sigma|\rho}$ is defined by deforming the joint probability distribution to $P_\rho(b,U) = \langle b|U\rho U^\dagger|b\rangle P(U)$. The measurement channel $\mathscr{M}$ can be formulated based on either of the snapshot ensembles equivalently, as stated in Eq. (4.1).

Furthermore, if this measurement channel $\mathscr{M}$ is tomographically complete, i.e. distinct states ρ can be distinguished by different measurement outcomes (b,U), then there is an inverse map $\mathscr{M}^{-1}$ called the *reconstruction map* such that $\rho = \mathscr{M}^{-1}(\sigma)$. The image of a given snapshot state $\hat{\sigma}$ under the reconstruction map is a *classical shadow* $\hat{\rho} = \mathscr{M}^{-1}(\hat{\sigma})$, which provides a single-shot classical representation of the initial quantum state ρ. The initial state ρ can be restored as the ensemble expectation of classical shadows $\rho = \mathbb{E}_{\hat{\sigma} \in \mathscr{E}_{\sigma|\rho}} \mathscr{M}^{-1}(\hat{\sigma})$, which enables us to predict physical observables of the initial state based on the classical snapshots collected

from randomized measurements,

$$\langle O \rangle = \text{Tr}(O\rho) = \mathop{\mathbb{E}}_{\hat{\sigma} \in \mathscr{E}_{\sigma|\rho}} \text{Tr}(O\mathscr{M}^{-1}(\hat{\sigma})). \tag{4.2}$$

This computation will be carried out on a classical computer. It will be efficient if the snapshot states are stabilizer states. Since the reconstruction map $\mathscr{M}^{-1}$ is not a quantum channel (because it is not a positivity-preserving map), the resulting shadow states $\hat{\rho}$ may not be positive semidefinite. Therefore, one may encounter unphysical single-shot estimation values of O for certain individual samples. Nevertheless, the ensemble average of all single-shot estimations is physical and unbiased.

In practice, quantum devices can have a significant amount of noise which can introduce randomness into meaurement outcomes. One solution to deal with noise is to incorporate the noise into the measurement channel itself. The idea is that the prior snapshot ensemble is composed of several layers of channels acting on an initial product state $\mathscr{E}_{\sigma} = \mathscr{E}_1 \circ \mathscr{E}_2 \ldots \mathscr{E}_L(|b\rangle\langle b|)$ which are averaged over to form the measurement channel. Incorporating noise amounts to simply introducing new noise layers resulting in a new measurement channel $\mathscr{M}'$ based on the statistical properties of the noise. If we assume that the noise is local and does not have a preferential basis, then the approach outlined in this paper would still be applicable and the resulting reconstruction map $(\mathscr{M}')^{-1}$ should have an efficient MPO description in the shallow circuit limit.

For any finite-sized classical shadow ensemble, the estimation exhibits statistical fluctuations around the true expectation value $\langle O \rangle$. The variance in the single-shot estimation of any observable O can be quantified through its *operator shadow norm* $\|O\|_{\text{shd}}^2$, introduced in Ref. [102]. The estimated expectation value obtained from averaging over M samples will be within the error bound of $\|O\|_{\text{shd}}^2/M$ with high probability. For a given class of observables, it is possible to choose the random unitary ensemble $\mathscr{U}$ wisely so as to reduce the shadow norm and to optimize the sample efficiency [98].

4.2.2 Locally-Scrambled Classical Shadow Tomography

Restricting the randomized measurement scheme to the scope of locally-scrambled ensembles proves to be advantageous [98]. Locally-scrambled ensembles [129] are random unitary ensembles which are invariant under both left and right local basis transformations, i.e. the probability distribution for all $U \in \mathscr{U}$ satisfies $P(U) = P(UV) = P(VU)$ where $V = \bigotimes_i V_i$ with $V_i \in \mathrm{U}(2)$ being any local basis transformation. Many typically studied random circuit models are in fact, local basis invariant, such as random Haar circuits [14, 45, 63, 156, 159, 240] and quantum Brownian circuits [42, 70, 131, 225, 235]. Furthermore, since the reconstruction map only depends on the second moment of the unitary ensemble, Clifford circuits, which are typically used for classical shadow tomography because of their classical simulability, are equivalent in terms of their tomographic properties to the random Haar circuit. Therefore, the formulation of locally-scrambled classical shadow tomography in Ref. [98] is applicable to generic random Clifford circuits of any depth and any gate arrangement.

In the locally-scrambled case, it was shown that the reconstruction map is determined entirely by the average purity (exponential of 2nd Rényi entropy) of the (prior) classical snapshot states [98] in *all* possible sub-regions, which are also known as the *entanglement features* [5, 63, 129, 229]. One might be daunted by the fact that there are exponentially many entanglement features to keep track of, one for every possible sub-region of the system. Thus the classical post-processing will not be scalable. This paper aims to answer the challenge. In particular, two questions will be addressed:

1. Is there any hope of representing the entanglement features efficiently?

2. Furthermore, can we utilize such representation to determine and apply the reconstruction map efficiently to large systems?

The answer to the first question is yes: average sub-region purities of locally-scrambled systems have efficient representations as matrix product states (MPS) due to the fact that purities

are completely positive and many-body states of positive wave functions are typically area-law entangled [79], such that the MPS bond dimension is a constant that does not scale with the system size n. The key idea is that the average purity of a n-qubit quantum system over all sub-regions can be compactly encoded in a fictitious many-body state vector, called the *entanglement feature state*. Given an ensemble $\mathscr{E}$ of random states (random density matrices) of the system, the entanglement feature state $|W_\mathscr{E}\rangle$ associated with $\mathscr{E}$ is a vector whose components are

$$(W_\mathscr{E})_A = \langle A|W_\mathscr{E}\rangle := \underset{\rho \in \mathscr{E}}{\mathbb{E}} \, \mathrm{Tr}_A(\mathrm{Tr}_{\bar{A}}\rho)^2, \qquad (4.3)$$

where A denotes a sub-region of the system (as a subset of qubits) and $\bar{A}$ is the complement region of A. For each state ρ in the ensemble $\mathscr{E}$, $\rho_A = \mathrm{Tr}_{\bar{A}}\rho$ is the reduced density matrix of ρ in the region A, by tracing out qubits in the region $\bar{A}$. The entanglement feature $(W_\mathscr{E})_A$ in the region A is simply the purity of ρ_A averaged over all random states ρ in the ensemble $\mathscr{E}$. The entanglement feature state $|W_\mathscr{E}\rangle$ is a superposition of the basis states $|A\rangle$ labeled by sub-regions, and the superposition coefficients are the average purities in the corresponding sub-regions. Note that the entanglement feature state is not a physical state of the quantum system. It merely encodes the entanglement properties of the underlying physical quantum states. For example, the ensemble of *pure product states* has the following entanglement feature state

$$|W_{\mathrm{prod}}\rangle = \sum_A |A\rangle, \qquad (4.4)$$

because the purity is unity in any sub-region of any pure product state. In this way, we can efficiently encode the purity information over all sub-regions in a single entanglement feature state $|W_\mathscr{E}\rangle$.

The advantage of the entanglement feature formalism lies in the fact that there exists an efficient numerical method to calculate the *time evolution* of the entanglement feature state as the underlying quantum state evolves in time together. Let $\mathscr{U}$ be a locally-scrambled random unitary

ensemble and $\mathscr{E}$ be a random state ensemble. One can define a new random state ensemble $\mathscr{E}' = \{U\rho U^\dagger \mid \rho \in \mathscr{E}, U \in \mathscr{U}\}$, which can be interpreted as the ensemble of the random states evolved by the random unitaries. Then the entanglement features of $\mathscr{E}'$ and $\mathscr{E}$ are related by

$$|W_{\mathscr{E}'}\rangle = \hat{W}_{\mathscr{U}} \hat{W}_{\mathbb{1}}^{-1} |W_{\mathscr{E}}\rangle, \tag{4.5}$$

where $\hat{W}_{\mathscr{U}}$ is called the *entanglement feature operator* associated with the random unitary ensemble $\mathscr{U}$, whose matrix components are given by $\langle A|\hat{W}_{\mathscr{U}}|B\rangle = \mathbb{E}_{U\in\mathscr{U}} \operatorname{Tr}_{A,B}(\operatorname{Tr}_{\bar{A},\bar{B}} U)^2$ with A and B being the sub-regions on the future and the past sides of U respectively. $\hat{W}_{\mathbb{1}}$ is the entanglement feature operator for the identity operator, and its inverse operator is denoted as $\hat{W}_{\mathbb{1}}^{-1}$. Therefore, each step of the unitary evolution of the physical state ensemble can be mapped to a step of transfer matrix evolution of the entanglement feature state.

An efficient numerical approach to compute the entanglement feature dynamics Eq. (4.5) has been developed [5, 63] based on the MPS representation of entanglement feature states.

$$(W_{\mathscr{E}})_A = \langle A|W_{\mathscr{E}}\rangle = \operatorname{Tr} \prod_i (\delta_{i\notin A} \mathsf{W}^0_{\mathscr{E},i} + \delta_{i\in A} \mathsf{W}^1_{\mathscr{E},i}). \tag{4.6}$$

Here $\mathsf{W}^0_{\mathscr{E},i}$ and $\mathsf{W}^1_{\mathscr{E},i}$ are two sets of matrices parametrizing the MPS, where i labels the qubit/site. If the state ensemble $\mathscr{E}$ is translation invariant, we may choose to drop the index i on the MPS matrices $\mathsf{W}^0_{\mathscr{E},i}, \mathsf{W}^1_{\mathscr{E},i}$. Furthermore, for the pure state entanglement feature, the MPS matrices must satisfy various constraints. For example, for pure states, the purity of the whole system is one i.e. $\operatorname{Tr}(\mathsf{W}^0_{\mathscr{E}})^n = \operatorname{Tr}(\mathsf{W}^1_{\mathscr{E}})^n = 1$. In addition, because the purity of complementary regions is the same, i.e. $(W_{\mathscr{E}})_A = (W_{\mathscr{E}})_{\bar{A}}$, the entanglement feature state has a $\mathbb{Z}_2 : A \to \bar{A}$ symmetry, so the MPS tensor (formed by combining the two matrices into a single tensor) must carry definite $\mathbb{Z}_2$ representation. These symmetries greatly constrain the form of the MPS matrices (see Ref. [5] for more discussions).

As a fictitious many-body state, the entanglement feature state is typically area-law

entangled due to its completely positive sign structure [79]. Thus it generally admits efficient MPS representations if the system is one-dimensional. For example, the entanglement feature state of pure product states Eq. (4.4) can be represented as a trivial MPS (a product state) with bond dimension one. Most typical entanglement feature states will have a low bond dimension even as the underlying physical state ensemble is of volume-law entanglement. For nearly all locally-scrambled dynamics with various entanglement production rates, it has been numerically observed [5] that a bond dimension $D_W = 2$ MPS state could capture the purity evolution in all sub-regions to high accuracy. The accuracy can be further improved with a larger bond dimension. The evolution of the entanglement feature state, for example, by applying alternating layers of brick wall unitaries, can also be performed at the MPS level using the entanglement feature transfer matrix following Eq. (4.5). To implement that, we use a time-evolved block decimation (TEBD) approach and truncate at a fixed bond dimension D_W (see Appendix C of Ref. [63] or Appendix A, C of Ref. [5] for further details). With these previously developed tools, one can efficiently calculate the evolution of the average purity given a unitary ensemble. Furthermore, given the same circuit geometry, all unitary two-designs have the same entanglement feature dynamics, and so the Clifford circuit entanglement feature transfer matrix is equivalent to the random Haar one.

For the purpose of classical shadow tomography, what we need to know is the entanglement feature of the classical snapshots. Given any randomized measurement scheme (as specified by the random unitary ensemble $\mathscr{U}$), all possible classical snapshots form a random state ensemble, denoted as

$$\mathscr{E}_\sigma = \{U^\dagger |b\rangle\langle b|U \mid b \in \{0,1\}^{\times n}, U \in \mathscr{U}\}. \tag{4.7}$$

This is also the prior snapshot ensemble defined below Eq. (4.1). Following the definition in Eq. (4.3), we can define the entanglement feature state $|W_{\mathscr{E}_\sigma}\rangle$ for the classical snapshot ensemble

$\mathscr{E}_\sigma$, whose components are

$$(W_{\mathscr{E}_\sigma})_A = \langle A | W_{\mathscr{E}_\sigma} \rangle := \underset{\hat{\sigma} \in \mathscr{E}_\sigma}{\mathbb{E}} \operatorname{Tr}_A(\operatorname{Tr}_{\bar{A}} \hat{\sigma})^2. \tag{4.8}$$

Note that the classical snapshot state $\hat{\sigma} = U^\dagger |b\rangle\langle b| U$ is always constructed from a (reversed) unitary evolution $U^\dagger$ on a pure product state $|b\rangle$. So to obtain $|W_{\mathscr{E}_\sigma}\rangle$, we only need to take the MPS representation of $|W_{\mathrm{prod}}\rangle$ in Eq. (4.4) as the initial state and use the TEBD algorithm to evolve the entanglement feature MPS step by step following the entanglement feature dynamics in Eq. (4.5), as the underlying snapshot state gets evolved by the (reversed) Clifford circuit $U^\dagger$ layer by layer. Computing the entanglement feature $|W_{\mathscr{E}_\sigma}\rangle$ from the random unitary ensemble $\mathscr{U}$ is denoted as "EF Solver" in Fig. 4.2. The algorithm complexity is linear in the circuit depth L and independent of the number of qubit n (assuming translation symmetry).

According to Ref. [98], the entanglement feature of classical snapshots $|W_{\mathscr{E}_\sigma}\rangle$ is all we need to obtain the classical shadow reconstruction map $\mathscr{M}^{-1}$ for the corresponding randomized measurement scheme, as long as the random unitary ensemble $\mathscr{U}$ is locally scrambled (which applies to random Clifford circuits). The steps to obtain $\mathscr{M}^{-1}$ are as follows:
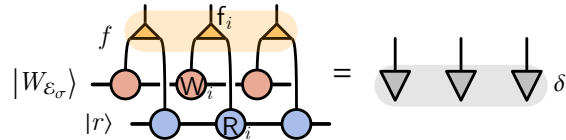
Figure 4.3. Illustration of the MPS equation Eq. (4.9). The fusion tensor f_i is defined in Tab. 4.1. The reconstruction coefficient $|r\rangle$ can be determined from the entanglement feature $|W_{\mathscr{E}_\sigma}\rangle$ by solving this MPS equation.

First, one solves for the *reconstruction coefficients* $|r\rangle$ (or $r_A = \langle A|r\rangle$) from $|W_{\mathscr{E}_\sigma}\rangle$ through the following consistency equation derived in Ref. [98] (which is essentially requiring $\mathscr{M}^{-1}\mathscr{M} = \mathbb{1}$ at the quantum channel level)

$$\sum_{A,C} r_A f_{A,B,C} (W_{\mathscr{E}_\sigma})_C = \delta_{B=\Omega}. \tag{4.9}$$

Here, A, C are summed over all sub-regions of the system (as subsets of qubits). Ω stands for the full system (as the full set of qubits) and $\delta_B = 1$ if $B = \Omega$ otherwise, $\delta_B = 0$. Also, r_A and $(W_{\mathscr{E}_\sigma})_C$ are components of the corresponding vectors $|r\rangle$ and $|W_{\mathscr{E}_\sigma}\rangle$, and $f_{A,B,C}$ is a fusion tensor that can be factorized to each site as $f = \otimes_{i=1}^{n} \mathsf{f}_i$, where components of the on-site tensor f_i are given by Tab. 4.1.

Table 4.1. Components of the f_i tensor (for qubit systems). The systematic formula for f_i in generic qudit systems can be found in Ref. [98].

f	$i \notin A$		$i \in A$	
	$i \notin C$	$i \in C$	$i \notin C$	$i \in C$
$i \notin B$	2	0	$\frac{8}{3}$	$-\frac{4}{3}$
$i \in B$	0	0	$-\frac{2}{3}$	$\frac{4}{3}$

Using MPS representations for both $|r\rangle$ and $|W_{\mathscr{E}_\sigma}\rangle$, and the tensor product structure of f, the left hand side of Eq. (4.9) may be represented as an MPS equation, as shown in Fig. 4.3. The left-hand side of Eq. (4.9) is made up of an MPS of bond dimension $D_r \times D_W$, whereas the right-hand side of Eq. (4.9) is a trivial MPS (a product state). We may then solve for the MPS tensors describing $|r\rangle$ by minimizing the following loss function

$$
\begin{aligned}
\mathscr{L} &= \|\mathscr{M}^{-1}\mathscr{M} - \mathbb{1}\|^2 \\
&= \sum_B \left(\sum_{A,C} r_A f_{A,B,C} (W_{\mathscr{E}_\sigma})_C - \delta_{B,\Omega} \right)^2.
\end{aligned}
\tag{4.10}
$$

The optimization is achieved by the gradient descent algorithm using the PyTorch package [177] for auto-differentiation. The loss function can be suppressed to $\mathscr{L} \lesssim 10^{-3}$ with a modest bond dimension $D_r = 6$ for circuit depths $L \le 5$. In summary, for shallow circuits, we numerically observe that the reconstruction coefficients can be well described by a low bond-dimension MPS. Solving the reconstruction coefficient $|r\rangle$ from the entanglement feature $|W_{\mathscr{E}_\sigma}\rangle$ is denoted as the "$\mathscr{M}^{-1}$ Solver" in Fig. 4.2. The algorithm complexity is *linear* in the number n of qubits and *independent* of the circuit depth L.

Then, the reconstruction map is a linear functional of the reconstruction coefficients.

$$\mathcal{M}^{-1}(\sigma) = 2^n \sum_A r_A \sigma_A \tag{4.11}$$

Each reconstruction coefficient is weighted by a partial trace of the input σ embedded back into the Hilbert space so that the sum is over normalized density matrices on the full Hilbert space. More specifically, $\sigma_A = 2^{|A|-n}(\text{Tr}_{\bar{A}}\,\sigma) \otimes \mathbb{1}^{\otimes(n-|A|)}$, which is a linear functional on σ. In practice, one works with individual snapshot states as opposed to their ensemble average. For Clifford circuit-based randomized measurement, each individual snapshot state is a stabilizer state, which admits efficient post-processing on a classical computer.

We would like to mention that the analytic expression for the reconstruction coefficient $|r\rangle$ is known from [28] by solving Eq. (4.9)

$$r_A = 2^{-n}(-1)^{|A|} \sum_{C|C \supseteq A} \frac{3^{|C|}}{\sum_{B|B \subseteq C}(-2)^{|B|}(W_{\mathscr{E}_\sigma})_B}. \tag{4.12}$$

Nevertheless, it does not seem to help our purpose, as we do not know how to encode this expression as a MPS efficiently. Directly evaluating Eq. (4.12) for every sub-region A is not scalable. Therefore, we will still follow our approach to solve Eq. (4.12) by MPS optimization.

In the next section, we will answer the second question: how can we leverage the scalability and efficiency of matrix product states for prediction using classical shadow tomography?

4.3 Applications for scalable prediction

4.3.1 Matrix Product State Representation of Reconstruction Coefficients

We consider that the quantum system is made of n qubits arranged on a one-dimensional lattice. We assume that there exists an efficient MPS representation for the reconstruction coefficients r_A. The existence of an efficient representation for $|r\rangle$ is implied from the tensor

network structure of the self-consistency equation and the efficient MPS representation for the entanglement feature. Let $R_i^j, j = 0, 1, i \in \{1 \cdots n\}$ be the MPS matrices for $|r\rangle$ at site i of type j, i.e. R_i^j is a $D_r \times D_r$ real matrix, satisfying

$$r_A = \langle A|r \rangle = \text{Tr}\left(B \prod_{i=1}^{n} (\delta_{i \notin A} R_i^0 + \delta_{i \in A} R_i^1) \right), \tag{4.13}$$

where B is a $D_r \times D_r$ real matrix serving as a twisted periodic boundary condition for the MPS. Here the site index $i = 1, 2, \cdots, n$ labels the qubits on the one-dimensional lattice. For $L = 0, \infty$, we may drop the site index altogether because there is translation invariance. For $0 < L < \infty$, there is only a two-site translation invariance, so only the parity of the site is important. For example, consider the deep circuit limit ($L = \infty$), which corresponds to random Clifford measurements. In this case, $r_\emptyset = -1$ and $r_\Omega = 1 + 2^{-n}$, where 2 is the local Hilbert space dimension and n is the number of qubits. One choice of tensors here is

$$R^j = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & j = 0 \\ \begin{pmatrix} 0 & 0 \\ 0 & r_\Omega^{1/n} \end{pmatrix} & j = 1 \end{cases} \qquad B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.14}$$

As the MPS matrix R_i^j is translationally invariant, we will drop the site index i and denote it as R^j. Introducing the boundary tensor can help simplify calculations. For example, without inserting the boundary tensor (or equivalently setting $B = \mathbb{1}$), we have that $\lambda_1^n + \lambda_2^n = -1$ for the $\emptyset$ component, where λ_1, λ_2 are the eigenvalues of R^0. When n is even, the equation has no real solutions, so the eigenvalues must be complex. In practice, this makes solving for the R matrices more difficult, as one has to optimize over complex tensors; in addition, one has to perform complex tensor network contractions, which can also significantly slow down calculations. This can be avoided, as suggested above, through a boundary tensor.

On the other hand, in the shallow circuit limit $L = 0$, the circuit reduces to random Pauli measurements. In this case, a trivial product state MPS suffices with

$$\mathsf{R}^j = \begin{cases} -1 & j = 0 \\ 3/2 & j = 1 \end{cases} \tag{4.15}$$

and $\mathsf{B} = 1$. Note that translation invariance is restored is both limits because $\mathsf{R}^0, \mathsf{R}^1$ does not depend on the site index i explicitly. For finite-depth circuits, $0 < L < \infty$, the brick wall structure breaks the translation invariance into the subgroup of two-site translations.

Also, as expected, the MPS representation of $|r\rangle$ both the shallow and deep circuit limits have low bond dimensions, although the underlying snapshot states have very different entanglement scalings (from area-law in $L = 0$ to volume-law in $L \to \infty$). The MPS representation of the reconstruction coefficients r_A allows us to apply the reconstruction map Eq. (4.11) in prediction algorithms in a scalable way, which we will see next. Even simply knowing the reconstruction coefficients exactly does not solve this problem because there are an extensive number of them. Therefore, the MPS representation of the reconstruction coefficients is key for classical shadow tomography (beyond the Pauli measurement) to be feasible on near-term quantum devices with large numbers of qubits; otherwise, we cannot reconstruct classical shadows efficiently.
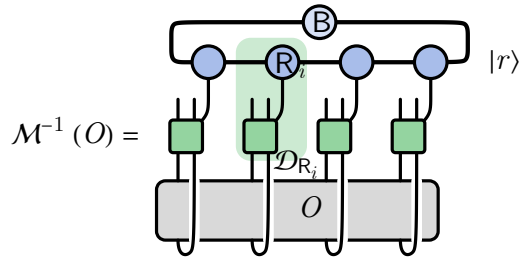
Figure 4.4. Tensor network representation of the reconstruction map $\mathscr{M}^{-1}(O)$ acting on a generic operator O, as defined in Eq. (4.17). Thick lines represent virtual bonds in the MPS. Each green box represents a dephasing channel $\mathscr{D}_{\mathsf{R}_i}$ at the qubit i in Eq. (4.16). The dephasing channels are correlated together by the MPS to form the reconstruction map.

Given the MPS representation of the reconstruction coefficient $|r\rangle$, the reconstruction map $\mathscr{M}^{-1}$ may be viewed as a matrix product operator (MPO). The local tensors are a generalized dephasing channel $\mathscr{D}_{\mathsf{R}_i}$ defined in terms of the MPS tensor for the reconstruction coefficients at a particular site i. It acts on a generic operator O as

$$\mathscr{D}_{\mathsf{R}_i}(O) := \mathsf{R}_i^0 \otimes (\mathbb{1}\,\mathrm{Tr}_i(O)/2) + \mathsf{R}_i^1 \otimes O \tag{4.16}$$

Then the reconstruction map on an arbitrary operator can be given as a super-operator defined in terms of the MPS tensors for the reconstruction coefficients, see Fig. 4.4,

$$\mathscr{M}^{-1} = 2^n \mathrm{Tr}_{\mathrm{virtual}} \left(\mathsf{B} \prod_{i=1}^{n} \mathscr{D}_{\mathsf{R}_i} \right) \tag{4.17}$$

where the trace and product are over the virtual indices of the MPS. When acting on a single-qubit Pauli operator P_i, where $P_i \in \{\mathbb{1}, X, Y, Z\}$, the generalized dephasing channel evaluates to

$$\mathscr{D}_{\mathsf{R}_i}(P_i) = (\delta_{P_i=\mathbb{1}} \mathsf{R}_i^0 + \mathsf{R}_i^1) \otimes P_i. \tag{4.18}$$

Hence, single-qubit Pauli operators are eigen-operators of the generalized dephasing channel, and Pauli strings are eigen-operators of the reconstruction map. On a arbitrary n-qubit Pauli string $P = \bigotimes_{i=1}^{n} P_i$, where $P_i \in \{\mathbb{1}, X, Y, Z\}$ the reconstruction map evaluates to

$$\mathscr{M}^{-1}(P) = 2^n \mathrm{Tr} \left(\mathsf{B} \prod_i (\delta_{P_i=\mathbb{1}} \mathsf{R}_i^0 + \mathsf{R}_i^1) \right) P. \tag{4.19}$$

The above formula is most useful for designing prediction algorithms. As we will see below, we can use the self-adjoint property of the reconstruction map to apply it to the operator whose expectation value we want to predict. This then allows us to take the expectation value of the transformed observable under the snapshot state, which has an efficient representation as a stabilizer state.

4.3.2 Pauli Estimation

Suppose we wish to estimate a Pauli observable $P = \bigotimes_{i=1}^{n} P_i$, where $P_i \in \{\mathbb{1}, X, Y, Z\}$. Efficient classical shadow tomography schemes for the prediction of such operators only exist using random Pauli measurements and become quickly infeasible as the weight of the Pauli string grows. Here, we show how the scheme can be extended to finite-depth circuits given the reconstruction coefficients. In practice, the expectation value is given by an average over the snapshot states $\hat{\sigma}_j$, $j = 1 \cdots M$.

$$\langle P \rangle = \frac{1}{M} \sum_{m=1}^{M} \mathrm{Tr}(P \mathcal{M}^{-1}(\hat{\sigma}_m)). \tag{4.20}$$

Since the measurement channel $\mathcal{M}$ defined in Eq. (4.1) is self-adjoint, so as its inverse

$$\mathrm{Tr}(P \mathcal{M}^{-1}(\hat{\sigma})) = \mathrm{Tr}(\mathcal{M}^{-1}(P)\hat{\sigma}), \tag{4.21}$$

we may move $\mathcal{M}^{-1}$ onto P, and apply Eq. (4.19) to simplify $\mathcal{M}^{-1}(P)$, then Eq. (4.20) becomes

$$\langle P \rangle = 2^n \mathrm{Tr}\left(B \prod_i (\delta_{P_i=\mathbb{1}} R_i^0 + R_i^1) \right) \langle P \rangle_\sigma, \tag{4.22}$$

where $\langle P \rangle_\sigma = \frac{1}{M} \sum_{m=1}^{M} \mathrm{Tr}(\hat{\sigma}_m P)$. This is a key result of this work. It enables us to extend classical shadow tomography beyond the Pauli measurement limit without losing the classical post-processing efficiency. Every term above is efficiently computable — the trace is over a product of n low-dimensional matrices and can be thought of as an inner product over MPS states, and the expectation value of the Pauli observable P on a stabilizer state $\hat{\sigma}_j$ can also be computed efficiently according to the Gottesmann-Knill theorem [2, 77], which is implemented in [99].

4.3.3 Generic Observable and Fidelity Estimation

What about more general operators? To effectively utilize the results of the previous section for scalable prediction, we need local tensor network descriptions of operators. One approach is to view the general operator O as a state in the Pauli basis i.e.

$$(O)_P = \langle P|O \rangle = \text{Tr}(PO). \tag{4.23}$$

The operator O may be viewed as a state in the 4^n dimensional Hilbert space with basis states labeled by the different Pauli strings. Given this state description, we can now assign onsite matrices O_i^j where i labels the site and the j labels the Pauli at site i in the matrix-product expansion of the operator. If O has little operator entanglement, then the corresponding matrices O_i^j will also have a small bond dimension, and the decomposition can be considered efficient.

$$(O)_P = \text{Tr} \left(\prod_i (O_i^0 \delta_{P_i=\mathbb{1}} + O_i^1 \delta_{P_i=X} + O_i^2 \delta_{P_i=Y} + O_i^3 \delta_{P_i=Z}) \right). \tag{4.24}$$

One context in which more general operator estimation is useful is when estimating fidelity. The fidelity $F(\rho,\rho')$ provides a way to characterize the closeness of one state ρ to another ρ'. It is defined as $F(\rho,\rho') = (\text{Tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}})^2 \in [0,1]$. The greater the value of F, the closer the two states are. If any of the states ρ or ρ' is pure, their fidelity simplifies to

$$F(\rho,\rho') = \text{Tr}(\rho\rho') \tag{4.25}$$

By viewing $F(\rho,\rho')$ as the expectation value of the operator $O = \rho'$ on the state ρ, the fidelity estimation amounts to predicting a low-rank, non-local observable $O = \rho'$ using the means described in this paper. For fidelity estimation, one might be worried that the operator entanglement of the reference state ρ' is too large to represent efficiently as MPS. Though this may seem like an obstacle, ground states, low-entangled states, and stabilizer states of shallow, local

126

circuits, all have an efficient MPS description of the form above. The last one, in particular, is useful in fidelity estimation, since we will use the MPS description of snapshot states to convert the trace into a tensor network. We will lay out the general formula given the matrix-product representation of ρ' per Eq. (4.24), and leave details about how to construct the MPS for $\hat{\sigma}$ in the Appendix C.1. A straightforward application of Eq. (4.19) gives the desired formula.

$$F(\rho,\rho') = \mathbb{E}_{\hat{\sigma}}\left[\sum_P (\hat{\sigma})_P (\rho')_P \operatorname{Tr}\left(B \prod_i (\delta_{P_i=\mathbb{1}} R_i^0 + R_i^1)\right)\right] \tag{4.26}$$

Using the MPS representation for the Pauli coefficients of $\hat{\sigma}$ and ρ', the summation above can be viewed as a tensor network, as shown in Fig. 4.5. The idea is to take the MPS matrix for $\hat{\sigma}, \rho'$ and r and contract them along an appropriate vertex tensor $u = \bigotimes_{i=1}^n u_i$. If we label the indices of u_i as $(u_i)_{P_i, P_i', j}$, where P, P' are Pauli strings in the basis expansion of $\hat{\sigma}, \rho'$, respectively, and j labels the local reconstruction MPS matrix R_i^j, then

$$(u_i)_{P_i, P_i', j} = \delta_{P_i=P_i'}(\delta_{P_i \neq \mathbb{1}} \delta_{j=1} + \delta_{P_i=\mathbb{1}}). \tag{4.27}$$

Note that this procedure applies also for predicting generic local operators O, by expanding the operator in its Pauli basis representation first. One only needs to replace ρ' by O in Eq. (4.26), then $\langle O \rangle = \operatorname{Tr}(\rho O)$ can be evaluated by a similar tensor network of the same structure in Fig. 4.5. The procedure outlined above only relies on an efficient Pauli basis representation for the state or operator in question.

4.3.4 Locally-Scrambled Shadow Norm

Apart from estimating the expectation value of an observable O from the classical shadow data, we also want to bound the variance of our estimation. This is provided by the operator shadow norm $\|O\|_{\text{shd}}^2$ introduced in Ref. [102]. In the following, we will focus on the case that O is a *traceless* operator, i.e. $\operatorname{Tr} O = 0$. Let $\operatorname{var}\langle O \rangle$ be the variance of the estimation

$$F\left(\rho,\rho'\right) = \mathbb{E}_{\hat{\sigma}}$$
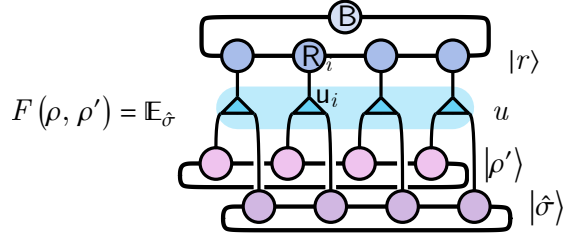
Figure 4.5. Tensor network representation of the fidelity $F(\rho,\rho')$ in Eq. (4.26). The vertex tensor u_i is defined in Eq. (4.27). $|\rho'\rangle$ (or $|\hat{\sigma}\rangle$) is the vector representation of the density matrix ρ' (or the snapshot state $\hat{\sigma}$) in the Pauli operator basis.

$1/M \sum_{m=1}^{M} \text{Tr}(O\mathcal{M}^{-1}(\hat{\sigma}_m))$ obtained from averaging M samples (for finite M). The variance of the mean decays as

$$\text{var}\langle O\rangle \lesssim \|O\|_{\text{shd}}^2/M, \tag{4.28}$$

where $\|O\|_{\text{shd}}^2 = \max_\rho \mathbb{E}_{\hat{\sigma}\in\mathscr{E}_\sigma} 2^n \text{Tr}(\hat{\sigma}\rho)(\text{Tr}(\hat{\sigma}\mathcal{M}^{-1}(O)))^2$ is the shadow norm. A more precise statement of Eq. (4.28) and its proof can be found in Ref. [102].

The conventional shadow norm $\|O\|_{\text{shd}}^2$ is defined as a maximum over all possible states ρ of the quantum system. This definition requires information about the third moment of the snapshot state ensemble $\mathscr{E}_\sigma$, which could take some effort to compute in general. In the context of locally-scrambled ensembles, Ref. [98] introduces an alternative version of the shadow norm, called the *locally-scrambled shadow norm*, denoted as $\|O\|_{\mathscr{E}_\sigma}^2$. It only requires the second moment of $\mathscr{E}_\sigma$, which can be conveniently read out from the entanglement feature $|W_{\mathscr{E}_\sigma}\rangle$ computed already. The key modification in the definition is to replace the maximization $\max_\rho$ by an average of ρ over its locally scrambled ensemble $\mathscr{E}_\rho = \{V\rho V^\dagger | V = \bigotimes_i V_i \in \text{U}(2)^{\otimes n}\}$ with contains all states related to ρ by local basis transformations V. The locally-scrambled shadow norm is defined as

$$\begin{aligned}
\|O\|_{\mathscr{E}_\sigma}^2 &:= \mathbb{E}_{\rho\in\mathscr{E}_\rho} \mathbb{E}_{\hat{\sigma}\in\mathscr{E}_\sigma} 2^n \text{Tr}(\hat{\sigma}\rho)(\text{Tr}(\hat{\sigma}\mathcal{M}^{-1}(O)))^2 \\
&= \mathbb{E}_{\hat{\sigma}\in\mathscr{E}_\sigma} (\text{Tr}(\hat{\sigma}\mathcal{M}^{-1}(O)))^2,
\end{aligned} \tag{4.29}$$

where the dependence on ρ drops from the result due to the locally scrambled property of $\mathscr{E}_\rho$. So $\|O\|^2_{\mathscr{E}_\sigma}$ is only a function of the randomized measurement scheme specified by the prior snapshot ensemble $\mathscr{E}_\sigma$ (defined in Eq. (4.7)) and the observable O in question.

Using the definition of the measurement channel $\mathscr{M}$ in Eq. (4.1) and the reconstruction map $\mathscr{M}^{-1}$ in Eq. (4.11), Eq. (4.29) can be further reduced to

$$\begin{aligned}
\|O\|^2_{\mathscr{E}_\sigma} &= 2^{-n} \operatorname{Tr}(O\mathscr{M}^{-1}(O)) \\
&= \sum_A 2^{|A|-n} r_A (W_O)_A,
\end{aligned} \tag{4.30}$$

where r_A is the reconstruction coefficient given by the solution of Eq. (4.9). $(W_O)_A$ is the entanglement feature of the operator O, which follows the definition of the entanglement features of a state ensemble in Eq. (4.3), with the density matrix ρ simply replaced by the operator O:

$$(W_O)_A = \langle A|W_O\rangle := \operatorname{Tr}_A(\operatorname{Tr}_{\bar{A}} O)^2. \tag{4.31}$$

It is also possible to represent $|W_O\rangle$ as an MPS, similar to the state entanglement feature in Eq. (4.6),

$$(W_O)_A = \operatorname{Tr} \prod_i (\delta_{i\notin A} \mathsf{W}^0_{O,i} + \delta_{i\in A} \mathsf{W}^1_{O,i}). \tag{4.32}$$

This defines the MPS tensor $\mathsf{W}^j_{O,i}$ for any operator O. Hence, for a generic operator O, the shadow norm $\|O\|^2_{\mathscr{E}_\sigma}$ in Eq. (4.30) can be evaluated as a simple overlap given the MPS representations of two objects: the MPS matrices $\mathsf{R}^0, \mathsf{R}^1$ for the reconstruction coefficients $|r\rangle$ is given in Eq. (4.13); and the MPS matrices $\mathsf{W}^0_O, \mathsf{W}^1_O$ for $|W_O\rangle$ as in Eq. (4.32). Fig. 4.6 illustrates the tensor network contraction, which can be computed efficiently with complexity linear in the system size.

The locally-scrambled shadow norm is upper bounded by the conventional shadow norm as $\|O\|^2_{\mathscr{E}_\sigma} \leq \|O\|^2_{\mathrm{shd}}$. It plays a similar role in variance estimation,

$$\operatorname{var}\langle O\rangle \simeq \|O\|^2_{\mathscr{E}_\sigma}/M. \tag{4.33}$$
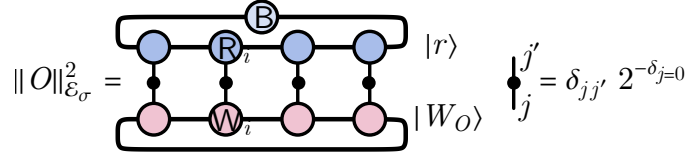
Figure 4.6. Tensor network representation of the locally-scrambled shadow norm $\|O\|_{\mathcal{E}_\sigma}^2$ defined in Eq. (4.30).

The shadow norm $\|O\|_{\mathcal{E}_\sigma}^2$ has two complementary physical meanings: (i) it characterizes the estimation variance $\mathrm{var}\langle O\rangle$ given a fixed number of samples M (operators with larger shadow norms will have proportionally higher variances), (ii) it determines the sample complexity given the variance level (i.e. $M \sim \|O\|_{\mathcal{E}_\sigma}^2/\varepsilon^2$ samples are needed in order to control the variance below a given threshold $\mathrm{var}\langle O\rangle \le \varepsilon^2$ set by a small ε). For Pauli estimation $O = P$, the operator entanglement feature $|W_P\rangle$ can be computed easily from its support. We define the support of a Pauli operator P as the subset of sites on which it acts non-trivially, denoted as $\mathrm{supp}\, P = \{i|P_i \ne \mathbb{1}\}$. Define the *operator weight* $k = |\mathrm{supp}\, P|$ as the size of the support of P. We find that the operator entanglement feature of a Pauli string P is simply given by

$$(W_P)_A = \delta_{\mathsf{supp}\ P \subseteq A} 2^{2n-|A|}. \tag{4.34}$$

The corresponding MPS matrices for Pauli observables $\mathsf{W}_{P,i}^0, \mathsf{W}_{P,i}^1$ take different values depending on if site i is inside the support of P or not. Given the general MPS form in Eq. (4.32), it is easy to check that

$$\mathsf{W}_{P,i}^0 = 4\delta_{i\notin\mathsf{supp}\ P}, \qquad \mathsf{W}_{P,i}^1 = 2. \tag{4.35}$$

Nonetheless, the matrices have bond dimension one because of the tensor product structure of P.

Plugging in (4.34) into (4.30), we obtain the shadow norm of a Pauli string as a function of the reconstruction coefficients,

$$\|P\|_{\mathcal{E}_\sigma}^2 = 2^n \sum_{A|A\supseteq\mathsf{supp}\ P} r_A. \tag{4.36}$$

130

Using the solution of $|r\rangle$ in Eq. (4.12), we have

$$\|P\|^2_{\mathscr{E}_\sigma} = \frac{(-3)^k}{\sum_{B|B\subseteq\text{supp } P}(-2)^{|B|}(W_{\mathscr{E}_\sigma})_B}, \tag{4.37}$$

where $k = |\text{supp } P|$. This result shows how the shadow norm of a Pauli operator P depends on both its support and the entanglement feature of the snapshot state ensemble.

In the following, we will directly evaluate the locally-scrambled shadow norm $\|O\|^2_{\mathscr{E}_\sigma}$ for several choices of the circuit depth L following Eq. (4.30). For notational convenience, we will treat $|W_O\rangle$ as translation-invariant unless evaluating the shadow norm for specific operator types that explicitly break translation symmetry, e.g. Pauli operators.

$L = 0$

Here, the unitary ensemble is simply Pauli measurement. The vector $|r\rangle$ is a product states, and the reconstruction coefficients are given in (4.15) as $R^0 = -1, R^1 = 3/2$. In terms of the matrices for $|W_O\rangle$, the expression for the shadow norm reduces the trace of a matrix power:

$$\|O\|^2_{L=0} = \text{Tr}\left[\left(\frac{-W^0_O + 3W^1_O}{2}\right)^n\right]. \tag{4.38}$$

We can see that in this limit, the shadow norm is closely related to the locality of the operator, since if O is the identity on some site i, then the corresponding contribution to the shadow norm is a factor of 1, i.e. it doesn't affect the shadow norm.

Consider the case $O = P$. In the translation-symmetry breaking version of Eq. (4.38), the power of n breaks into individual factors for each site, with each term in the product depending on the operator entanglement feature matrix at that site:

$$\|P\|^2_{L=0} = \prod_{i=1}^{n}\left(\frac{-W^0_{P,i} + 3W^1_{P,i}}{2}\right). \tag{4.39}$$

Using Eq. (4.35), the product contributes a factor of 1 where P_i is identity and a factor of 3 where

131

P_i is not. Hence we see that the locally-scrambled shadow norm for P grows exponentially in the weight $k = |\mathrm{supp}\, P|$,

$$\|P\|^2_{L=0} = 3^k. \tag{4.40}$$

This result agrees with the conventionally defined shadow norm. It can also be seen to agree with (4.37) since the entanglement feature state of the unitary ensemble $|W_{\mathscr{E}_\sigma}\rangle$ is a trivial product state in the $L=0$ case.

Furthermore, although the locally scrambled shadow norm is less than the conventionally defined shadow norm, for locally-scrambled unitary ensembles, it provides accurate bounds on the single-shot variance of Pauli observables [98]. Support for this claim is given in Sec. 4.4.

$L = 1$

Another important case we may check is $L = 1$. In this case, the ensemble is a product of two-local, two-design unitary gates. Each gate effectively scrambles site $2i$ with its neighbor at $2i + 1$, and generates no entanglement outside this unit cell. Within each unit cell, the unitary is Clifford random. Thus the MPS for reconstruction coefficients in each block for the $L = 1$ case is given by (4.14). Therefore,

$$\|O\|^2_{L=1} = \mathrm{Tr}\left[\left(\frac{-(\mathsf{W}^0_O)^2 + 5(\mathsf{W}^1_O)^2}{4}\right)^{n/2}\right]. \tag{4.41}$$

Consider next the case $O = P$, and suppose n is even for simplicity. Then the translation-symmetry breaking version of (4.41) for a Pauli operator is

$$\|P\|^2_{L=1} = \prod_{i=1}^{n/2}\left(\frac{-\mathsf{W}^0_{P,2i-1}\mathsf{W}^0_{P,2i} + 5\mathsf{W}^1_{P,2i-1}\mathsf{W}^1_{P,2i}}{4}\right). \tag{4.42}$$

There are two types of factors depending on the support of the operator P in that unit cell. If the Pauli operator P is trivial in the unit cell, then the factor is one. Otherwise, if the Pauli operator P has any support in the unit cell, then the factor is five. Hence, for a contiguous Pauli operator

of weight k, there is a (staggered) exponential growth in the shadow norm due to the unit cell structure of the brick wall circuit i.e.

$$\|P\|_{L=1}^2 = 5^{\lceil k/2 \rceil}. \tag{4.43}$$

Thus, as the operator weight k grows, the shadow norm, and hence the number of samples required for reliable prediction, is already exponentially smaller for $L = 1$ than $L = 0$, since $\frac{\|P\|_{L=1}^2}{\|P\|_{L=0}^2} \sim (\frac{\sqrt{5}}{3})^k$. We find that this pattern, that longer operators can be better estimated by deeper circuits, continues for $L > 1$ in Sec. 4.4. For a given operator weight k, there is an ideal short circuit depth $L^* \geq 0$ such that the shadow norm and, therefore, the singles-shot variance are minimized. This shows that classical shadow tomography beyond Pauli measurements has practical applications since we can use circuits of varying depths to minimize the number of samples one has to take.

When $L \geq 2$, the shadow norm is more difficult to evaluate analytically. This is because both the entanglement features of the state ensemble and the reconstruction coefficients are more complicated. While the cases $L = 0, 1, \infty$ can be helpful, for $1 < L < \infty$, we resort to numerical evaluation of the shadow norm. See Sec. 4.4. We find that the shadow norm is a reliable indicator of the estimation variance in the case where O is a tensor product of local operators.

$L = \infty$

When $L \to \infty$, there are only two non-zero reconstruction coefficients $r_\emptyset = -1$ and $r_\Omega = 1 + 2^{-n}$, which correspond to the empty set and the total system, respectively. Then the shadow norm $\|O\|_{L=\infty}^2$ for a traceless Hermitian operator reduces to

$$\|O\|_{L=\infty}^2 = (1 + 2^{-n}) \operatorname{Tr}(O^2). \tag{4.44}$$

This is reminiscent of the conventional shadow norm introduced in [102], as both are proportional to $\operatorname{Tr}(O^2)$, indicating that the locally-scrambled shadow norm in the deep circuit

limit also depends strongly on the rank of the operator. However, the locally-scrambled shadow norm has a smaller proportionality factor since the conventional shadow norm is defined by taking a maximum over all state ρ. We also note that afterwards in [11], some closed form analytic expressions were developed for the reconstruction map in brickwall circuits that are consistent with these results.

4.4 Numerical Demonstrations

In this section, we demonstrate the prediction algorithms for various system sizes based on numerical simulation[1]. We evolve the system using random two-local Clifford gates arranged in a brick wall pattern with periodic boundary conditions. The depth L of the circuit refers to the number of layers of Clifford gates. For example, for $L = 2$, we apply one even layer and then one odd layer of independently random two-qubit local Clifford gates. After the Clifford circuit evolution, we then measure the qubits and store the resulting snapshot state $\hat{\sigma} = U^{\dagger}|b\rangle\langle b|U$ for later use in predicting various quantities. For demonstration purposes, we collect the classical snapshots by numerically simulating the randomized measurement. Suppose the underlying original state ρ is a stabilizer state; the simulation can be efficiently carried out on a classical computer using the stabilizer table algorithm, thanks to the Gottesmann-Knill theorem.

In our numerical demonstrations, we consider two example initial states: the cluster state ρ_{ZXZ} and the Greenberger-Horne-Zeilinger (GHZ) state ρ_{GHZ}. They can be specified by their stabilizer groups,

$$\mathscr{G}_{\text{ZXZ}} = \langle Z_1 X_2 Z_3, \cdots, Z_n X_1 Z_2 \rangle, \tag{4.45}$$

$$\mathscr{G}_{\text{GHZ}} = \langle Z_1 Z_2, \cdots, Z_{n-1} Z_n, \prod_{i=1}^{n} X_i \rangle, \tag{4.46}$$

Each stabilizer state $\rho = |\mathscr{G}|^{-1} \sum_{g \in \mathscr{G}} g$ is defined as the invariant state of the corresponding stabilizer group $\mathscr{G}$.

---

[1] Most of the numerical codes used in this paper can be found at this github repository.

In the following figures, the initial state will be indicated using squares (for ρ_{ZXZ}) and triangles (for ρ_{GHZ}), respectively. The cluster state ρ_{ZXZ} is the ground state of a gapped, local, stabilizer Hamiltonian, and therefore carries strictly area-law entanglement. The GHZ state ρ_{GHZ} is the symmetric ground state of the Ising model, which contains long-range entanglement and is more challenging for prediction based on the randomized measurement on shallow local circuits. For example, the extensive generators (i.e. $\prod_i X_i$ in the $\mathscr{G}_{GHZ}$ stabilizer group) will be hard to probe by local circuits of a finite depth L, because the generators of the snapshot state are at most $\sim 2L$ in length.

We then use the reconstruction coefficients, together with the algorithms outlined above, to perform prediction using the snapshot states. The reconstruction coefficients are calculated using PyTorch, specifically the AdamW optimizer, by minimizing the loss function $\mathscr{L}$ in Eq. (4.10). The minimization procedure is continued until $\mathscr{L} \leq 10^{-3}$.

We would like to briefly comment on the flexibility of this approach with respect to different circuit evolutions. As explained in Fig. 4.1, the input for determining the reconstruction coefficients is the entanglement features of the underlying unitary ensemble. One might be wondering if the brick-wall Clifford unitary circuit is the only specialized case which admits efficient entanglement feature dynamics. For example, one may be interested doing tomography using other locally-scrambled unitary ensembles, such as random Hamiltonian-generated evolution, which was proposed in [98]. It was found in prior work that locally-scrambled dynamics generated by continuously scrambling gates is also well described by a simple $D = 2$ MPS ansatz [5]. Similar efficiency was found in [63]. It is believed that this locally-scrambled purity dynamics is generically efficiently representable as MPS away from finely tuned points like the measurement-induced entanglement transition. Therefore, the shallow-circuit tomography scheme presented in this paper can be adapted to other circuits which are also locally-scrambled at the two-design level. We will leave the analysis of the tomography scheme in these more general contexts to future work.

Lastly, a similar approach using MPS to do efficient shallow circuit tomography was

developed concurrently to ours in [20]. In this paper, the Pauli weight instead of the entanglement feature is computed. Since these quantities are related by local basis transformation in the entanglement-feature Hilbert space, we suspect that the efficiency of their approach will be comparable to ours.

4.4.1 Pauli Estimation

To demonstrate the efficiency of our approach in estimating Pauli observables, we will take $Z^{\otimes k} := \bigotimes_{i=1}^{k} Z_i$ as the target observable without loss of generality. The operator weight (operator size) k can be adjusted. Since the randomized measurement schemes are invariant under local Clifford transformations, any Pauli string P of the same weight k will exhibit the same sample efficiency (as the shadow norm $\|P\|_{\mathscr{E}_\sigma}^2$ only depends on the support of P). Nevertheless, the expectation value $\langle P \rangle$ will still depend on the choice of the observable P and the underlying state. On the GHZ state, the expectation value of $Z^{\otimes k}$ is 1 or 0 depending on if k is even or odd, respectively. On the cluster state, it is 0 for all $k \geq 1$.
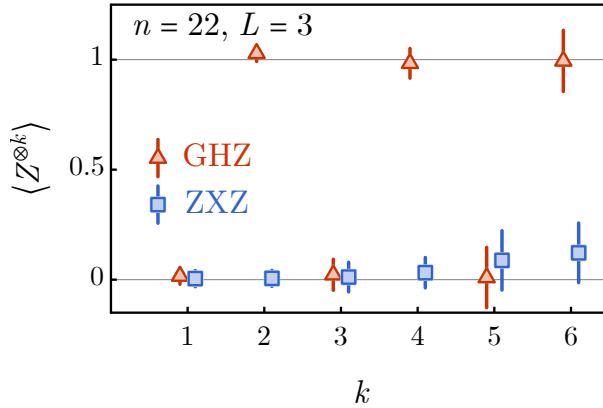
Figure 4.7. The estimated Pauli observable $Z^{\otimes k} := \bigotimes_{i=1}^{k} Z_i$ of different weight k with the underlying state ρ being the cluster (squares) and GHZ (triangles) states. Each point is based on 50000 measurement samples collected from a brick-wall circuit of depth $L = 3$ measurement protocol. The system size is $n = 22$ qubits. The error bar indicates two standard deviations of the estimated value.

In Fig. 4.7, we demonstrate that our method can reliably predict the expectation value of the Pauli string observable $Z^{\otimes k}$ on a system of $n = 22$ qubits via randomized measurements on

the random Clifford circuit of $L = 3$ layers. Our numerical result is obtained by first simulating the randomized measurement on a classical computer and then processing the data using the algorithm described in Sec. 4.3.2. Our classical post-processing approach provides unbiased estimations of $Z^{\otimes k}$ on both the cluster state (blue squares) and the GHZ state (red triangles) for various operator weights k, although the estimation variance is growing with k (which will be analyzed soon). Here we only demonstrate the case of circuit depth $L = 3$, but our method is applicable for other circuit depths, and the quality of estimation remains similar.

The variance of the estimation is associated with the sample complexity. A larger variance (a larger shadow norm) means more samples are needed to reduce the variance to the desired accuracy level. The traditional classical shadow tomography with Pauli measurements ($L = 0$) quickly requires an extensive number of samples to estimate a weight-k Pauli string observable (as the shadow norm $\|P\|_{L=0}^2 = 3^k$ grows exponentially with k). However, with shallow circuits $L \geq 1$, we can easily achieve the same accuracy with far fewer samples. We have analytically proved this statement in Sec. 4.3.4 for $L = 1$ by calculating the locally-scrambled shadow norm. For large L, it is difficult to obtain a closed-form analytic expression for the shadow norm, but we can numerically investigate the shadow norm as a function of L according to Eq. (4.30) using tensor network techniques. Our results are presented in Fig. 4.8.

We first verify that the locally-scrambled shadow norm $\|Z^{\otimes k}\|_L^2$ computed from Eq. (4.30) correctly captures the variance $\mathrm{var}\, Z^{\otimes k}$ of the estimated expectation value observed in our numerical simulation. Based on Eq. (4.33), we expect the variance and the shadow norm to match each other, as $\mathrm{var}\langle Z^{\otimes k}\rangle = \|Z^{\otimes k}\|_L^2/M$, where M is the number of samples. This is indeed the case as shown in Fig. 4.8. The numerically simulated variance and the computed shadow norm are in good agreement as we vary L and k, even as k approaches a quarter of the system size. This confirms the correctness of our approach to computing the shadow norm. The shadow norm is useful as it can give us an idea of the number of samples needed to control the estimation variance to the desired level, even before we start performing the classical shadow tomography experiment.
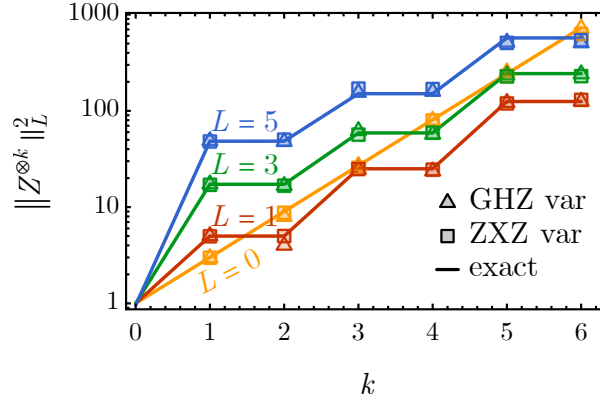
Figure 4.8. The locally-scrambled shadow norms for circuit depth $L = 0, 1, 3, 5$ (yellow, red, green, and blue lines, respectively), with $n = 22$ qubits. We also plot the corresponding variances over all samples for the cluster (squares) and GHZ (triangles) states. These agree with the shadow norms. The staggered behavior of the shadow norm for $L > 0$ is because these circuits break translation invariance.

As shown in Fig. 4.8, the shadow norm $\|Z^{\otimes k}\|_L^2$ always grows with the operator weight k, meaning that we always need to perform more measurements to estimate longer Pauli strings accurately. However, the growth rates of shadow norms are different for different circuit depths L. For example, the $L = 1$ shadow norm increases with k slower than that of $L = 0$, meaning that the $L = 1$ circuit will be more advantageous for probing Pauli string operators of sufficiently large weight k.
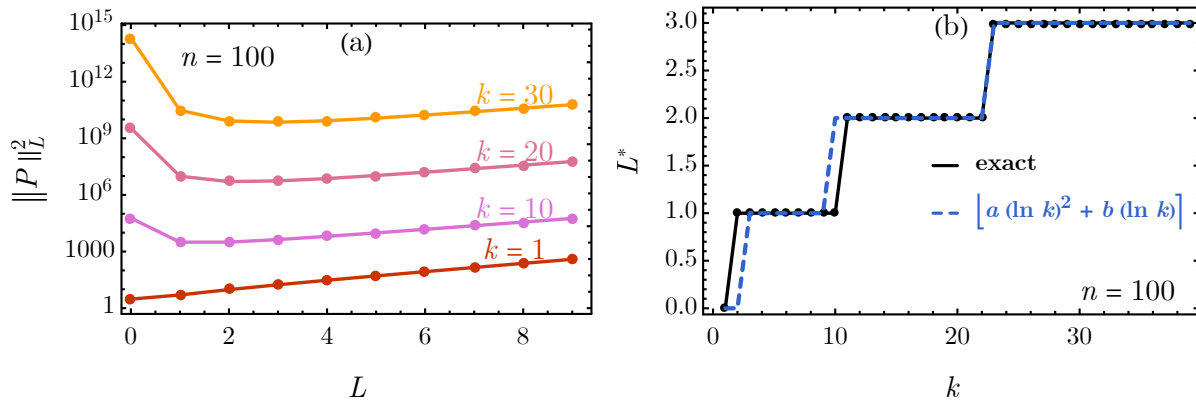
Figure 4.9. (a) The locally-scrambled shadow norm for a general Pauli operator P of size k v.s. the circuit depth L of the randomized measurement scheme. (b) The optimal value of circuit depth L^* to predict a contiguous Pauli operator of weight k. We use a system size of $n = 100$ qubits here. The results do not change as we further increase n.

To study this trend more quantitatively, we use (4.37) to compute the shadow norm of Pauli operators for various weights k under various circuit depths L. The result is presented in Fig. 4.9(a). Given the Pauli operator size k, it is clear that there is an optimal circuit depth L^* that minimizes the shadow norm. We conjecture that L^* can be fitted by an empirical formula

$$L^* \sim a(\ln k)^2 + b(\ln k).$$ (4.47)

The formula fits the numerically found L^* nicely with parameters $a = 0.14, b = 0.35$, as shown in Fig. 4.9(b), where the rounded value of the fitting function is shown since L^* must always be an integer. This confirms the sub-polynomial growth in the optimal circuit depth L^* as proposed in Eq. (4.47). The result indicates that a relatively shallow circuit can efficiently optimize the sample complexity for classical shadow tomography of large Pauli strings, which demonstrates the advantage of our shallow-circuit measurement approach. The logarithmic scaling of the optimal circuit depth was confirmed afterwards in [108]. In this work, the authors were able to derive a bound on the shadow norm by studying the evolution of the weight distribution of contiguous operators under the random brick-wall circuit.

We would like to point out that for estimating the Pauli string with a fixed operator weight k, the shadow norm remains constant with the system size n, meaning that classical shadow tomography is scalable for this task. Our approach of using shallow-circuit measurements enables further improvement in the sample complexity compared to the Pauli measurement without losing the scalability and feasibility on near-term quantum devices.

4.4.2 Fidelity Prediction

Fidelity estimation is another important task in quantum information. We can also compare the fidelity of ρ_{GHZ} and ρ_{ZXZ} against various states using classical shadow tomography on shallow circuits. A high fidelity $F = 1$ ensures that the state produced in the lab is close to the desired state, whereas a low fidelity means that the state produced in the lab has very

little overlap with the desired state. To demonstrate our approach, we simulate the shallow-circuit randomized measurement on a classical computer and then use the algorithm outlined in Sec. 4.3.3 to process the measurement data collected from the simulation. We estimate the fidelity of the classical shadow reconstructed state with the underlying original state and show that our approach successfully reconstructs the original quantum state from measurement outcomes with high fidelity. Although the shallow-circuit classical shadow tomography still has exponential sample complexity for the fidelity estimation task, increasing the circuit depth a bit can significantly reduce the base of this exponential complexity, which makes our approach useful for quantum state tomography on near-term quantum devices.
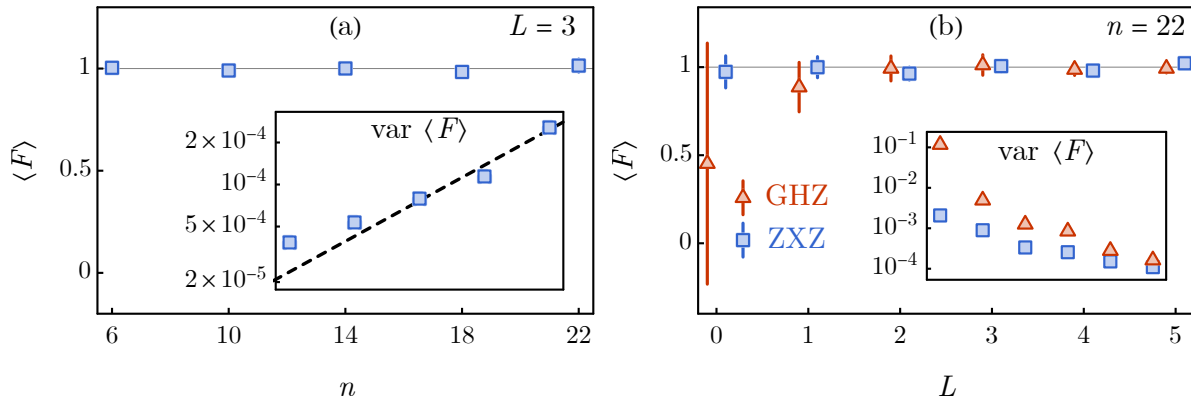
Figure 4.10. (a) The fidelity and its variance as a function of system size, for systems of size $n = 6, 10, 14, 18, 22$ for a fixed circuit depth of size $L = 3$. We use the cluster state as the reference state. Each point is the average of 50000 measurement samples. The mean fidelity is tightly centered around one, which is expected. The growth of the sample variance is indicated in the inset. (b) The fidelity and its variance versus circuit depth L for a system of $n = 22$ qubits. The two states are the cluster state (squares) and the GHZ state (triangles). We use a median-of-means average to estimate the fidelity, using 50000 total samples broken up into 12 groups. The error bar corresponds to two standard deviations. The inset shows that the variance decays with L rapidly.

First, we compare the fidelity of the cluster state with itself as a function of the qubit number n for $n = 6, 10, 14, 18, 22$, using a circuit depth of $L = 3$, using 50000 samples. The result is shown in Fig. 4.10(a). We find the reconstructed states have close-to-one fidelities with negligible error bars up to 22 qubits. We also observe a (rather slow) exponential growth in the

140

variance of the fidelity, in agreement with Ref. [98]. In that paper, a scaling form is proposed with var $F \propto \exp(cn/(L+1)^{\alpha})$. Fixing $\alpha = 0.72$, per Ref. [98], we can fit the parameter c based on our data. We find that the fitted value $c = 0.4$ agrees with the value $c = 0.47 \pm 0.08$ previously observed in Ref. [98]. This empirical relation suggests that although the variance of fidelity estimation (and hence the sample complexity) grows exponentially with the number of qubits n, increasing the circuit depth L in the shallow circuit regime can significantly reduce the variance (by making the exponential growth very slow).

Next, we study the efficiency of the protocol as a function of the circuit depth L for a system of size $n = 22$, as shown in Fig. 4.10(b). We plot the fidelity of both the cluster state and the GHZ states with themselves and find that both are centered around one, as expected. For both states, we observe a rapid decline in the variance of the sample fidelity as we increase circuit depth. However, we notice that the GHZ state has a much higher sample variance. This is likely due to the fact that the GHZ state has long-range correlations, unlike the cluster state which only has finite-range correlations. Thus quasi-local measurements will not be very efficient in probing the GHZ state. For example, at $L = 0$ (Pauli measurements), the reconstructed state has very low fidelity $F < 0.5$ and very large sample variance $\text{var}F \sim 1$ for the GHZ state. We see a sharp (more than an order of magnitude) decline in the sample variance for $L = 1$, and a continued exponential decline as we increase L. By $L = 3$, the fidelity approaches one for both states within a variance less than 10^{-3}. This is because a deeper scrambling circuit will enable the final computational basis measurements to probe more complex (higher weight) operators in the original basis, which benefits the fidelity estimation. This demonstrates the great advantage of entanglement-assisted shallow-circuit measurements. The rapid decline of the variance in fidelity versus circuit depth L makes it promising to use classical shadow tomography for fidelity estimation with shallow circuits achievable on near-term quantum devices.

4.5 Conclusion and Open Questions

In this paper, we demonstrate a novel approach to predicting expectation values of observables using shallow-depth circuit classical shadow tomography. Our work extends the well-known Pauli measurement and Clifford measurement protocols to the case of finite-depth brick wall circuits, which are available on near-term devices on many qubits. This allows us to use shallow circuit measurement to "approximate" the Clifford measurement limit to do fidelity estimation efficiently. It also allows us to use the scrambling properties of circuits with $L > 0$ to more efficiently probe large-weight operators, e.g. long Pauli strings. The approach that we take is based on our entanglement features formalism, which allows us to simply represent the purity data as MPS and consequently determine the reconstruction coefficients. Given the reconstruction coefficients and that they can be represented with a low-bond dimension MPS, we propose a scalable approach to predict many properties of quantum states.

There are still several questions that remain and that we wish to explore in future works. Firstly, is there an exact and efficient representation of the reconstruction coefficients as MPS? In our paper, we determine the reconstruction coefficients by learning their MPS tensors via gradient descent on the consistency equation Eq. (4.9). This procedure only needs to be done once, but it is bound to produce some error that can propagate through the prediction algorithm. In Ref. [28], the exact, individual reconstruction coefficients were determined. In practice, we want to be able to organize all 2^n coefficients efficiently — is there some way to represent the exact individual coefficients scalably? Such an approach would also give us insight into how the bond dimension or complexity of the reconstruction map evolves with circuit depth. We know that it is small in the shallow circuit limit and also small at $L = \infty$. Is it possible that the bond dimension remains small and finite for all values of the circuit depth L? Furthermore, is there some way to generalize the results in this paper to higher dimensions? For example, do the entanglement feature states and reconstruction coefficients now a low-dimensional PEPS structure?

Moreover, the prediction procedures described in Sec. 4.3 assume that the operator whose expectation value we want to predict has an efficient Pauli basis expansion. This means that the operator with a large amount of operator entanglement will be difficult to predict. Furthermore, this procedure is limited by the growth in operator entanglement of the snapshot states over time. The snapshot states' bond dimension grows exponentially in L, which can be seen from the fact that the typical size of the generators of the snapshot states grows linearly in L. This means this procedure will be costly for deep circuits. What approach can we take in this case?

In addition, we only considered measuring quantities that depend on the first moment of ρ i.e. $\text{Tr}(O\rho)$. We could also consider quantities that depend on higher moments of ρ, e.g. $\text{Tr}(O\rho \otimes \rho)$, where O is an operator on the doubled Hilbert space. How can we generalize the approach developed here to expectation values based on higher moments of ρ?

In this paper, we only considered a particular class of circuit ensembles formed by varying the circuit depth L. However, the procedure we outlined is independent of the particular circuit geometry. How might other geometries be useful in other contexts? For example, instead of considering depth L circuits, we could also consider circuits composed of adjacent k-qubit unitaries. Such circuits should also be feasible on near-term devices, provided that k is small. More ambitiously, one could also consider circuits with holographic geometries. Such circuit geometries may be optimal for prediction on gapless or scale-invariant states. We leave these cases to future research.

There are still many conceptual questions that have yet to be answered. For example, what is the physical meaning of the locally-scrambled shadow norm? The norm is closely related to the scrambling properties of the unitary ensemble, which begs the question of how it can be related to other information-theoretic and thermodynamic quantities. Our setup assumes a qubit system, where the local degrees of freedom are finite and commuting. How do we apply our approach in Sec. 4.3 to the fermionic context or other experimental contexts which may not be easily described in terms of qubits?

One may also consider different types of locally-scrambled or approximately locally-

scrambled ensembles than the Clifford ensemble. For example, one may consider instead quantum Brownian motion ensembles, where the evolution is generated by a random Hamiltonian coupling nearby sites. The entanglement features of such ensembles have been extensively studied in prior work [5, 129, 229]. What are the reconstruction coefficients for such ensembles? What are their prediction properties? Classical shadow tomography is a rapidly developing field with many interesting applications and open questions. We hope that our paper provides some insights that can motivate further interesting developments.

Chapter 4, in full, is a reprint of the material as it appears in Scalable and Flexible Classical Shadow Tomography with Tensor Networks 2023. Akhtar, Ahmed; You, Hu, Hong-Ye; Yi-Zhuang, Quantum, 2023. The dissertation author was the first author and primary investigator of this paper.

Chapter 5

Classical Shadows and the Measurement-Induced Phase Transition

Classical shadow tomography [83, 103, 168] offers an efficient randomized measurement scheme for extracting physically relevant information from a quantum state. Much research [6, 12, 21, 27, 59, 60, 86, 89, 98–100, 104, 109, 136, 196] primarily concentrates on the randomized measurement protocol that entails random unitary evolution, followed by the final stage of local measurements on all qubits. This process is akin to halting the universe's time evolution to measure every qubit. A more realistic measurement scheme involves conducting local measurements intermittently while the entire quantum system continues to evolve, which more closely imitates how we observe the quantum universe surrounding us. This situation can be represented by hybrid quantum circuits [65, 139, 182, 200] formed by randomly interspersing local measurements among unitary gates in a quantum circuit. Notably, hybrid quantum circuits reveal a phase transition [15, 16, 46, 63, 81, 113, 141, 152, 220, 232] in the quantum entanglement among qubits when the measurement rate surpasses a critical threshold, known as the measurement-induced entanglement transition or the purification transition. Our focus in this work is to explore the hybrid circuit as a randomized measurement scheme for classical shadow tomography and investigate the reconstruction of the quantum state using measurement outcomes obtained from intermittent measurements during the hybrid circuit's evolution, as illustrated in Fig. 5.1.

145
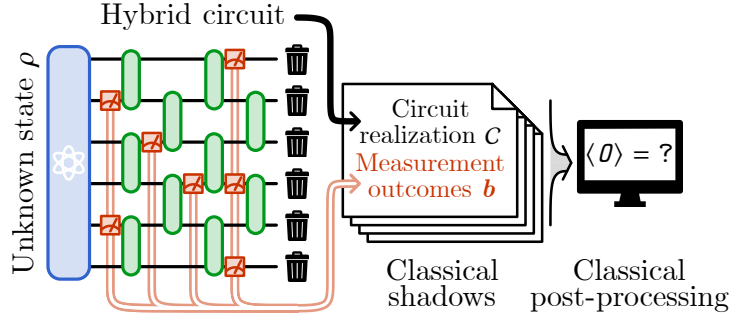
Figure 5.1. Using hybrid quantum circuit as a randomized measurement scheme for classical shadow tomography. Starting from an unknown quantum state ρ, evolve the system by layers of random local Clifford gates, and measure each qubit with probability p in random Pauli basis in each layer. The final state is trashed, but the circuit realization $\mathscr{C}$ (the gate choices and measurement observables) and the measurement outcomes $\boldsymbol{b}$ are recorded as a classical shadow. Repeated randomized measurements of copies of ρ will collect a dataset of classical shadows, which can be used to predict the physical properties of the state ρ through classical post-processing.

The primary scientific question we aim to address concerns the efficiency of extracting information about the initial quantum state from intermittent measurement outcomes collected from the hybrid quantum circuit, within the context of classical shadow tomography. To address this problem, we first expanded the existing classical shadow tomography framework to accommodate more general scenarios where measurements can occur at any spacetime position throughout the quantum evolution. In particular, we introduced a systematic classical post-processing method for reconstructing the quantum state from the classical data of random circuit realizations and measurement outcomes in repeated experiments. Numerical simulations were conducted to validate the proposed reconstruction formula.

Subsequently, we defined the locally-scrambled shadow norm [27, 98] for the hybrid quantum circuit measurement scheme, which quantifies the typical number M of experiments required to estimate the expectation value of an observable accurately, also referred to as the *sample complexity* in quantum state tomography. Utilizing the tensor network method [6, 21], we found that the sample complexity M scales with the operator size $k = |\operatorname{supp} O|$ of the observable as $M \simeq \beta^k \operatorname{poly}(k)$, with the base β depending on the measurement rate p of the hybrid quantum

circuit. We noted that β is minimized (yielding optimal sample complexity scaling) when the measurement rate $p = p_c$ is tuned to the critical point of the measurement-induced transition in the hybrid quantum circuit. The minimal value is found to be around $\beta_{\min} \approx 2.2$. Therefore, measurement-induced criticality is tomographically optimal within the scope of the hybrid quantum circuit measurement scheme.

5.1   Generalized Classical Snapshots

Our first goal is to extend the existing theoretical framework of classical shadow tomography to accommodate more general randomized measurement schemes that permit intermittent and partial measurements throughout random quantum evolutions. Conceptually, the idea is as follows: irrespective of how single-qubit measurements are arranged and implemented in a single-shot experiment, the experimental result must be a string of classical bits, denoted as $\boldsymbol{b} = (b_1, b_2, \cdots)$, which represents the measurement outcome $b_n \in \{0, 1\}$ for the nth measurement in the process. Given an initial quantum state ρ and a particular measurement circuit $\mathscr{C}$ (specified by both the circuit structure and gate choices), the entire measurement protocol can be characterized by the conditional probability $p(\boldsymbol{b}|\rho, \mathscr{C})$.

The linearity of quantum mechanics implies that there must exist a measurement operator $\sigma_{\boldsymbol{b}|\mathscr{C}}$ associated with each possible string of measurement outcomes $\boldsymbol{b}$, such that:

$$p(\boldsymbol{b}|\rho, \mathscr{C}) \propto \mathrm{Tr}(\sigma_{\boldsymbol{b}|\mathscr{C}}\rho). \tag{5.1}$$

We will call the operator $\sigma_{\boldsymbol{b}|\mathscr{C}}$ a *classical snapshot*. In the conventional classical shadow tomography, where the randomized measurement is implemented by first applying a random unitary transformation U to the initial state ρ (as $\rho \to U\rho U^{\dagger}$) and then measuring every qubit separately in the Z-basis, the classical snapshot $\sigma_{\boldsymbol{b}|\mathscr{C}}$ reduces to the standard form of $\sigma_{\boldsymbol{b}|\mathscr{C}} = U^{\dagger}|\boldsymbol{b}\rangle\langle\boldsymbol{b}|U$. Beyond this conventional setup, Eq. (5.1) provides a more general definition of classical snapshots when the measurement protocol is more involved. The classical snapshot

$\sigma_{\boldsymbol{b}|\mathscr{C}}$ should be a Hermitian positive semi-definite operator to ensure the real positivity of the conditional probability $p(\boldsymbol{b}|\rho,\mathscr{C})$. Given this property, it is natural to normalize $\sigma_{\boldsymbol{b}|\mathscr{C}}$ such that $\mathrm{Tr}\,\sigma_{\boldsymbol{b}|\mathscr{C}} = 1$ [1], and view $\sigma_{\boldsymbol{b}|\mathscr{C}}$ as another density matrix, called the classical snapshot state.

5.2 Hybrid Quantum Circuit Measurement

The hybrid quantum circuit measurement scheme is depicted in Fig. 5.1. Starting from an N-qubit unknown quantum state ρ of interest, apply the measurement and unitary layers alternately, where:

- Measurement layer: For each qubit independently, with probability p, choose to measure it in one of the three Pauli bases randomly. In the l-th measurement layer, suppose A_l is the subset of qubits chosen to be measured. For each chosen qubit $i \in A_l$, let $P_i^{(l)} \in \{X_i, Y_i, Z_i\}$ be the choice of Pauli observable and $b_i^{(l)} \in \{0,1\}$ be the corresponding measurement outcome. The measurement layer is described by the Kraus operator

$$K_l^M = \prod_{i \in A_l} \frac{\mathbb{1} + (-)^{b_i^{(l)}} P_i^{(l)}}{2}. \tag{5.2}$$

- Unitary layer: For every other nearest-two-qubit bond independently, apply a Clifford gate [75, 76] uniformly drawn from the two-qubit Clifford group. The Kraus operator for the l-th unitary layer is

$$K_l^U = \begin{cases} \prod_i U_{2i,2i+1}^{(l)} & l \in \text{even}, \\ \prod_i U_{2i-1,2i}^{(l)} & l \in \text{odd}, \end{cases} \tag{5.3}$$

which alternates between even and odd bonds with the layer index l (such that the unitary gates form a brick-wall pattern as shown in Fig. 5.1).

Packing the choice of measurement subsets A_l, Pauli observables $\boldsymbol{P}^{(l)}$ and Clifford gates $\boldsymbol{U}^{(l)}$ (for $l = 1, 2, \cdots$) altogether into the specification of a measurement circuit $\mathscr{C}$, and gathering all the

---

[1]See Appendix D.0.3 for more rigorous treatment of the normalization.

measurement outcomes $\boldsymbol{b} = \{\boldsymbol{b}^{(l)}\}$ together as a classical bit-string, the probability to observe $\boldsymbol{b}$ given $\mathscr{C}$ is

$$p(\boldsymbol{b}|\boldsymbol{\rho},\mathscr{C}) = \mathrm{Tr}(K_{\boldsymbol{b}|\mathscr{C}} \, \boldsymbol{\rho} \, K_{\boldsymbol{b}|\mathscr{C}}^{\dagger}), \qquad (5.4)$$

where $K_{\boldsymbol{b}|\mathscr{C}} = \prod_l K_l^U K_l^M$ is the overall Krause operator.
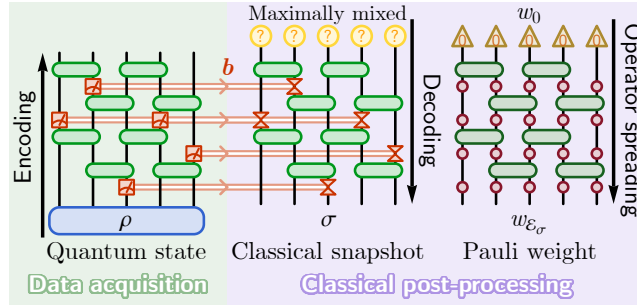
Figure 5.2. Protocol of classical shadow tomography for hybrid quantum circuits. The quantum state ρ is efficiently encoded as classical information by randomized measurements in the data acquisition phase. A classical snapshot state σ is decoded by backward evolution from a maximally mixed state, given the circuit structure and measurement outcomes $\boldsymbol{b}$. On the other hand, its prior Pauli weights $w_{\mathcal{E}_\sigma}(P)$ are inferred following the operator spreading dynamics.

Then following the assertion in Eq. (5.1), the classical snapshot associated with such measurement outcome should be identified as

$$\sigma_{\boldsymbol{b}|\mathscr{C}} = \frac{K_{\boldsymbol{b}|\mathscr{C}}^{\dagger} K_{\boldsymbol{b}|\mathscr{C}}}{\mathrm{Tr}(K_{\boldsymbol{b}|\mathscr{C}}^{\dagger} K_{\boldsymbol{b}|\mathscr{C}})}, \qquad (5.5)$$

where the denominator normalizes the classical snapshot as a state. Since the measurement circuit $\mathscr{C}$ is composed of Clifford gates and Pauli measurements, every classical snapshot $\sigma_{\boldsymbol{b}|\mathscr{C}}$ is a stabilizer state and can be efficiently represented and reconstructed on classical computers [75, 76]. As illustrated in Fig. 5.2, to reconstruct $\sigma_{\boldsymbol{b}|\mathscr{C}}$, one starts with a maximally mixed state (described by the density matrix $\mathbb{1}/(\mathrm{Tr}\,\mathbb{1})$) and traces back the measurement circuit: inverting every unitary gate, replacing every measurement by projection to the measurement outcome, and normalizing the final state in the end.

5.3 Posterior and Prior Distributions

We can interpret the hybrid quantum circuit measurement process as a measure-and-prepare quantum channel that measures the initial state ρ and prepares the classical snapshot state σ with the *posterior* probability:

$$p(\sigma|\rho) := \sum_{\boldsymbol{b},\mathscr{C}} \delta_{\sigma,\sigma_{\boldsymbol{b}|\mathscr{C}}} p(\boldsymbol{b}|\rho,\mathscr{C}) p(\mathscr{C}), \tag{5.6}$$

where $p(\mathscr{C})$ denotes the probability of realizing a specific measurement circuit $\mathscr{C}$. Assuming that all Pauli measurements and Clifford gates are chosen uniformly, $p(\mathscr{C}) \propto \prod_l p^{|A_l|}(1-p)^{N-|A_l|}$ will only be affected by the measurement rate p of the hybrid quantum circuit. By conducting the hybrid quantum circuit measurements of the target state ρ repeatedly, one can sample classical snapshot states σ from the *posterior* distribution $p(\sigma|\rho)$, forming an ensemble $\mathscr{E}_{\sigma|\rho} = \{\sigma|\sigma \sim p(\sigma|\rho)\}$. The objective of classical shadow tomography is to predict properties of ρ based on the samples of $\mathscr{E}_{\sigma|\rho}$ collected from experiments as classical data.

We introduce the *prior* distribution $p(\sigma)$ of the classical snapshot [6], defined as $p(\sigma) := p(\sigma|\rho = \mathbb{1}/(\mathrm{Tr}\,\mathbb{1}))$. This distribution describes our knowledge about classical snapshots before observing the quantum state ρ (as if ρ is maximally mixed). The prior distribution solely characterizes the statistical properties of the randomized measurement scheme, reflecting our uncertainty about the measurement circuit structures and gate choices.

5.4 Pauli weight

A crucial property of the prior classical snapshot ensemble $\mathscr{E}_\sigma = \{\sigma|\sigma \sim p(\sigma)\}$ is its *Pauli weight* [21, 27]

$$w_{\mathscr{E}_\sigma}(P) := \mathop{\mathbb{E}}_{\sigma \sim p(\sigma)} (\mathrm{Tr}\,P\sigma)^2, \tag{5.7}$$

defined for any Pauli operator $P = \prod_i P_i$ (where $P_i \in \{I,X,Y,Z\}$ denotes the Pauli operator on the i-th qubit). The Pauli weight $w_{\mathscr{E}_\sigma}(P)$ fully characterizes the second-moment statistical

feature of the prior distribution $p(\sigma)$. It represents the probability for a Pauli observable P to be transformed to the measurement basis and observed directly by the randomized measurement. It plays an important role in performing and analyzing classical shadow tomography.

For hybrid quantum circuits, the Pauli weight can be computed following the operator dynamics [5, 24, 35, 43, 95, 118, 125, 127, 157, 160, 176, 186, 188, 215, 226, 236, 241]. For every step of the physical evolution of a random quantum state ρ through a random quantum channel $\mathcal{K}$, the Pauli weight will be updated by the Markov process [2]

$$w_{\mathcal{E}_{\mathcal{K}(\rho)}}(P) = \sum_{P'} w_{\mathcal{E}_{\mathcal{K}}}(P,P') w_{\mathcal{E}_{\rho}}(P'), \tag{5.8}$$

where $w_{\mathcal{E}_{\mathcal{K}}}(P,P') := \mathbb{E}_{\mathcal{K} \in \mathcal{E}_{\mathcal{K}}}(\mathrm{Tr}(P\mathcal{K}(P'))/\mathrm{Tr}\,\mathbb{1})^2$ is the Pauli transfer matrix of the random channel ensemble $\mathcal{E}_{\mathcal{K}}$. For every two-qubit random Clifford unitary channel $\mathcal{U}$ and every probabilistic single-qubit random Pauli measurement channel $\mathcal{M}$, the corresponding Pauli transfer matrices are

$$\begin{aligned} w_{\mathcal{E}_{\mathcal{U}}}(P,P') &= \check{\delta}_{P,\mathbb{1}}\check{\delta}_{P',\mathbb{1}} + \frac{1}{15}(1-\check{\delta}_{P,\mathbb{1}})(1-\check{\delta}_{P',\mathbb{1}}), \\ w_{\mathcal{E}_{\mathcal{M}}}(P,P') &= \frac{p}{9}(1+2\check{\delta}_{P,\mathbb{1}})(1+2\check{\delta}_{P',\mathbb{1}}) + (1-p)\check{\delta}_{P,P'}, \end{aligned} \tag{5.9}$$

where $\check{\delta}$ denotes the Kronecker delta symbol restricted to the support of the corresponding quantum channel. Starting from the initial Pauli weight $w_0(P) = \delta_{P,\mathbb{1}}$ of the maximally mixed state and applying the Pauli transfer matrix in accordance with the measurement circuit structure (see Fig. 5.2), the classical snapshot Pauli weight $w_{\mathcal{E}_{\sigma}}(P)$ can be evaluated following Eq. (5.8). In the end, the Pauli weight should be normalized to $w_{\mathcal{E}_{\sigma}}(\mathbb{1}) = 1$ to be consistent with the normalization of the classical snapshot states defined in Eq. (5.5).

---

[2]See Appendix D.0.3 for a brief review of the Markov evolution of Pauli weights.

5.5 Observable Estimation

We now present a key result of our study: given any Pauli observable P, its expectation value on the initial quantum state ρ can be inferred from the posterior classical snapshots via [21, 27]

$$\langle P \rangle := \mathrm{Tr}(P\rho) = \mathop{\mathbb{E}}_{\sigma \sim p(\sigma|\rho)} \frac{\mathrm{Tr}(P\sigma)}{w_{\mathscr{E}_\sigma}(P)}. \tag{5.10}$$

For more general observable $O = \sum_P o_P P$, the expectation value can be similarly predicted by $\langle O \rangle = \mathbb{E}_{\sigma \sim p(\sigma|\rho)} O_\sigma$, where $O_\sigma := \sum_P o_P \mathrm{Tr}(P\sigma)/w_{\mathscr{E}_\sigma}(P)$ is the *single-shot estimation* [103] of the observable O given a particular classical snapshot σ, defined based on Eq. (5.10). This allows us to decode the quantum information about the original state ρ from the classical shadows collected from the hybrid quantum circuit measurement. In practice, the expectation $\mathbb{E}_{\sigma \sim p(\sigma|\rho)}$ is often estimated by the median of means over a finite number of classical snapshots collected from experiments.
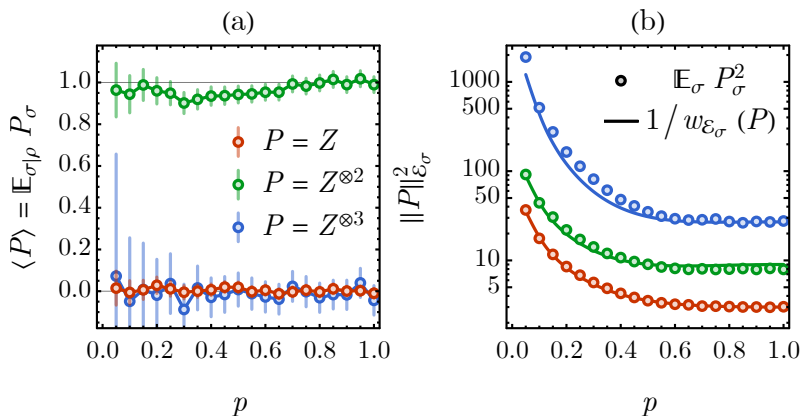
Figure 5.3. Demonstration of hybrid quantum circuit classical shadow tomography on a 12-qubit GHZ state. (a) Predicted observable expectation values $\langle P \rangle$ and (b) locally-scrambled shadow norm $\|P\|_{\mathscr{E}_\sigma}^2$ as functions of the measurement rate p. Colors label different Pauli observables $P = Z^{\otimes k}$.

To demonstrate the validity of Eq. (5.10), we carried out numerical experiments. We take the Greenberger-Horne-Zeilinger (GHZ) [78] state of $N = 12$ qubits, described by $\rho = |\psi\rangle\langle\psi|$

with $|\psi\rangle = (|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$. We consider a randomized measurement scheme implemented by shallow hybrid circuits, which contain three layers of random Clifford gates, together with random Pauli measurements inserted before each unitary layer with probability p on each qubit. We simulate the protocol numerically on a classical computer by repeatedly preparing the GHZ state, applying the hybrid circuit, and collecting the measurement outcomes. For every given measurement rate p, we collect $M = 50000$ samples and estimate the Pauli observables $P = Z^{\otimes k}$ based on the measurement outcomes using Eq. (5.10). Our results, shown in Fig. 5.3(a), demonstrate that the estimated observable expectation values are consistent with their theoretical expectation on the GHZ state throughout the full range of p, i.e., $\langle Z^{\otimes k} \rangle = \frac{1}{2}(1 + (-1)^k)$.

5.6 Sample Complexity Scaling

The statistical uncertainty in the estimation, indicated by the error bar in Fig. 5.3(a), is due to the finite number of samples. The typical variance $\mathrm{var}\langle O \rangle \sim \|O\|_{\mathscr{E}_\sigma}^2/M$ scales inversely with the number M of samples. The coefficient $\|O\|_{\mathscr{E}_\sigma}^2 := \mathbb{E}_{\sigma \sim p(\sigma)} O_\sigma^2$ is the *locally-scrambled shadow norm*, introduced in Ref. [98]. It upper-bounds the variance of the single-shot estimation O_σ over the prior classical snapshot ensemble $\mathscr{E}_\sigma$. For Pauli observable P, the shadow norm has a simple expression [21, 27]

$$\|P\|_{\mathscr{E}_\sigma}^2 = \frac{1}{w_{\mathscr{E}_\sigma}(P)}. \tag{5.11}$$

In Fig. 5.3(b), the second moment of the single-shot estimation $\mathbb{E}_{\sigma \sim p(\sigma)} P_\sigma^2$ is compared with the inverse Pauli weight $1/w_{\mathscr{E}_\sigma}(P)$ calculated from operator spreading dynamics. The results indicate a close match between the two measures. For generic observable $O = \sum_P o_P P$, the shadow norm is given by $\|O\|_{\mathscr{E}_\sigma}^2 = \sum_P |o_P|^2 \|P\|_{\mathscr{E}_\sigma}^2$. The shadow norm quantifies the number M of samples needed to control the estimation variances $\mathrm{var}\langle O \rangle \lesssim \delta^2$ below a desired level set by a small δ, which scales as $M \sim \|O\|_{\mathscr{E}_\sigma}^2/\delta^2$. Therefore, the shadow norm measures the sample complexity for classical shadow tomography to predict the observable O based on the randomized measurement scheme characterized by $\mathscr{E}_\sigma$.

To study how the shadow norm scales with the operator size, we use the matrix product state (MPS) based approach developed in Ref. [6, 21] to compute the Pauli weight $w_{\mathscr{E}_\sigma}(P)$ following the operator spreading dynamics and determine the shadow norm $\|P\|_{\mathscr{E}_\sigma}^2 = 1/w_{\mathscr{E}_\sigma}(P)$ for Pauli observables P of different operator sizes $k = |\mathrm{supp}\,P|$. The result is plotted in Fig. 5.4(a). The shadow norm scales with the operator size k exponentially with a base β at the leading level

$$\|P\|_{\mathscr{E}_\sigma}^2 \simeq \beta^k \mathrm{poly}(k), \tag{5.12}$$

where $\mathrm{poly}(k)$ stands for sub-leading correction that is polynomial in k. This is consistent with the intuition that longer Pauli observable will require exponentially more local measurements to determine. However, the base β depends on the measurement rate p, as shown in Fig. 5.4(b).
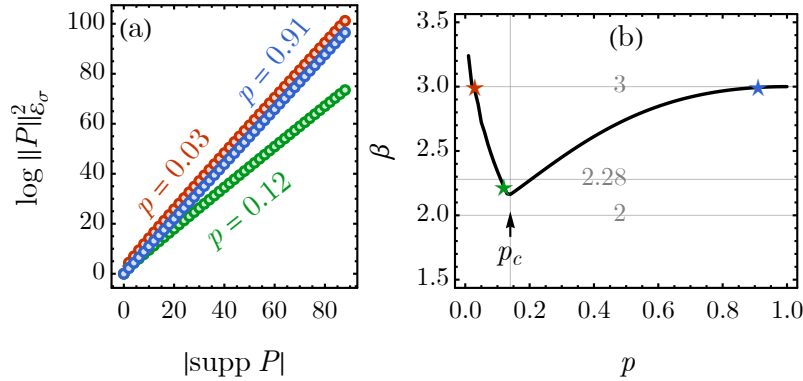
Figure 5.4. (a) Dependence of log shadow norm $\log\|P\|_{\mathscr{E}_\sigma}^2$ on the operator size $|\mathrm{supp}\,P|$ at different measurement rates, demonstrating a leading linear behavior. (b) The base β minimizes at a measurement rate p_c that matches the measurement-induced transition of hybrid circuits. The measurement rates exemplified in (a) are highlighted as stars in (b).

We find that β is minimized at $p = p_c$ when the hybrid quantum circuit operates at the measurement-induced criticality, and the shadow norm scales as

$$\|P\|_{\mathscr{E}_\sigma}^2|_{p=p_c} \simeq \beta_{\min}^k k^{2\Delta}, \tag{5.13}$$

where $\beta_{\min} = 2.23 \pm 0.006$ and $\Delta = 0.33 \pm 0.02$ are determined by fitting. We expect the critical

exponent Δ to be universal, corresponding to the scaling dimension of a defect operator in the boundary conformal field theory (CFT) for the measurement-induced transition [3]. The minimal $\beta_{\min}$ enters the region between $3^{3/4} \approx 2.28$ and 2, which is the range of optimal scaling achievable by shallow circuit classical shadows [109].

The minimization of β can be understood by examining it from both sides of the phase transition. In the area-law phase ($p > p_c$), β should decrease with decreasing measurement rate p. This is because a lower measurement rate allows for a few more local measurements to be deferred to deeper layers of the unitary circuit, enabling larger-size observables to be probed more efficiently by leveraging the scrambling power of shallow circuits. However, in the volume-law phase ($p < p_c$), if the measurement rate continues to decrease, β will instead increase. Because the circuit's scrambling power becomes so strong that it begins to hide the quantum information of the initial state from local measurements deep in the circuit [46, 63, 81], which renders the measurements increasingly inefficient. As the measurement rate approaches zero ($p \to 0$), the shadow norm must diverge, because it becomes impossible to reconstruct the initial state in the absence of measurements. Therefore, the optimal scaling of the shadow norm (or the sample complexity) can only occur at the transition point p_c, where observables of all scales are probed efficiently [4].

5.7 Summary and Discussions

In this work, we present the classical shadow tomography approach for decoding quantum information from measurement outcomes of hybrid quantum circuits. This method involves computing classical snapshots associated with measurement outcomes and using them to infer properties of the initial quantum state. The Pauli weight of the prior classical snapshot ensemble characterizes the statistical properties of the randomized measurement scheme, and the shadow norm quantifies the sample complexity for predicting observables. The log shadow norm

---

[3]See Appendix D.1 for a statistical mechanical model interpretation for the scaling behavior.
[4]See Appendix D.2 for a more detailed quantitative analysis using toy models.

scales linearly with the operator size of the observable and exhibits optimal scaling at a critical measurement rate of the hybrid circuit that corresponds to the measurement-induced criticality.

Hybrid quantum circuits are known for their error correction encoding in the volume-law phase [46, 63, 81]. To use them as a random quantum error correction code, the ability to decode quantum information from measurement outcomes is essential. Classical shadow tomography provides a systematic and universal way to decode hybrid quantum circuits, making them suitable for more exciting quantum information applications.

Measurement-induced transition in hybrid quantum circuits was originally proposed as an entanglement transition. However, measuring entanglement entropy is a difficult task that requires post-selections. With classical shadow tomography, we can directly benchmark the prior classical snapshot Pauli weight $w_{\mathscr{E}_\sigma}(P)$ on a known quantum states ρ (assuming $\mathrm{Tr}(P\rho) \neq 0$),

$$w_{\mathscr{E}_\sigma}(P) = \frac{\mathbb{E}_{\sigma \sim p(\sigma|\rho)} \mathrm{Tr}(P\sigma)}{\mathrm{Tr}(P\rho)}. \tag{5.14}$$

where $p(\sigma|\rho)$ can be sampled by performing the hybrid circuit measurement on ρ. Then β can be extracted by fitting the dependence of $w_{\mathscr{E}_\sigma}(P)$ with respect to $k = |\operatorname{supp} P|$. It is supposed to exhibit a kink at the measurement-induced transition as Fig. 5.4, which provides another method to detect the transition without post-selections apart from the cross-entropy benchmark [142].

Note added. — Up on finishing this work, we become aware that a related work [107] appeared.

$$\rule{2cm}{0.4pt}\!\!\blacktriangleright\!\!\blacklozenge\!\!\blacktriangleleft\!\!\rule{2cm}{0.4pt}$$

Chapter 5, in part, has been submitted for publication as it appears in Measurement-Induced Criticality is Tomographically Optimal 2023. Akhtar, Ahmed; You, Hu, Hong-Ye; Yi-Zhuang, ArXiv, 2023. The dissertation author was the first author and primary investigator of this paper.

Chapter 6

Summary and outlook

In this dissertation, we applied the concepts of entanglement, measurement, and randomness in quantum mechanics to explore the consequences of several recurring themes. In every chapter, we utilized some notion of local-basis invariance at an ensemble level. In Chapter 2, Chapter 3, we used local-scrambling at the state and operator ensembles to prove markovianness in the average purity dynamics. This is a useful property, because it allows us to study the entanglement dynamics directly i.e. without reference to the original quantum state ensembles. The dynamics in the entanglement features are also often simpler than the dynamics of the quantum states, mapping onto Ising spin systems. The transfer matrix evolving the entanglement features state carries a $\mathbb{Z}_2$ symmetry due to the systems being closed, and can be used to define an entanglement feature Hamiltonian, whose spectral properties define the entanglement velocity, rate of thermalization, fixed points, etc. This is only possible by erasing local information, by which we are able to isolate information that characterizes correlations only. This is the same strategy employed by Wigner in the derivation of spectral properties of heavy nuclei: simplify by averaging over irrelevant details.

The entanglement features state itself is also much simpler, being strictly positive and having generically low-entanglement. This simplicity allows efficient MPS descriptions of the entanglement dynamics, which enable us to numerically and analytically study the whole entanglement "spectrum" and how information propogates through the system. The spectrum has

several interesting features, including a lower, area-law edge from single-region entanglement, an upper, volume-law edge for multi-region entanglement from a lattice of domain wall pairs, and a multi-region entanglement continuum separated by a finite gap to the single-region entanglement. Via the MPS ansatz, each of these features can be neatly described analytically. Furthermore, we show that this entanglement gap provides a correction to the entanglement velocity which was missed from previous models of entanglement dynamics which did not consider contibutions from multi-region entanglements.

As a statistical symmetry of ensembles, local basis invariance also greatly constrains the form of the measurement channel in classical shadow tomography. Since at the two design level, local basis invariance is equivalent to Pauli invariance, such ensembles lead to measurement channels which are diagonal in the Pauli basis. The eigenvalues of these ensembles is also determined in terms of the entanglement features, and therefore they have efficient MPS reprsentations. In 4, we find also that the eigenvalues of the reconstruction map also have efficent descriptions as MPS for shallow circuits, which allows us to devise powerful schemes to predict a wide range of operators. We find that shallow circuit tomography of depth $\log l$ is optimal for predicting operators of size $\mathscr{O}(l)$, and that the shallow circuits can serve as a more feasible alternative to Clifford measurements for fidelity estimation.

Using the structure of the measurement channel outline in Chapter 4, we extend our study to hybrid circuits, where we study the competition between randomized measurements and entangling unitaries on the locally-scrambled shadow norm. The locally-scrambled shadow norm measures the sample complexity of predicting Pauli operators in locally-scrambled ensembles, and potentially serves as a more operational order parameter in the study of measurement-induced phase transitions. To simplify our analysis in the thermodynamic limit, we focus on the scaling exponent in the shadow norm, which is the coefficient to the volume law term in the log shadow norm. In the area law phase ($p > p_c$), we find that lowering the measurement rate p increases the size of the effective scrambling region, allowing for more efficient prediction. Similarly, in the volume law phase $p < p_c$, we find that increasing the measurement rate decreases the error

158

correcting capacity of measurement channel, thereby extracting information for efficiently for prediction. Therefore, the critical point optimizes the shadow norm.

Another recurring theme is the average purity or entanglement features $e^{-S^{(2)}}$. In Chapter 2, Chapter 3, the entanglement features serve as a measure for the average entanglement; in Chapter 4, Chapter 5, the entanglement features is related to the Pauli weight, the eigenvalues of the measurement channel, by a local basis transformation. The usefulness of the entanglement features lies in their simplicity; because they only depend on the second moment of the random unitary ensemble, the entanglement feature Hilbert space is an Ising space. For higher Renyi moments $n > 2$, we'd have to deal with spaces with much larger local dimension $n!$ and more complicated dynamics. Another simplification of the using the purity is that because it only depends the second moment, it coincides for all t-designs with $t \geq 2$. We use this to our advantage in Chapter 4, Chapter 5, where the shadows are formed by Clifford gates which are not strictly locally-scrambled but are equivalent at the two-design level. Yet another theme is applying techniques from many-body quantum physics to unphysical states: we study the entanglement structure of entanglement feature states; we study the spectral properties of a non-hermitian entanglement feature Hamiltonian; we study a phase transition the pauli weight.

Looking forward, there are several interesting research questions which the author wishes to explore in future works. Firstly, we wish to have a better understanding of the efficiency solving the consistency equation in shallow circuit tomography. In our works, we mostly looked at the shallow circuit limit, we could demonstrate the low bond dimension MPS works to minimize the loss function. What about as we increase the circuit depth? What can we say about how the required bond dimension scales? How does this bond dimension relate to the optimal bond dimension for the entanglement feature? In general, can we devise a technique by which to invert a matrix product state i.e. given a strictly positive MPS state $|\Psi\rangle$ with on-site tensors A^i, can we determine the MPS with on-site tensors B^i such that its corresponding state has reciprocal components to that of $|\Psi\rangle$? In general, this seems like a hard problem, but it would help us understand the efficiency of our approach much better.

Secondly, we wish to understand better understanding of the Pauli weight near the measurement-induced phase transition. In Chapter 5, we make some initial progress at bounding the cusp in Fig. 5.4, but have yet to develop a systematic understanding of the whole phase. We have made some initial progress here which we hope to flesh out in later work. Furthermore, we wish to understand better the effect of different approximations made in deriving this result. For example, we assume that the numerator and denominator can be averaged separately in the definition of the snapshot state in Appendix D.0.5. This approximation is not valid at the critical point, so it would be good to assess what differences arise as a result of this assumption.

Lastly, classical shadow tomography has created an exciting landscape to revisit concepts such as operator complexity, operator spreading, quantum chaos, etc., in a more operational setting. How do quantities such as the butterfly velocity or the OTOC relate to quantities such as the shadow norm? Can we extract these quantities from the Pauli weight transfer matrix itself, similar to the approach outlined in Chapter 2? What effect do symmetries have? Can we develop a new approach to classical shadow tomography in the presence of symmetries? There's a world of new protocols, and consequently, new physical insights, waiting to be discovered.

Appendix A

Theoretical Framework of Entanglement Features

A.1 Entanglement Feature of Page State

The Page state can be considered as a single random tensor. According to Ref. [93], the 2nd Rényi entanglement feature of a random tensor network can be calculated as the partition function of an Ising model,

$$W_{\text{RTN}}[\boldsymbol{\sigma}] = \frac{1}{Z} \sum_{\boldsymbol{\tau}} e^{-E_{\text{RTN}}[\boldsymbol{\sigma}, \boldsymbol{\tau}]}, \tag{A.1}$$

where each random tensor is mapped to an Ising spin τ_i coupled together via the network, and the boundary condition pinned by external Zeeman field along the direction specified by $\boldsymbol{\sigma}$. Applying this result to the Page state,

$$W_{\text{Page}}[\boldsymbol{\sigma}] = \frac{1}{Z} \sum_{\tau} e^{\eta \sum_{i=1}^{L} \sigma_i \tau}, \tag{A.2}$$

where there is only one Ising spin τ because the Page state is only a single random tensor. The τ spin couples to the boundary condition $\boldsymbol{\sigma}$ via uniform field strength $\eta = \frac{1}{2} \log d$, which is determined by the qudit Hilbert space dimension d (see Ref. [93] for derivation). Complete the

summation over Ising spin τ, we obtain

$$W_{\text{Page}}[\boldsymbol{\sigma}] = \frac{2}{Z}\cosh\left(\eta\sum_{i=1}^{L}\sigma_i\right). \tag{A.3}$$

The normalization constant Z is determined by the condition that $W_{\text{Page}}[\Uparrow] \equiv 1$, such that $Z = 2\cosh(\eta L)$, hence

$$W_{\text{Page}}[\boldsymbol{\sigma}] = \frac{\cosh(\eta\sum_{i=1}^{L}\sigma_i)}{\cosh(\eta L)}, \tag{A.4}$$

which can be rewritten as the EF state $|W_{\text{Page}}\rangle$ in Eq. (2.11).

A.2 Entanglement Feature of Two-Qudit Haar Random Unitary

Here we derive the ensemble averaged EF operator for two-qudit Haar random unitary gate. We start with the definition

$$W_{U_{ij}}[\boldsymbol{\sigma},\boldsymbol{\tau}] = \text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}}U_{ij}^{\otimes 2}\mathscr{X}_{\boldsymbol{\tau}}U_{ij}^{\dagger\otimes 2}). \tag{A.5}$$

U_{ij} is a two-qudit gate acting on qudits labeled by i and j. Focusing on these two qudits, the Ising variables $\boldsymbol{\sigma} = (\sigma_i,\sigma_j)$ and $\boldsymbol{\tau} = (\tau_i,\tau_j)$ both contain only two components. Consider averaging the EF $W_{U_{ij}}$ over unitary gates U_{ij} in the Haar random unitary ensemble,

$$\begin{aligned}
\mathbb{E}_{U_{ij}\in\text{Haar}}W_{U_{ij}}[\boldsymbol{\sigma},\boldsymbol{\tau}] &= \mathbb{E}_{U_{ij}\in\text{Haar}}\text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}}U_{ij}^{\otimes 2}\mathscr{X}_{\boldsymbol{\tau}}U_{ij}^{\dagger\otimes 2}) \\
&= \sum_{g,h\in S_2}\text{Wg}(g^{-1}h,d^2)\,\text{Tr}(\mathscr{X}_g\mathscr{X}_{\sigma_1})\,\text{Tr}(\mathscr{X}_g\mathscr{X}_{\sigma_2})\,\text{Tr}(\mathscr{X}_h\mathscr{X}_{\tau_1})\,\text{Tr}(\mathscr{X}_h\mathscr{X}_{\tau_2}),
\end{aligned} \tag{A.6}$$

where Wg is the Weingarten function[52, 219] and g, h are S_2 group elements

$$\text{Wg}(g^{-1}h, d^2) = \begin{cases} \frac{1}{d^4 - 1} & g^{-1}h = || \\ -\frac{1}{d^2(d^4 - 1)} & g^{-1}h = \times \end{cases}.$$ (A.7)

The cycle counting function $\text{Tr}(\mathscr{X}_g \mathscr{X}_h)$ follows

$$\text{Tr}(\mathscr{X}_g \mathscr{X}_h) = \begin{cases} d^2 & gh = || \\ d & gh = \times \end{cases}.$$ (A.8)

Substitute Eq. (A.7) and Eq. (A.8) into Eq. (A.6), we can evaluate $\mathbb{E}_{U_{ij} \in \text{Haar}} W_{U_{ij}}[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ for all configurations of $\boldsymbol{\sigma}, \boldsymbol{\tau}$. In terms of Ising variables (following the identification $|| \Leftrightarrow \uparrow$ and $\times \Leftrightarrow \downarrow$), we can summarize the result as the following matrix in the Ising basis $\boldsymbol{\sigma}, \boldsymbol{\tau} = \uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$

$$\mathbb{E}_{U_{ij} \in \text{Haar}} \hat{W}_{U_{ij}} = \begin{bmatrix} d^4 & d^3 & d^3 & d^2 \\ d^3 & \frac{2d^4}{d^2+1} & \frac{2d^4}{d^2+1} & d^3 \\ d^3 & \frac{2d^4}{d^2+1} & \frac{2d^4}{d^2+1} & d^3 \\ d^2 & d^3 & d^3 & d^4 \end{bmatrix},$$ (A.9)

which is also the matrix representation of the (ensemble averaged) EF operator $\hat{W}_{U_{ij}}$. The matrix can as well be written in terms of Pauli operators as

$$\mathbb{E}_{U_{ij} \in \text{Haar}} \hat{W}_{U_{ij}} = d^2(d + X_i)(d + X_j) - \frac{d^2(d^2 - 1)}{2(d^2 + 1)}(1 - Z_i Z_j)(d^2 - X_i X_j),$$ (A.10)

as claimed in Eq. (2.13). For simplicity, we have omitted $\mathbb{E}_{U_{ij} \in \text{Haar}}$ notation in Eq. (2.13), with the understanding that the EF for an ensemble of unitaries is implicitly averaged.

A.3 Relation Between State and Unitary Entanglement Features

Here we prove Eq. (2.19). Consider a many-body state (multi-qudit) state $|\Psi\rangle$ and an unitary operator U_t supported in the same Hilbert space. Suppose that $|\Psi'\rangle = U_t|\Psi\rangle$, our goal is to derive the time evolution of the corresponding EF state. In general, this is not tractable since the unitary operator U_t contains many non-universal features that are specific to the choice of local basis. Such features may affect the entanglement of the final state but such features are not captured in EF formalism. By this property, we instead consider an ensemble of unitary operator U,

$$\mathscr{E}_U = \left\{ V^\dagger U V \middle| V = \bigotimes_{i=1}^{L} V_i, V_i \in \text{Haar} \right\}, \tag{A.11}$$

where each V_i independently follows Haar random unitary distribution on the ith qudit. Our goal is to compute $\mathbb{E}_{U' \in \mathscr{E}_U} W_{U'|\Psi\rangle}[\boldsymbol{\sigma}]$,

$$\begin{aligned}
\mathbb{E}_{U' \in \mathscr{E}_U} W_{U'|\Psi\rangle}[\boldsymbol{\sigma}] &= \mathbb{E}_{U' \in \mathscr{E}_U} \text{Tr}[\mathscr{X}_{\boldsymbol{\sigma}} (U'|\Psi\rangle\langle\Psi|U'^\dagger)^{\otimes 2}] \\
&= \mathbb{E}_{V \in \text{Haar}} \text{Tr}[\mathscr{X}_{\boldsymbol{\sigma}} (V^\dagger U V |\Psi\rangle\langle\Psi| V^\dagger U^\dagger V)^{\otimes 2}] \\
&= \mathbb{E}_{V \in \text{Haar}} \langle\Psi|^{\otimes 2} V^{\dagger \otimes 2} U^{\dagger \otimes 2} V^{\otimes 2} \mathscr{X}_{\boldsymbol{\sigma}} V^{\dagger \otimes 2} U^{\otimes 2} V^{\otimes 2} |\Psi\rangle^{\otimes 2} \\
&= \mathbb{E}_{V \in \text{Haar}} \langle\Psi|^{\otimes 2} V^{\dagger \otimes 2} U^{\dagger \otimes 2} \mathscr{X}_{\boldsymbol{\sigma}} U^{\otimes 2} V^{\otimes 2} |\Psi\rangle^{\otimes 2} \\
&= \sum_{\boldsymbol{\tau}, \boldsymbol{\tau}'} \text{Tr}(\mathscr{X}_{\boldsymbol{\tau}} |\Psi\rangle\langle\Psi|^{\otimes 2}) \text{Tr}(\mathscr{X}_{\boldsymbol{\tau}'} U^\dagger(t)^{\otimes 2} \mathscr{X}_{\boldsymbol{\sigma}} U_t^{\otimes 2}) \prod_i \text{Wg}(\tau_i'^{-1} \tau_i, d) \\
&= \sum_{\boldsymbol{\tau}, \boldsymbol{\tau}'} W_{|\Psi\rangle}[\boldsymbol{\tau}] W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}'] \prod_i \text{Wg}(\tau_i'^{-1} \tau_i, d),
\end{aligned} \tag{A.12}$$

where Wg denotes the Weingarten function[52, 219] originated from the Haar ensemble average of $V^{\dagger \otimes 2} V^{\otimes 2}$, and $\boldsymbol{\tau}, \boldsymbol{\tau}'$ are new set of Ising variables. The derivation in Eq. (A.12) can also be diagrammatically represented as Fig. A.1.

By definition, the Weingarten function, when viewed as a matrix indexed by $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$, is
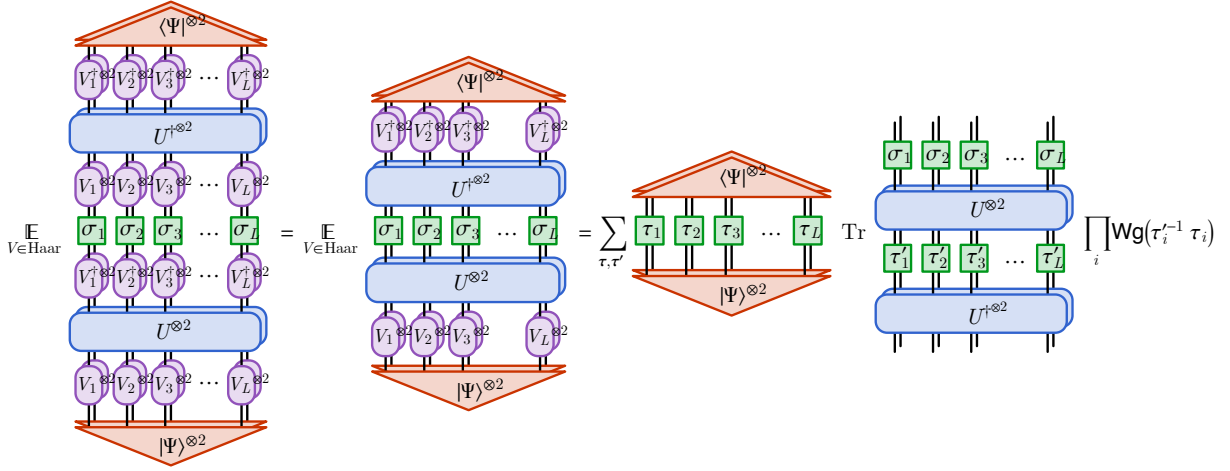
Figure A.1. Diagrammatic proof of Eq. (2.19)

the inverse of the Gram matrix $\text{Tr} \, \mathscr{X}_{\boldsymbol{\tau}} \mathscr{X}_{\boldsymbol{\tau}'} = \langle \boldsymbol{\tau}' | \hat{W}_{\mathbb{1}} | \boldsymbol{\tau} \rangle$, which is simply the matrix representation of the EF operator $\hat{W}_{\mathbb{1}}$ of the identity operator. So the Weingarten function is given by the matrix element of $\hat{W}_{\mathbb{1}}^{-1}$ as

$$\prod_i \text{Wg}(\tau_i'^{-1} \tau_i, d) = \langle \boldsymbol{\tau}' | \hat{W}_{\mathbb{1}}^{-1} | \boldsymbol{\tau} \rangle. \tag{A.13}$$

Therefore, in operator form, we have

$$\begin{aligned}
\mathbb{E}_{U' \in \mathscr{E}_U} | W_{U'\Psi} \rangle &= \sum_{\boldsymbol{\sigma}} W_{U'|\Psi\rangle}[\boldsymbol{\sigma}] |\boldsymbol{\sigma}\rangle \\
&= \sum_{\boldsymbol{\sigma}, \boldsymbol{\tau}', \boldsymbol{\tau}} \left(|\boldsymbol{\sigma}\rangle W_U[\boldsymbol{\sigma}, \boldsymbol{\tau}'] \langle \boldsymbol{\tau}'| \right) \hat{W}_{\mathbb{1}}^{-1} \left(W_{|\Psi\rangle}[\boldsymbol{\tau}] |\boldsymbol{\tau}\rangle \right) \\
&= \hat{W}_U \hat{W}_{\mathbb{1}}^{-1} | W_\Psi \rangle,
\end{aligned} \tag{A.14}$$

as stated in Eq. (2.19).

A.4   Spectral Properties of Entanglement Hamiltonian

Let us start with the most general form of the EF Hamiltonian $\hat{H}_{\text{EF}}$ given in Eq. (2.32) and investigate its spectral properties.

$$\hat{H}_{\text{EF}} = \sum_{i,j} \hat{H}_{ij}, \qquad \hat{H}_{ij} = g_{ij} \frac{1 - Z_i Z_j}{2} e^{-\beta_{ij} X_i X_j - \delta(X_i + X_j)}, \tag{A.15}$$

with $\coth \delta = d$. Our first goal is to show that $\hat{H}_{\text{EF}}$ is positive semi-definite. The trick is to first deform $\hat{H}_{\text{EF}}$ to a Hermitian version $\hat{H}'_{\text{EF}}$, given by

$$\hat{H}'_{\text{EF}} = \hat{W}_{\mathbb{1}}^{-1/2} \hat{H}_{\text{EF}} \hat{W}_{\mathbb{1}}^{1/2}. \tag{A.16}$$

Because $\hat{H}_{\text{EF}}$ and $\hat{H}'_{\text{EF}}$ are related by similar transformation, they share the same set of eigenvalues. So the positivity of the original EF Hamiltonian $\hat{H}_{\text{EF}}$ is equivalent to the positivity of the transformed Hermitian version $\hat{H}'_{\text{EF}}$. The later turns out to be easier to prove. By the way, to see that $\hat{H}'_{\text{EF}}$ is Hermitian (or real symmetric to be more precise), we use $\hat{W}_{\mathbb{1}}^{\mathsf{T}} = \hat{W}_{\mathbb{1}}$ and Eq. (2.35) that $\hat{W}_{\mathbb{1}} \hat{H}_{\text{EF}}^{\mathsf{T}} = \hat{H}_{\text{EF}} \hat{W}_{\mathbb{1}}$, then

$$\hat{H}'^{\mathsf{T}}_{\text{EF}} = \hat{W}_{\mathbb{1}}^{1/2} \hat{H}_{\text{EF}}^{\mathsf{T}} \hat{W}_{\mathbb{1}}^{-1/2} = \hat{W}_{\mathbb{1}}^{-1/2} (\hat{W}_{\mathbb{1}} \hat{H}_{\text{EF}}^{\mathsf{T}}) \hat{W}_{\mathbb{1}}^{-1/2}$$
$$= \hat{W}_{\mathbb{1}}^{-1/2} (\hat{H}_{\text{EF}} \hat{W}_{\mathbb{1}}) \hat{W}_{\mathbb{1}}^{-1/2} = \hat{W}_{\mathbb{1}}^{-1/2} \hat{H}_{\text{EF}} \hat{W}_{\mathbb{1}}^{1/2} = \hat{H}'_{\text{EF}}, \tag{A.17}$$

meaning that $\hat{H}'_{\text{EF}}$ is transpose symmetric. Moreover $\hat{H}'_{\text{EF}}$ is real by definition, so $\hat{H}'_{\text{EF}}$ is real and symmetric and therefore Hermitian. As a real symmetric operator, $\hat{H}'_{\text{EF}}$ admits the following spectral decomposition

$$\hat{H}'_{\text{EF}} = \sum_a |V_a\rangle \lambda_a \langle V_a|, \tag{A.18}$$

with $|V_a\rangle = \langle V_a|^{\mathsf{T}}$ being the eigenvector corresponding to the eigenvalue λ_a. If we can show that the expectation value $\langle V|\hat{H}'_{\text{EF}}|V\rangle \geq 0$ is non-negative on *any* state $|V\rangle$ in the EF Hilbert space

(including the eigenstates $|V_a\rangle$), we will be able to prove that all eigenvalues $\lambda_a = \langle V_a|\hat{H}'_{\text{EF}}|V_a\rangle \geq 0$ are non-negative, hence $\hat{H}'_{\text{EF}}$ will be positive semi-definite.

We can show $\langle V|\hat{H}'_{\text{EF}}|V\rangle \geq 0$ by finding the Cholesky decomposition for each terms in $\hat{H}'_{\text{EF}}$. A useful trick is to note that $d(d+X_i) = e^{\delta X_i}/(\tanh \delta \sinh \delta)$ given $d = \coth \delta$, so $\hat{W}_{\mathbb{1}}$ can be rewritten as

$$\hat{W}_{\mathbb{1}} = \prod_{i=1}^{L} d(d+X_i) = \prod_{i=1}^{L} \frac{e^{\delta X_i}}{\tanh \delta \sinh \delta} = \frac{1}{(\tanh \delta \sinh \delta)^L} \prod_{i=1}^{L} e^{\delta X_i}, \tag{A.19}$$

such that any $\hat{W}_{\mathbb{1}}^{\alpha}$ can be simply calculated,

$$\hat{W}_{\mathbb{1}}^{\alpha} = (\tanh \delta \sinh \delta)^{-\alpha L} \prod_{i=1}^{L} e^{\alpha \delta X_i}. \tag{A.20}$$

With this, and substitute Eq. (A.15) in Eq. (A.16), we can show that

$$\begin{aligned}
\hat{H}'_{\text{EF}} &= \sum_{i,j} \hat{H}'_{ij}, \\
\hat{H}'_{ij} &= \hat{W}_{\mathbb{1}}^{-1/2} \hat{H}_{ij} \hat{W}_{\mathbb{1}}^{1/2} \\
&= \prod_{i=1}^{L} e^{-\frac{\delta}{2} X_i} \hat{H}_{ij} \prod_{i=1}^{L} e^{\frac{\delta}{2} X_i} \\
&= e^{-\frac{\delta}{2}(X_i+X_j)} \hat{H}_{ij} e^{\frac{\delta}{2}(X_i+X_j)} \\
&= g_{ij} e^{-\frac{\delta}{2}(X_i+X_j)} \frac{1-Z_iZ_j}{2} e^{-\beta_{ij} X_iX_j} e^{-\frac{\delta}{2}(X_i+X_j)} \\
&= g_{ij} e^{-\frac{\delta}{2}(X_i+X_j)} \frac{1-Z_iZ_j}{2} e^{-\beta_{ij} X_iX_j} \frac{1-Z_iZ_j}{2} e^{-\frac{\delta}{2}(X_i+X_j)}.
\end{aligned} \tag{A.21}$$

In the last step, we use the fact that $\frac{1-Z_iZ_j}{2}$ is a projection operator, so $(\frac{1-Z_iZ_j}{2})^2 = \frac{1-Z_iZ_j}{2}$. Also $\frac{1-Z_iZ_j}{2}$ and $e^{-\beta_{ij} X_iX_j}$ commute with each other, so we are free to move $e^{-\beta_{ij} X_iX_j}$ through $\frac{1-Z_iZ_j}{2}$. The final form of $\hat{H}'_{ij}$ admits the following Cholesky decomposition explicitly

$$\hat{H}'_{ij} = \hat{A}_{ij}^{\mathsf{T}} \hat{A}_{ij}, \qquad \hat{A}_{ij} = g_{ij}^{1/2} e^{-\frac{\beta_{ij}}{2} X_iX_j} \frac{1-Z_iZ_j}{2} e^{-\frac{\delta}{2}(X_i+X_j)}. \tag{A.22}$$

167

For any state $|V\rangle$ in EF Hilbert space, the expectation value $\langle V|\hat{H}'_{ij}|V\rangle = \langle V|\hat{A}^\mathsf{T}_{ij}\hat{A}_{ij}|V\rangle \geq 0$ is non-negative, therefore $\hat{H}'_{ij}$ is positive semi-definite. In consequence, the transformed EF Hamiltonian $\hat{H}'_{\mathrm{EF}} = \sum_{i,j}\hat{H}'_{ij}$ is also positive semi-definite, as it is the sum of positive semi-definite terms $\hat{H}'_{ij}$. Recall that the similar transformation does not affect the eigenvalues, so $\hat{H}_{\mathrm{EF}} = \hat{W}_{\mathbb{1}}^{1/2}\hat{H}'_{\mathrm{EF}}\hat{W}_{\mathbb{1}}^{-1/2}$ is also positive semi-definite.

We can further show that $\hat{H}_{\mathrm{EF}}$ always has two zero modes: one is even under $\mathbb{Z}_2$ Ising symmetry, and the other is odd. Using the left-null-state property $\langle\Uparrow|\hat{H}_{\mathrm{EF}} = 0$ given in Eq. (2.34), it is ensured that $\langle\Uparrow|$ is an left-eigenstate of $\hat{H}_{\mathrm{EF}}$ with zero eigenvalue. Since $\hat{H}_{\mathrm{EF}}$ is $\mathbb{Z}_2$ symmetric, the $\mathbb{Z}_2$ related state $\langle\Downarrow| = \langle\Uparrow|\prod_i X_i$ is also a left zero mode. So by explicit construction, we have shown that $\hat{H}_{\mathrm{EF}}$ has at least two zero eigenvalues. The left zero mode subspace is spanned by $\langle\Uparrow|$ and $\langle\Downarrow|$. Using the relation between left- and right-eigenstate $|R_a\rangle \propto (\langle L_a|\hat{W}_{\mathbb{1}})^\mathsf{T}$, the corresponding right zero mode subspace is spanned by $\hat{W}_{\mathbb{1}}|\Uparrow\rangle$ and $\hat{W}_{\mathbb{1}}|\Downarrow\rangle$.

Since we are most interested about the EF of pure states, we should focus on the $\mathbb{Z}_2$ even state in the zero mode subspace. In that regard, the left and right zero modes in the $\mathbb{Z}_2$ even sector are given by

$$
\begin{aligned}
\langle L_0| &= \frac{\langle\Uparrow| + \langle\Downarrow|}{2}, \\
|R_0\rangle &\propto \hat{W}_{\mathbb{1}}\frac{|\Uparrow\rangle + |\Downarrow\rangle}{2} \\
&= \frac{1}{2}\left(\prod_{i=1}^{L}d(d+X_i)\right)(|\Uparrow\rangle + |\Downarrow\rangle) \\
&= \frac{1}{2}d^{3L/2}\left(\prod_{i=1}^{L}(e^\eta + e^{-\eta}X_i)\right)(|\Uparrow\rangle + |\Downarrow\rangle) \qquad (\eta \equiv \tfrac{1}{2}\log d) \\
&= \frac{1}{2}d^{3L/2}\sum_{\boldsymbol{\sigma}}\left(\prod_{i=1}^{L}e^{\eta\sigma_i}|\boldsymbol{\sigma}\rangle + \prod_{i=1}^{L}e^{-\eta\sigma_i}|\boldsymbol{\sigma}\rangle\right) \\
&= d^{3L/2}\sum_{\boldsymbol{\sigma}}\cosh\left(\eta\sum_{i=1}^{L}\sigma_i\right)|\boldsymbol{\sigma}\rangle.
\end{aligned}
\tag{A.23}
$$

168

The normalization of $|R_0\rangle$ is determined by the condition $\langle L_0|R_0\rangle = 1$, such that

$$|R_0\rangle = \sum_{\boldsymbol{\sigma}} \frac{\cosh\left(\eta \sum_{i=1}^{L} \sigma_i\right)}{\cosh(\eta L)} |\boldsymbol{\sigma}\rangle = |W_{\text{Page}}\rangle. \tag{A.24}$$

In summary, we have shown that in the $\mathbb{Z}_2$ even sector, the EF Hamiltonian $\hat{H}_{\text{EF}}$ has at least one zero eigenvalue $\lambda_0 = 0$, whose left- and right-eigenstates are given by

$$\langle L_0| = \frac{\langle \Uparrow | + \langle \Downarrow |}{2}, \qquad |R_0\rangle = |W_{\text{Page}}\rangle, \tag{A.25}$$

as claimed in Eq. (2.41).

A.5 Derivation of the Dispersion Relation for two-domain-wall Ansatz

Here, we show the derivation of Eq. (2.52). Our goal here is to obtain the analytical expression of excited state energy, namely dispersion relation, $\omega(k)$. Note that the left and right eigenstates are *not* simply each other's conjugate transpose due to the non-hermitian nature of EF Hamiltonian (Eq. (2.32)). For simplicity, we focus on the excitation of left eigenstates and construct the corresponding right eigenstate with $|R\rangle = (\langle L|\hat{W}_1)^\mathsf{T}$. From the discussion in Sec. 2.2.6, the universal left ground state for any parameters g_{ij}, β is the linear combination of all spin-up and spin-down state,

$$\langle L_0| = \frac{\langle \Uparrow | + \langle \Downarrow |}{2}. \tag{A.26}$$

Based on our ED result in Fig. 2.5, the low energy left excited state mainly consists of two-domain-wall states. The generic form of two-domain-wall state can be expressed as

$$\langle k| \equiv C_k \sum_{n,a} \langle k_n, a| \equiv C_k \sum_{n,a} e^{ikn} \phi^*(a) \langle \Uparrow | \prod_{i=n}^{n+a} X_i \tag{A.27}$$

where C_k is the normalization constant and $\langle k_n, a |$ represents the two-domain-wall state ranging from n to $n+a$.

First, we start from deriving the normalization constant.

$$\langle k|k\rangle = |C_k|^2 \sum_{n,m,a,b} e^{ik(n-m)} \phi^*(a)\phi(b) \langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} \prod_{i=m}^{m+b} X_i | \Uparrow \rangle = 1. \qquad \text{(A.28)}$$

Next step is to evaluate $\langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle$. The physical meaning is the transition amplitude between two Bethe string states with evolution as $\hat{W}_{\mathbb{1}}$. There are two possibilities for each site. When both Bethe strings have/do not have excitation at site i, the answer would be $\langle \Uparrow |(\hat{W}_{\mathbb{1}})_i| \Uparrow \rangle = d^2$. When either Bethe string has excitation on site i, the result becomes $\langle \Uparrow |X_i(\hat{W}_{\mathbb{1}})_i| \Uparrow \rangle = d$. To evaluate this quantity, we perform perturbative expansion as $1/d$ series. To obtain analytical expression of $|C_k|^2$, we also approximate $\phi(a)$ as plane wave $\sim e^{-ika/2}$. Physical intuition is that we assume these domain walls have little interaction with each other and thus they can move through each other almost freely. Consequently, plane wave solution is assumed and $a/2$ represents the center location of domain wall. Let's start evaluating the normalization constant up to the order of $1/d^2$,

$$\langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle =$$

$$\delta_{n,m}\delta_{a,b}d^{2N} + (\delta_{n,m}\delta_{a,b+1} + \delta_{n,m}\delta_{a,b-1} + \delta_{n,m+1}\delta_{a,b-1} + \delta_{n,m-1}\delta_{a,b+1})d^{2N-1}$$

$$+ (\delta_{n,m}\delta_{a,b+2} + \delta_{n,m}\delta_{a,b-2} + \delta_{n,m-2}\delta_{a,b+2} + \delta_{n,m+2}\delta_{a,b-2}$$

$$+ \delta_{n,m+1}\delta_{a,b} + \delta_{n,m+1}\delta_{a,b-2} + \delta_{n,m-1}\delta_{a,b} + \delta_{n,m-1}\delta_{a,b+2})d^{2N-2} + \mathcal{O}(d^{2N-3}) \quad \text{(A.29)}$$

For $b = 0, 1$, we would have different terms,

$$\langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} X_m | \Uparrow \rangle = \delta_{n,m} \delta_{a,0} d^{2N} + (\delta_{n,m} \delta_{a,1} + \delta_{n,m-1} \delta_{a,1}) d^{2N-1}$$

$$+ (\delta_{n,m} \delta_{a,2} + \delta_{n,m-2} \delta_{a,2} + \delta_{n,m+1} \delta_{a,0} + \delta_{n,m-1} \delta_{a,0} + \delta_{n,m-1} \delta_{a,2}) d^{2N-2} + \mathcal{O}(d^{2N-3}) \quad \text{(A.30)}$$

$$\langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} X_m X_{m+1} | \Uparrow \rangle =$$

$$\delta_{n,m} \delta_{a,1} d^{2N} + (\delta_{n,m} \delta_{a,2} + \delta_{n,m} \delta_{a,0} + \delta_{n,m+1} \delta_{a,0} + \delta_{n,m-1} \delta_{a,2}) d^{2N-1}$$

$$+ (\delta_{n,m} \delta_{a,3} + \delta_{n,m-2} \delta_{a,3} + \delta_{n,m+1} \delta_{a,1} + \delta_{n,m-1} \delta_{a,1} + \delta_{n,m-1} \delta_{a,3}) d^{2N-2} + \mathcal{O}(d^{2N-3})$$

$$\text{(A.31)}$$

Put them back to Eq. (A.28) and we can obtain

$$|C_k|^2 d^{2N} N^2 \{ \frac{N-2}{N} [1 + \frac{4}{d} \cos \frac{k}{2} + \frac{1}{d^2} (2 + 6 \cos k)] + \frac{1}{N} [1 + \frac{2}{d} \cos \frac{k}{2} + \frac{1}{d^2} (1 + 4 \cos k)]$$

$$+ \frac{1}{N} [1 + \frac{4}{d} \cos \frac{k}{2} + \frac{1}{d^2} (1 + 4 \cos k)] + \mathcal{O}(d^{-3}) \} = 1 \quad \text{(A.32)}$$

To simplify the whole calculation, the thermodynamics limit is taken $N \to \infty$. The main effect of thermodynamics limit is that the contribution from short two domain wall states (e.g. single-site or two-site excitations) is fully suppressed. Thus, up to $\mathcal{O}(\frac{1}{d^2})$, we have

$$|C_k|^2 d^{2N} N^2 [1 + \frac{4}{d} \cos \frac{k}{2} + \frac{1}{d^2} (2 + 6 \cos k)] = 1 \quad \text{(A.33)}$$

Now, we are ready to evaluate the energy expectation value of our two-domain-wall state,

$\langle k|H_{EF}|k\rangle$. For simplicity, we assume $g_{ij} = 1, \beta_{ij} = \beta$ and reorganize the EF Hamiltonian

$$
\begin{aligned}
\hat{H}_{\mathrm{EF}} &= \sum_{i,j} \frac{1-Z_iZ_j}{2} e^{-\beta X_i X_j - \delta(X_i+X_j)} \\
&= \sum_{i,j} \frac{1-Z_iZ_j}{2} \frac{d^2}{d^2-1} [\cosh\beta - \sinh\beta X_i X_j][1 - \frac{1}{d}(X_i+X_j) + \frac{1}{d^2} X_i X_j] \\
&\equiv \sum_i \frac{1-Z_iZ_{i+1}}{2} \left[a(\beta,d) + b(\beta,d)(X_i+X_{i+1}) + c(\beta,d)X_i X_{i+1} \right]
\end{aligned}
\tag{A.34}
$$

where

$$
\begin{aligned}
a(\beta,d) &= \frac{d^2}{d^2-1}(\cosh\beta - \frac{\sinh\beta}{d^2}) = \cosh\beta + \frac{\cosh\beta - \sinh\beta}{d^2} + \mathscr{O}(\frac{1}{d^4}) \\
b(\beta,d) &= -\frac{d}{d^2-1}(\cosh\beta - \sinh\beta) = -\frac{1}{d}(\cosh\beta - \sinh\beta) + \mathscr{O}(\frac{1}{d^3}) \\
c(\beta,d) &= \frac{d^2}{d^2-1}(\frac{\cosh\beta}{d^2} - \sinh\beta) = \frac{\cosh\beta - \sinh\beta}{d^2} - \sinh\beta + \mathscr{O}(\frac{1}{d^4}).
\end{aligned}
\tag{A.35}
$$

The first term in $\langle k|H_{EF}|k\rangle$ is

$$
\begin{aligned}
&|C_k|^2 a(\beta,d) \sum_{n,m,a,b} e^{ik(n-m)} e^{ik(a-b)/2} \langle \Uparrow | \prod_{i=n}^{n+a} X_i \sum_l \frac{1-Z_lZ_{l+1}}{2} \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle \\
&= 2a(\beta,d)|C_k|^2 \sum_{n,m,a,b} e^{ik(n-m)} e^{ik(a-b)/2} \langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle = 2a(\beta,d)
\end{aligned}
\tag{A.36}
$$

As for the second term, since $b(\beta,d)$ already contains $1/d$ power, we just compute the terms up to $1/d$ order and the result is

$$
\begin{aligned}
&|C_k|^2 b(\beta,d) \sum_{n,m,a,b} e^{ik(n-m)} e^{ik(a-b)/2} \langle \Uparrow | \prod_{i=n}^{n+a} X_i \sum_l \frac{1-Z_lZ_{l+1}}{2} (X_l+X_{l+1}) \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle \\
&= |C_k|^2 b(\beta,d) \sum_{n,m,a,b} e^{ik(n-m)} e^{ik(a-b)/2} \\
&\quad \times [h(n-1,m,a+1,b) + h(n+1,m,a-1,b) + h(n,m,a+1,b) + h(n,m,a-1,b)]
\end{aligned}
\tag{A.37}
$$

where

$$h(n,m,a,b) = \langle \Uparrow | \prod_{i=n}^{n+a} X_i \hat{W}_{\mathbb{1}} \prod_{j=m}^{m+b} X_j | \Uparrow \rangle. \tag{A.38}$$

For each $h(n,m,a,b)$, the boundary terms would have different result. For example, the results of $h(n-1,m,a+1,b)$ are as follows

$$
\begin{aligned}
h(n-1,m,a+1,0) &= (\delta_{n,m+1}\delta_{a,0} + \delta_{n,m}\delta_{a,0})d^{2N-1} + \mathcal{O}(d^{2N-2}) \\
h(n-1,m,a+1,1) &= \delta_{n,m+1}\delta_{a,0}d^{2N} + (\delta_{n,m+1}\delta_{a,1} + \delta_{n,m}\delta_{a,1})d^{2N-1} + \mathcal{O}(d^{2N-2}) \\
h(n-1,m,a+1,2) &= \delta_{n,m+1}\delta_{a,1}d^{2N} \\
&\quad + (\delta_{n,m+1}\delta_{a,2} + \delta_{n,m+1}\delta_{a,0} + \delta_{n,m+2}\delta_{a,0} + \delta_{n,m}\delta_{a,2})d^{2N-1} + \mathcal{O}(d^{2N-2}) \\
h(n-1,m,a+1,b &\neq 0,1,2) = \delta_{n,m+1}\delta_{a,b-1}d^{2N} + (\delta_{n,m+1}\delta_{a,b} + \delta_{n,m+1}\delta_{a,b-2} + \\
&\quad \delta_{n,m+2}\delta_{a,b-2} + \delta_{n,m}\delta_{a,b})d^{2N-1} + \mathcal{O}(d^{2N-2}).
\end{aligned}
\tag{A.39}
$$

Since the thermodynamics limit would be taken ($N \to \infty$), the "boundary effect" from short two-domain-wall states would be suppressed. Consequently, we only keep the last term in our calculation. Combine these four terms and compute the sum with thermodynamic limit,

$$
\begin{aligned}
&|C_k|^2 b(\beta,d)d^{2N}N^2[4\cos\frac{k}{2} + \frac{8}{d}(1+\cos k)] + \mathcal{O}(\frac{1}{d^2}) \\
&= b(\beta,d)[4\cos\frac{k}{2} - \frac{16}{d}\cos^2\frac{k}{2} + \frac{4}{d}(2+2\cos k)] + \mathcal{O}(\frac{1}{d^2})
\end{aligned}
\tag{A.40}
$$

For the third term, the following quantity is computed

$$|C_k|^2 c(\beta,d) \sum_{n,m,a,b} e^{ik(n-m)}e^{ik(a-b)/2}\langle \Uparrow | \prod_{i=n}^{n+a} X_i \sum_l \frac{1-Z_lZ_{l+1}}{2}X_lX_{l+1}\hat{W}_I \prod_{j=m}^{m+b} X_j | \Uparrow \rangle. \tag{A.41}$$

The EF Hamiltonian would give extra X_iX_{i+1} term. In most two-domain-wall states (length > 1), the two-domain-wall structure would destroyed. However, for single site excitation, this X_iX_{i+1} term would only shift the position of excitation with one site. Due to the suppression of thermodynamic limit, we would also drop this term. Eventually, we can have

$$c(\beta,d)\left[\frac{4}{d}\cos\frac{k}{2} + \frac{1}{d^2}(8 + 8\cos k)\right]\left(1 - \frac{4}{d}\cos\frac{k}{2}\right) + \mathcal{O}(\frac{1}{d^3})$$
$$= c(\beta,d)\left[\frac{4}{d}\cos\frac{k}{2} + \frac{1}{d^2}(8 + 8\cos k) - \frac{16}{d^2}\cos^2\frac{k}{2}\right] + \mathcal{O}(\frac{1}{d^3}). \tag{A.42}$$

Combining Eq. (A.36), Eq. (A.40) and Eq. (A.42) and keeping terms up to $\mathcal{O}(\frac{1}{d^3})$, $\langle k|H_{EF}|k\rangle$ would be

$$\langle k|H_{EF}|k\rangle =$$
$$2\left[\cosh\beta + \frac{\cosh\beta - \sinh\beta}{d^2}\right] - \frac{1}{d}(\cosh\beta - \sinh\beta)\left[4\cos\frac{k}{2} - \frac{16}{d}\cos^2\frac{k}{2} + \frac{4}{d}(2 + 2\cos k)\right]$$
$$- \sinh\beta\left[\frac{4}{d}\cos\frac{k}{2} + \frac{1}{d^2}(8 + 8\cos k) - \frac{16}{d^2}\cos^2\frac{k}{2}\right] + \mathcal{O}(\frac{1}{d^3}) \tag{A.43}$$

A.6 Derivation of the Dispersion Relation for Single-Site Excitation ansatz

This appendix is similar with the calculation in Appendix A.5. The only difference is the ansatz state we use. The single-site excitation ansatz is defined as

$$\langle k| = C_k\langle\Uparrow|\sum_n X_n e^{ikn}, \quad |k\rangle = \hat{W}_{\mathbb{1}}\sum_n X_n e^{-ikn}|\Uparrow\rangle. \tag{A.44}$$

First, we start from the normalization condition $\langle k|k\rangle = 1$,

$$\langle k|k\rangle = 1 = C_k\langle\Uparrow|\sum_{n,m} e^{ik(n-m)}X_n\hat{W}_{\mathbb{1}}X_m|\Uparrow\rangle = C_k[Nd^{2(N-1)}(d^2 - 1) + N^2\delta_{k,0}d^{2(N-1)}]. \tag{A.45}$$

Following the expression in Eq. (A.46),

$$\hat{H}_{\text{EF}} = \sum_i \frac{1 - Z_iZ_{i+1}}{2}\left[a(\beta,d) + b(\beta,d)(X_i + X_{i+1}) + c(\beta,d)X_iX_{i+1}\right] \tag{A.46}$$

174

where

$$a(\beta,d) = \frac{d^2}{d^2-1}(\cosh\beta - \frac{\sinh\beta}{d^2}),$$

$$b(\beta,d) = -\frac{d}{d^2-1}(\cosh\beta - \sinh\beta), \qquad (A.47)$$

$$c(\beta,d) = \frac{d^2}{d^2-1}(\frac{\cosh\beta}{d^2} - \sinh\beta).$$

The first term is

$$\begin{aligned}
a(\beta,d)\langle k|\sum_l \frac{1-Z_lZ_{l+1}}{2}W_I|k\rangle &= a(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\sum_l \frac{1-Z_lZ_{l+1}}{2}\hat{W}_{\mathbb{1}}|m\rangle \\
&= a(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\sum_l(\delta_{l,n-1}+\delta_{l,n})\hat{W}_{\mathbb{1}}|m\rangle \\
&= 2a(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\hat{W}_{\mathbb{1}}|m\rangle \\
&= 2a(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}[\delta_{n,m}d^{2N}+(1-\delta_{n,m})d^{2(N-1)}] \\
&= 2a(\beta,d)C_k[Nd^{2(N-1)}(d^2-1)+N^2\delta_{k,0}d^{2(N-1)}] = 2a(\beta,d).
\end{aligned}$$

$$(A.48)$$

The second term is

$$b(\beta,d)\langle k|\sum_l \frac{1-Z_lZ_{l+1}}{2}(X_l+X_{l+1})W_I|k\rangle$$

$$= b(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\sum_l \frac{1-Z_lZ_{l+1}(X_l+X_{l+1})}{2}\hat{W}_{\mathbb{1}}|m\rangle$$

$$= b(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|(X_{n-1}+2X_n+X_{n+1})\hat{W}_{\mathbb{1}}|m\rangle$$

$$= b(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}(\langle n,n-1|W_I|m\rangle+\langle n,n+1|\hat{W}_{\mathbb{1}}|m\rangle+2\langle\Uparrow|\hat{W}_{\mathbb{1}}|m\rangle)$$

$$= b(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}[d^{2N-3}(1-\delta_{m,n})(1-\delta_{m,n-1})+d^{2N-3}(1-\delta_{m,n})(1-\delta_{m,n+1})$$

$$+d^{2N-1}(\delta_{m,n}+\delta_{m,n-1})+d^{2N-1}(\delta_{m,n}+\delta_{m,n+1})+2d^{2N-1}]$$

$$= b(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}[2d^{2N-3}+2d^{2N-1}+(d^{2N-1}-d^{2N-3})(\delta_{m,n-1}+2\delta_{m,n}+\delta_{m,n+1})]$$

$$= b(\beta,d)C_k[N^2\delta_{k,0}2d^{2N-3}(d^2+1)+2d^{2N-3}(d^2-1)N(1+\cos k)].$$

$$(A.49)$$

For single-site excitation, we focus on the region which $k\neq 0$. The result would be

$$2b(\beta,d)\frac{1+\cos k}{d}\times C_k[Nd^{2(N-2)}(d^2-1)]=\frac{2b(\beta,d)}{d}(1+\cos k).\qquad (A.50)$$

The third term is

$$
c(\beta,d)\langle k|\sum_l \frac{1-Z_lZ_{l+1}}{2}X_lX_{l+1}\hat{W}_{\mathbb{1}}|k\rangle
$$

$$
= c(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\sum_l \frac{(1-Z_lZ_{l+1})X_lX_{l+1}}{2}\hat{W}_{\mathbb{1}}|m\rangle
$$

$$
= c(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}\langle n|\sum_l(\delta_{l,n-1}+\delta_{l,n})X_lX_{l+1}\hat{W}_{\mathbb{1}}|m\rangle
$$

$$
= c(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}(\langle n-1|\hat{W}_{\mathbb{1}}|m\rangle+\langle n+1|\hat{W}_{\mathbb{1}}|m\rangle)
$$

$$
= c(\beta,d)C_k\sum_{m,n}e^{ik(n-m)}(2d^{2(N-1)}+(\delta_{m,n-1}+\delta_{m,n+1})(d^{2N}-d^{2(N-1)}))
$$

$$
= 2c(\beta,d)C_k[Nd^{2(N-1)}(d^2-1)\cos k+N^2d^{2(N-1)}\delta_{k,0}] = 2c(\beta,d)\cos k
$$

(A.51)

The overall result would be

$$
\langle k|\hat{H}_{EF}|k\rangle
$$

$$
= 2a(\beta,d)+\frac{2b(\beta,d)}{d}(1+\cos k)+2c(\beta,d)\cos k
$$

$$
= 2a(\beta,d)+\frac{2b(\beta,d)}{d}+\cos k[2c(\beta,d)+\frac{2b(\beta,d)}{d}]
$$

(A.52)

$$
= \frac{2d^2}{d^2-1}(\cosh\beta-\frac{\sinh\beta}{d^2})-\frac{2}{d^2-1}(\cosh\beta-\sinh\beta)
$$

$$
+\cos k[\frac{2d^2}{d^2-1}(\frac{\cosh\beta}{d^2}-\sinh\beta)-\frac{2}{d^2-1}(\cosh\beta-\sinh\beta)]
$$

A.7   Diagrammatic Expansion of Entanglement Feature Hamiltonian

In this appendix, we derive the EF Hamiltonian for the locally scrambled Hamiltonian dynamics. We start from the definition of the EF for $e^{-i\varepsilon H}$ following Eq. (2.5),

$$
W_{e^{-i\varepsilon H}}[\boldsymbol{\sigma},\boldsymbol{\tau}] = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}}(e^{-i\varepsilon H})^{\otimes 2}\mathscr{X}_{\boldsymbol{\tau}}(e^{i\varepsilon H})^{\otimes 2}\right) = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}}e^{-i\varepsilon\mathbb{H}}\mathscr{X}_{\boldsymbol{\tau}}e^{i\varepsilon\mathbb{H}}\right),
$$

(A.53)

where we have introduced $\mathbb{H} = H \otimes \mathbb{1} + \mathbb{1} \otimes H$ to denote the double Hamiltonian. Given the locality of $H = \sum_x H_x$, the double Hamiltonian $\mathbb{H}$ is also a sum of local terms $\mathbb{H} = \sum_x \mathbb{H}_x$ with $\mathbb{H}_x = H_x \otimes \mathbb{1} + \mathbb{1} \otimes H_x$ being the doubled version of H_x. Expand around $\varepsilon \to 0$ to the order of ε^2, we obtain

$$
\begin{aligned}
W_{e^{-i\varepsilon H}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] &= \mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}}) - \varepsilon^2 \, \mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}^2 - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}) + \mathscr{O}(\varepsilon^4), \\
&= W_{\mathbb{1}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] - \varepsilon^2 \sum_{x,x'} \mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x \mathbb{H}_{x'} - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{x'}) + \mathscr{O}(\varepsilon^4), \\
&= W_{\mathbb{1}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] - \varepsilon^2 \sum_x \mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x^2 - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x) + \mathscr{O}(\varepsilon^4)
\end{aligned}
\tag{A.54}
$$

where the first order term in ε vanishes by the cyclic identity of trace, confirming the argument in Sec. 2.2.5 that $W_{U(\varepsilon)}$ will be even in ε. The last equality in Eq. (A.54) relies on the fact that $\mathrm{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x \mathbb{H}_{x'} - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{x'}) = 0$ as long as $x \neq x'$. To prove this, we first consider the case when $x = \langle ij \rangle$ and $x' = \langle kl \rangle$ do not overlap,

$$
\begin{aligned}
\mathrm{Tr} \, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle kl \rangle} &= \mathrm{Tr} \, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \left(\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j} \mathscr{X}_{\tau_k} \mathscr{X}_{\tau_l} \bigotimes_{m \neq i,j,k,l} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle kl \rangle} \\
&= \mathrm{Tr} \, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_k} \mathscr{X}_{\tau_l} \bigotimes_{m \neq i,j,k,l} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle kl \rangle} (\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j}) \\
&= \mathrm{Tr}(\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j}) \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_k} \mathscr{X}_{\tau_l} \bigotimes_{m \neq i,j,k,l} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle kl \rangle} \\
&= \mathrm{Tr} \, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j} \mathscr{X}_{\tau_k} \mathscr{X}_{\tau_l} \bigotimes_{m \neq i,j,k,l} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle kl \rangle} \\
&= \mathrm{Tr} \, \mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle kl \rangle},
\end{aligned}
\tag{A.55}
$$

where we have used the fact that $[\mathbb{H}_{\langle ij \rangle}, \mathscr{X}_{\tau_k} \mathscr{X}_{\tau_l}] = 0$ for $i, j \neq k, l$, and $[\mathscr{X}_{\sigma_i}, \mathscr{X}_{\tau_j}] = 0$ for any i, j as the S_2 group is Abelian. We then consider the case when $x = \langle ij \rangle$ and $x' = \langle jk \rangle$ overlaps

178

on a single site j,

$$\begin{aligned}
\text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle jk \rangle} &= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \left(\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j} \mathscr{X}_{\tau_k} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle jk \rangle} \\
&= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_k} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\tau_j} \mathbb{H}_{\langle jk \rangle} \mathscr{X}_{\tau_i}.
\end{aligned} \tag{A.56}$$

At this point, it seems that $\mathscr{X}_{\tau_j}$ is caught between $\mathbb{H}_{\langle ij \rangle}$ and $\mathbb{H}_{\langle jk \rangle}$. The solution is to make use of the property that $\mathbb{H}_{\langle jk \rangle} = \mathscr{X}_{\alpha_j}^{-1} \mathscr{X}_{\alpha_k}^{-1} \mathbb{H}_{\langle jk \rangle} \mathscr{X}_{\alpha_k} \mathscr{X}_{\alpha_j}$ for any $\alpha_j = \alpha_k \in S_2$, due to the permutation symmetry to exchange the two replicas of the double Hamiltonian. Now we choose $\alpha_j = \alpha_k = \tau_j$, such that $\mathscr{X}_{\tau_j} \mathscr{X}_{\alpha_j}^{-1} = 1$, then

$$\begin{aligned}
\text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle jk \rangle} &= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_k} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\tau_j} \mathscr{X}_{\alpha_j}^{-1} \mathscr{X}_{\alpha_k}^{-1} \mathbb{H}_{\langle jk \rangle} \mathscr{X}_{\alpha_k} \mathscr{X}_{\alpha_j} \mathscr{X}_{\tau_i} \\
&= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_k} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\alpha_k}^{-1} \mathbb{H}_{\langle jk \rangle} \mathscr{X}_{\alpha_k} \mathscr{X}_{\alpha_j} \mathscr{X}_{\tau_i} \\
&= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_i} \mathscr{X}_{\alpha_j} \mathscr{X}_{\alpha_k} \mathscr{X}_{\tau_k} \mathscr{X}_{\alpha_k}^{-1} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle jk \rangle} \\
&= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \left(\mathscr{X}_{\tau_i} \mathscr{X}_{\tau_j} \mathscr{X}_{\tau_k} \bigotimes_{m \neq i,j,k} \mathscr{X}_{\tau_m} \right) \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle jk \rangle} \\
&= \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle jk \rangle}.
\end{aligned} \tag{A.57}$$

Hence we have shown that $\text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_{\langle ij \rangle} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle kl \rangle} = \text{Tr}\, \mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{\langle ij \rangle} \mathbb{H}_{\langle kl \rangle}$ as long as $\langle ij \rangle \neq \langle kl \rangle$, meaning that $\text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x \mathbb{H}_{x'} - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_{x'}) = \delta_{xx'} \text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x^2 - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x)$. Thus the derivation of Eq. (A.54) is justified.

If we consider the difference between $W_{e^{-i\varepsilon H}}$ and $W_{\mathbb{1}}$, denoted as δW,

$$\delta W[\boldsymbol{\sigma}, \boldsymbol{\tau}] \equiv W_{e^{-i\varepsilon H}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] - W_{\mathbb{1}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] = -\varepsilon^2 \sum_x \text{Tr}(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x^2 - \mathscr{X}_{\boldsymbol{\sigma}} \mathbb{H}_x \mathscr{X}_{\boldsymbol{\tau}} \mathbb{H}_x) + \mathscr{O}(\varepsilon^4). \tag{A.58}$$

$\delta W[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \sum_x \delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}] W_{\mathbb{1}_{\bar{x}}}[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ can be expressed as a sum of terms on each bond x (at least to the order of ε^2). To carry out the ε expansion more systematically, we choose to focus on a

single bond, and define the EF difference

$$\delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}] \equiv W_{e^{-i\varepsilon H_x}}[\boldsymbol{\sigma}, \boldsymbol{\tau}] - W_{\mathbb{1}_x}[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} e^{-i\varepsilon\mathbb{H}_x} \mathscr{X}_{\boldsymbol{\tau}} e^{i\varepsilon\mathbb{H}_x}\right) - \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}}\right), \quad \text{(A.59)}$$

where $\boldsymbol{\sigma} = (\sigma_i, \sigma_j)$ is restricted to the two sites i, j connected by the bond x and similarly for $\boldsymbol{\tau}$. $\delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}] = 0$ vanishes as long as $\sigma_i = \sigma_j$ or $\tau_i = \tau_j$, because in that case, $\mathscr{X}_{\boldsymbol{\sigma}}$ or $\mathscr{X}_{\boldsymbol{\tau}}$ will commute with $\mathbb{H}_x$ and hence the two traces will cancel with each other. Therefore there are only two independent non-trivial components of $\delta W[\boldsymbol{\sigma}, \boldsymbol{\tau}]$, which we denote as u and v:

$$u = \delta W\left[|\mathord{\uparrow}_i\mathord{\times}_j, |\mathord{\uparrow}_i\mathord{\times}_j\right] = \delta W\left[\mathord{\times}_i|\mathord{\uparrow}_j, \mathord{\times}_i|\mathord{\uparrow}_j\right],$$
$$v = \delta W\left[|\mathord{\uparrow}_i\mathord{\times}_j, \mathord{\times}_i|\mathord{\uparrow}_j\right] = \delta W\left[\mathord{\times}_i|\mathord{\uparrow}_j, |\mathord{\uparrow}_i\mathord{\times}_j\right]. \quad \text{(A.60)}$$

So we only need to focus on these terms and perform the ε expansion following the definition

$$\delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}] = \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} e^{-i\varepsilon H_x} \otimes e^{-i\varepsilon H_x} \mathscr{X}_{\boldsymbol{\tau}} e^{i\varepsilon H_x} \otimes e^{i\varepsilon H_x}\right) - \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} \mathscr{X}_{\boldsymbol{\tau}}\right)$$
$$= \sum_{k=1}^{\infty} \varepsilon^{2k} \sum_{n_1+n_2+n_3+n_4=2k} \frac{i^{-n_1-n_2+n_3+n_4}}{n_1! n_2! n_3! n_4!} \mathrm{Tr}\left(\mathscr{X}_{\boldsymbol{\sigma}} H_x^{n_1} \otimes H_x^{n_2} \mathscr{X}_{\boldsymbol{\tau}} H_x^{n_3} \otimes H_x^{n_4}\right). \quad \text{(A.61)}$$

The ε odd power terms must vanish because $\delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ must be real. To the ε^4 order, we found

$$
\begin{aligned}
u = {} & \varepsilon^2 \Big(-2\,\substack{\text{\tiny■}} + \big(2d\,\substack{\text{\tiny■}} + 2d\,\substack{\text{\tiny■}}\big) - \tfrac{1}{2!}\big(4d^2\,\substack{\text{\tiny■}}\big)\Big) \\
& + \varepsilon^4 \Big(\substack{\text{\tiny■}} + \tfrac{1}{2!}\big(4\,\substack{\text{\tiny■}} - 4\,\substack{\text{\tiny■}} - 4\,\substack{\text{\tiny■}}\big) + \tfrac{1}{(2!)^2}\big(2\,\substack{\text{\tiny■}} + 2d\,\substack{\text{\tiny■}} + 2d\,\substack{\text{\tiny■}}\big) \\
& \quad + \tfrac{1}{3!}\big(4\,\substack{\text{\tiny■}} - 4d\,\substack{\text{\tiny■}} - 4d\,\substack{\text{\tiny■}}\big) + \tfrac{1}{4!}\big(4d^2\,\substack{\text{\tiny■}}\big)\Big) + \mathcal{O}(\varepsilon^6) \\
= {} & -\varepsilon^2\big(2R^{(1)(2)}_{(1)(2)} - 2d\big(R^{(12)}_{(1)(2)} + R^{(1)(2)}_{(12)}\big) + 2d^2 R^{(12)}_{(12)}\big) \\
& + \varepsilon^4\big(R^{(13)(24)}_{(12)(34)} + 2\big(R^{(124)(3)}_{(123)(4)} - R^{(123)(4)}_{(123)(4)} - R^{(123)(4)}_{(12)(34)}\big) + \tfrac{1}{2}\big(R^{(12)(34)}_{(12)(34)} + dR^{(1234)}_{(12)(34)} + dR^{(12)(34)}_{(1234)}\big) \\
& \quad + \tfrac{2}{3}\big(R^{(123)(4)}_{(123)(4)} - dR^{(1234)}_{(123)(4)} - dR^{(123)(4)}_{(1234)}\big) + \tfrac{1}{6}d^2 R^{(1234)}_{(1234)}\big) + \mathcal{O}(\varepsilon^6), \\
v = {} & \varepsilon^4 \Big(\substack{\text{\tiny■}} + \tfrac{4-8}{2!}\,\substack{\text{\tiny■}} + \big(\tfrac{6}{(2!)^2} + \tfrac{4-8}{3!} + \tfrac{4}{4!}\big)\substack{\text{\tiny■}}\Big) + \mathcal{O}(\varepsilon^6) \\
= {} & \varepsilon^4\big(R^{(1432)}_{(1234)} - 2R^{(1243)}_{(1234)} + R^{(1234)}_{(1234)}\big) + \mathcal{O}(\varepsilon^6).
\end{aligned}
\tag{A.62}
$$

In the diagrams, each small red block represents a copy of the bond Hamiltonian H_x. Their legs are contracted according to the assignment of the permutations $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$. The result can be expressed in terms of the generalized spectral form factor $R^{g_j}_{g_i}$, labeled by two permutations $g_i, g_j \in S_n$ acting separately on sites i and j,

$$
R^{g_j}_{g_i} = \mathrm{Tr}(H^{\otimes n}_{ij}\, \mathscr{X}_{g_i g_j}).
\tag{A.63}
$$

where $\mathscr{X}_{g_i g_j} = \mathscr{X}_{g_i} \mathscr{X}_{g_j}$ is the representation of g_i and g_j in the n-replicated Hilbert space. For example, $R^{(1)(2)}_{(1)(2)} = (\mathrm{Tr}\,H_{ij})^2$, $R^{(12)}_{(1)(2)} = \mathrm{Tr}_j(\mathrm{Tr}_i H_{ij})^2$ (where Tr_i denotes the partial trace over site i), and $R^{(12)}_{(12)} = \mathrm{Tr}(H^2_{ij})$.

Given the components u and v, we can rewrite $\delta W_x[\boldsymbol{\sigma}, \boldsymbol{\tau}]$ in the operator form

$$
\delta\hat{W}_x = \frac{1 - Z_i Z_j}{2}(u + v X_i X_j)\frac{1 - Z_i Z_j}{2},
\tag{A.64}
$$

therefore the EF operator reads

$$\hat{W}_{e^{-i\varepsilon H}} = \hat{W}_{\mathbb{1}} + \sum_x \delta \hat{W}_x \otimes \hat{W}_{\mathbb{1}_{\bar{x}}} = \hat{W}_{\mathbb{1}} + \sum_{ij} \frac{1 - Z_i Z_j}{2} (u + v X_i X_j) \frac{1 - Z_i Z_j}{2} \otimes \hat{W}_{\mathbb{1}_{\overline{ij}}}. \tag{A.65}$$

The EF Hamiltonian is therefore given by

$$
\begin{aligned}
\hat{H}_{\text{EF}} &= \frac{1}{\varepsilon^2} (\mathbb{1} - \hat{W}_{e^{-i\varepsilon H}} \hat{W}_{\mathbb{1}}^{-1}) \\
&= -\frac{1}{\varepsilon^2} \sum_{ij} \frac{1 - Z_i Z_j}{2} (u + v X_i X_j) \frac{1 - Z_i Z_j}{2} \hat{W}_{\mathbb{1}_{ij}}^{-1} \\
&= -\frac{1}{\varepsilon^2} \sum_{ij} \frac{1 - Z_i Z_j}{2} (u + v X_i X_j) \frac{1 - Z_i Z_j}{2} \frac{1}{d^2(d^2 - 1)} e^{-\delta(X_i + X_j)} \\
&= -\sum_{ij} \frac{1 - Z_i Z_j}{2} \frac{u + v X_i X_j}{\varepsilon^2 d^2 (d^2 - 1)} e^{-\delta(X_i + X_j)}
\end{aligned}
\tag{A.66}
$$

Therefore the EF Hamiltonian generally take the form of

$$\hat{H}_{\text{EF}} = \sum_{ij} g \frac{1 - Z_i Z_j}{2} e^{-\beta X_i X_j - \delta(X_i + X_j)}, \tag{A.67}$$

consistent with the general form in Eq. (2.32). Comparing Eq. (A.66) with Eq. (A.67), we should identify

$$g e^{-\beta X_i X_j} = -\frac{u + v X_i X_j}{\varepsilon^2 d^2 (d^2 - 1)}, \tag{A.68}$$

which indicates

$$
\begin{aligned}
g \cosh \beta &= -\frac{u}{\varepsilon^2 d^2 (d^2 - 1)} = \frac{1}{d^2 (d^2 - 1)} (u_2 - u_4 \varepsilon^2 + \mathcal{O}(\varepsilon^4)) \\
g \sinh \beta &= \frac{v}{\varepsilon^2 d^2 (d^2 - 1)} = \frac{1}{d^2 (d^2 - 1)} (v_4 \varepsilon^2 + \mathcal{O}(\varepsilon^4)),
\end{aligned}
\tag{A.69}
$$

where the coefficients u_2, u_4, v_4 are defined in terms of generalized spectral form factors R^{gj}_{gi} as

$$u_2 = 2R^{(1)(2)}_{(1)(2)} - 2d\left(R^{(12)}_{(1)(2)} + R^{(1)(2)}_{(12)}\right) + 2d^2 R^{(12)}_{(12)},$$

$$u_4 = R^{(13)(24)}_{(12)(34)} + 2\left(R^{(124)(3)}_{(123)(4)} - R^{(12)(34)}_{(123)(4)} - R^{(123)(4)}_{(12)(34)}\right) + \frac{1}{2}\left(R^{(12)(34)}_{(12)(34)} + dR^{(1234)}_{(12)(34)} + dR^{(12)(34)}_{(1234)}\right)$$

$$\qquad + \frac{2}{3}\left(R^{(123)(4)}_{(123)(4)} - dR^{(1234)}_{(123)(4)} - dR^{(123)(4)}_{(1234)}\right) + \frac{1}{6}d^2 R^{(1234)}_{(1234)}, \tag{A.70}$$

$$v_4 = R^{(1432)}_{(1234)} - 2R^{(1243)}_{(1234)} + R^{(1234)}_{(1234)}.$$

For specific model of H_{ij}, we can evaluate the generalized spectral form factors, then we can determined the parameters g and β as well as the EF Hamiltonian. In the following, we will perform the calculation for random $U(d)$ spin model and the locally scrambled Ising model.

For two-qudit GUE Hamiltonians, the generalized spectral form factors, defined in Eq. (A.63), can be evaluated under the GUE average using the basic property that

$$\mathop{\mathbb{E}}_{\text{GUE}} H^{\otimes 2}_{ij} \equiv \mathop{\mathbb{E}}_{\text{GUE}} \boxed{H^{\otimes 2}_{ij}}$$

$$= \frac{1}{d^2}\left\Vert h_i\ h_j \right\Vert \equiv \frac{1}{d^2}\mathscr{X}_{(12)_i(12)_j}, \tag{A.71}$$

the GUE average of n-replicated Hamiltonian H_{ij} can be obtained by summing over Wick contractions

$$\mathop{\mathbb{E}}_{\text{GUE}} H^{\otimes n}_{ij} = \begin{cases} d^{-n}\sum_{h_i=h_j\in P_n} \mathscr{X}_{h_i h_j} & n \in \text{even}, \\ 0 & n \in \text{odd}, \end{cases} \tag{A.72}$$

where P_n denotes all possible pair-wise exchange of n replicas. Then the generalized spectral form factor reads

$$\mathop{\mathbb{E}}_{\text{GUE}} R^{gj}_{gi} = \frac{1}{d^n}\sum_{h\in P_n} \text{Tr}(\mathscr{X}_{g_i}\mathscr{X}_h)\,\text{Tr}(\mathscr{X}_{g_j}\mathscr{X}_h), \tag{A.73}$$

whose results are enumerated in Tab. A.1. Substitute these results to Eq. (A.70), we find $u_2 = 2(d^2 - 1)^2$, $u_4 = \frac{11}{6}(d^2 - 1)^2$, and $v_4 = 2(d^2 - 1)^2/d^2$. By solving Eq. (A.69), we can determine

the parameters g and β to the order of ε^2,

$$g = 2(1 - d^{-2})\big(1 - \tfrac{11}{12}\varepsilon^2 + \mathcal{O}(\varepsilon^4)\big),$$
$$\beta = \varepsilon^2/d^2 + \mathcal{O}(\varepsilon^4). \tag{A.74}$$

In conclusion, as we consider the locally scrambled quantum dynamics by alternatively applying the small unitary $e^{-i\varepsilon H}$ and the local scramblers, the evolution of the corresponding EF state will be governed by $\partial_t |W_{\Psi_t}\rangle = -\hat{H}_{\text{EF}}|W_{\Psi_t}\rangle$, with the EF Hamiltonian $\hat{H}_{\text{EF}}$ given by Eq. (A.67). The random U(d) spin model H in Eq. (2.74) corresponds to the set of parameters in Eq. (A.74) for $\hat{H}_{\text{EF}}$.

Table A.1. Spectral form factor of two-qudit GUE Hamiltonian

$R^{(1)(2)}_{(1)(2)}$		1	$R^{(12)}_{(1)(2)}$		d	$R^{(1)(2)}_{(12)}$		d
$R^{(123)(4)}_{(123)(4)}$		3	$R^{(124)(3)}_{(123)(4)}$		3	$R^{(1234)}_{(123)(4)}$		$2d + \frac{1}{d}$
$R^{(12)(34)}_{(123)(4)}$		$d^2 + 2$	$R^{(1243)}_{(1234)}$		$d^2 + 2$	$R^{(123)(4)}_{(12)(34)}$		$d^2 + 2$
$R^{(1234)}_{(1234)}$		$2d^2 + \frac{1}{d^2}$	$R^{(1432)}_{(1234)}$		$2d^2 + \frac{1}{d^2}$	$R^{(12)(34)}_{(1234)}$		$d^3 + d + \frac{1}{d}$
$R^{(12)(34)}_{(12)(34)}$		$d^4 + 2$						
$R^{(12)}_{(12)}$		d^2						
$R^{(123)(4)}_{(1234)}$		$2d + \frac{1}{d}$						
$R^{(13)(24)}_{(12)(34)}$		$2d^2 + 1$						
$R^{(1234)}_{(12)(34)}$		$d^3 + d + \frac{1}{d}$						

184

Appendix B

MPS Ansatz

B.1   Efficacy of the $D = 2$ MPS

We can use the TEBD approach to evolve a $D = 2$ MPS state to calculate the entanglement dynamics governed by $-\partial_t |W_\Psi\rangle = H_{\text{EF}} |W_\Psi\rangle$. By comparing MPS the result with the exact numerical solution the differential equation, we found that $D = 2$ successfully captures the evolution of all the multi-region entanglement over the entire process of thermalization from a product state, as shown in Fig. B.1.

We find that the $D = 2$ MPS captures the full entanglement continuum regardless of the parameters g, β of locally scrambled evolution. This is expected from the sign structure of the entanglement feature states[80]. Since generic positive vectors have constant law entanglement, it is expected that generic states (states that are not fine-tuned) will be well described by MPS. The exception is near the entanglement transition where the bond dimension is expected to diverge [62].

TEBD on the MPS ansatz works as in Fig. B.2. We apply the two-local EF transfer matrix to the product of local MPS tensors. We separate the result into two tensors, L and R which are given by doing an SVD truncation on the product, taking only the two largest singular values. These tensors are then updated again, but in reversed order, by the transfer matrix according to the brick wall arrangement of the circuit. The resulting two tensors are identical and become the next MPS tensor M.
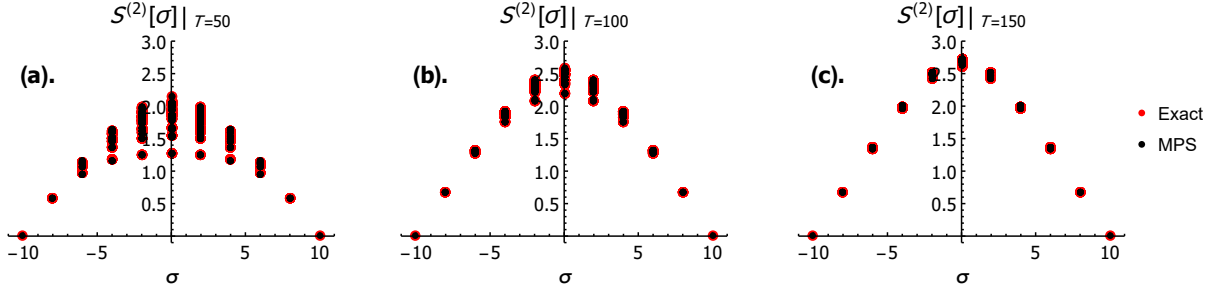
Figure B.1. The three figures show the evolution of the Renyi entropy in the fractional swap gate for swap probability $x = 0.1$ at three different time slices: $T = 50, 100$ and 150 for figures (a),(b), and (c) respectively. The x-axis is the magnetization, which is related to the size of the entangling region by $|A| = (\sigma + N)/2$. Initially, at $T = 0$, the state starts off as a product state and all bipartite entropies are zero. Then, as it evolves, it takes the generic form of the multi-region entanglement continuum, lower-bounded by the single-region entanglement which is area-law. This can be observed in (a). It gradually becomes dominated by the volume law term, at which point the area law plateau begins to vanish, as seen in (b). As it approaches the Page state, which is a pure volume law state, the multi-region entanglement continuum collapses into a single curve so that $S^{(2)}(A) = S^{(2)}_{\min}(|A|)$, as we can see in (c). The red points are given by the exact numerics, whereas the black points are given by the $D = 2$ MPS evolved by TEBD. We see nearly perfect alignment between the two for the full multi-region entanglement continuum for arbitrary choice of EF parameters g, β.
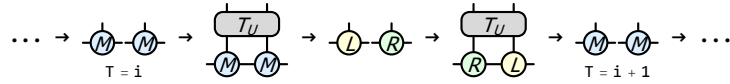
Figure B.2. The Time Evolved Block Decimation (TEBD) procedure for updating the tensor M in the translationally invariant MPS EF state. First, we apply an even layer of transfer matrices on the current state which is given by a product of M tensors. After SVD truncation on the resulting tensor, we procure two new tensors L and R. This is by taking the left (right) legs as the input (output) legs of a matrix and calculating its SVD decomposition $usv^{\dagger}$, truncating all but the two largest singular values ($D = 2$), separating the result into a product of two matrices $u\sqrt{s}$ and $\sqrt{s}v^{\dagger}$, and then reshaping these resulting matrices into L and R, respectively. Note that these tensors may break the translational symmetry in the system. Applying the odd layer now, the right tensor of the previous layer is the left tensor of the new layer. We now do SVD truncation once more to get the new M tensors. These resulting tensors should be the same on the left and right, and so it suffices to just take the left or the right one. In principle, because the circuit breaks the translational symmetry into a two-site translational symmetry, it is possible that the tensors are not the same. However, we find that translational symmetry in the MPS is not broken as we evolve the circuit.

B.2 Derivation of Edges of Multi-Region Continuum

We provide the detailed derivation of $S_{\min}(|A|)$ and $S_{\max}(|A|)$ here. Our starting point is the MPS ansatz $M^{\sigma}_{(\alpha,\theta)}$ in Eq. (3.7), which can be written as

$$M^{\sigma} = e^{\alpha(\sin\theta\, X + \sigma\cos\theta\, Z)}. \tag{B.1}$$

We have suppressed the subscript (α, θ) for simplicity. We first evaluate the following matrix product

$$
\begin{aligned}
(M^{\downarrow})^m (M^{\uparrow})^n &= e^{\alpha m(\sin\theta\, X - \cos\theta\, Z)} e^{\alpha n(\sin\theta\, X + \cos\theta\, Z)} \\
&= \big(\cosh\alpha m + \sinh\alpha m(\sin\theta\, X - \cos\theta\, Z)\big)\big(\cosh\alpha n + \sinh\alpha n(\sin\theta\, X + \cos\theta\, Z)\big) \\
&= \cosh\alpha m \cosh\alpha n \big(1 + \tanh\alpha m(\sin\theta\, X - \cos\theta\, Z)\big)\big(1 + \tanh\alpha n(\sin\theta\, X + \cos\theta\, Z)\big) \\
&= \cosh\alpha m \cosh\alpha n \big(1 + (\tanh\alpha m + \tanh\alpha n)\sin\theta\, X + (\tanh\alpha m - \tanh\alpha n)\cos\theta\, Z \\
&\qquad\qquad \tanh\alpha m \tanh\alpha n(-\cos 2\theta - i\sin 2\theta\, Y)\big) \\
&= \cosh\alpha m \cosh\alpha n(c_0 + c_1 X + ic_2 Y + c_3 Z),
\end{aligned}
\tag{B.2}
$$

where we have introduced the coefficients $c_{0,1,2,3}$ as

$$
\begin{aligned}
c_0 &= 1 - \tanh\alpha m \tanh\alpha n \cos 2\theta, \\
c_1 &= (\tanh\alpha m + \tanh\alpha n)\sin\theta, \\
c_2 &= -\tanh\alpha m \tanh\alpha n \sin 2\theta, \\
c_3 &= (\tanh\alpha m - \tanh\alpha n)\cos\theta.
\end{aligned}
\tag{B.3}
$$

The eigenvalues of $(M^{\downarrow})^m (M^{\uparrow})^n$ are given by $\mu_{\pm} = c_{\pm}\cosh\alpha m \cosh\alpha n$ with

$$c_{\pm} = c_0 \pm \sqrt{c_1^2 - c_2^2 + c_3^2}. \tag{B.4}$$

We notice that

$$c_1^2 - c_2^2 + c_3^2$$

$$= (\tanh \alpha m + \tanh \alpha n)^2 \sin^2 \theta + (\tanh \alpha m - \tanh \alpha n)^2 \cos^2 \theta - \tanh^2 \alpha m \tanh^2 \alpha n \sin^2 2\theta$$

$$= (\tanh^2 \alpha m + \tanh^2 \alpha n) - 2\tanh \alpha m \tanh \alpha n \cos 2\theta - \tanh^2 \alpha m \tanh^2 \alpha n \sin^2 2\theta$$

$$= (1 - \tanh \alpha m \tanh \alpha n \cos 2\theta)^2 - (1 - \tanh^2 \alpha m)(1 - \tanh^2 \alpha n)$$

$$= c_0^2 - \text{sech}^2 \alpha m \, \text{sech}^2 \alpha n.$$

$$(B.5)$$

So we have the following relation

$$\sqrt{c_+ c_-} = \sqrt{c_0^2 - (c_1^2 - c_2^2 + c_3^2)} = \text{sech}\, \alpha m \, \text{sech}\, \alpha n. \qquad (B.6)$$

These results are useful to evaluate the trace of $((M^\downarrow)^m (M^\uparrow)^n)^p$. Given the eigenvalues $\mu_\pm$ of $(M^\downarrow)^m (M^\uparrow)^n$,

$$\text{Tr}((M^\downarrow)^m (M^\uparrow)^n)^p = \mu_+^p + \mu_-^p$$

$$= \cosh^p \alpha m \cosh^p \alpha n (c_+^p + c_-^p)$$

$$= \cosh^p \alpha m \cosh^p \alpha n (\sqrt{c_+ c_-})^p \left(\left(\tfrac{c_+}{c_-}\right)^{p/2} + \left(\tfrac{c_-}{c_+}\right)^{p/2} \right) \qquad (B.7)$$

$$= \cosh^p \alpha m \cosh^p \alpha n \, \text{sech}^p \alpha m \, \text{sech}^p \alpha n \left(\left(\tfrac{c_+}{c_-}\right)^{p/2} + \left(\tfrac{c_-}{c_+}\right)^{p/2} \right)$$

$$= \left(\tfrac{c_+}{c_-}\right)^{p/2} + \left(\tfrac{c_-}{c_+}\right)^{p/2}.$$

If we define a new parameter η via $e^\eta = (\frac{c_+}{c_-})^{1/2}$, the trace in Eq. (B.7) can be written as $\text{Tr}((M^\downarrow)^m(M^\uparrow)^n)^p = e^{\eta p} + e^{-\eta p} = 2\cosh \eta p$. In particular,

$$
\begin{aligned}
\cosh \eta &= \tfrac{1}{2}\left(\left(\tfrac{c_+}{c_-}\right)^{1/2} + \left(\tfrac{c_-}{c_+}\right)^{1/2}\right) \\
&= \frac{c_+ + c_-}{2\sqrt{c_+ c_-}} \\
&= c_0 \cosh \alpha m \cosh \alpha n \\
&= \cosh \alpha m \cosh \alpha n - \sinh \alpha m \sinh \alpha n \cos 2\theta \\
&= (\cosh \alpha m \cosh \alpha n + \sinh \alpha m \sinh \alpha n)\sin^2 \theta + (\cosh \alpha m \cosh \alpha n - \sinh \alpha m \sinh \alpha n)\cos^2 \theta \\
&= \sin^2 \theta \cosh \alpha(m+n) + \cos^2 \theta \cosh \alpha(m-n).
\end{aligned}
$$

$$(\text{B.8})$$

In conclusion, we have arrived at a trace formula

$$
\text{Tr}((M^\downarrow)^m(M^\uparrow)^n)^p = 2\cosh \eta p,
$$
$$(\text{B.9})$$
$$
\text{with } \cosh \eta = \sin^2 \theta \cosh \alpha(m+n) + \cos^2 \theta \cosh \alpha(m-n).
$$

Using the trace formula, we can now evaluate the bottom and upper edge entanglement entropies based on the MPS ansatz. We first calculate $S_{\min}(|A|)$,

$$
\begin{aligned}
S_{\min}(|A|) &= -\ln \text{Tr}(M^\downarrow)^{|A|}(M^\uparrow)^{N-|A|} + S_0 \\
&= -\ln\left(2(\sin^2 \theta \cosh \alpha N + \cos^2 \theta \cosh \alpha(N - 2|A|))\right) + \ln(2\cosh \alpha N) \\
&= -\ln\left(\sin^2 \theta + \cos^2 \theta \frac{\cosh \alpha(N-2|A|)}{\cosh \alpha N}\right),
\end{aligned}
$$
$$(\text{B.10})$$

which gives Eq. (3.10). We then calculate $S_{\max}(|A|)$,

$$
\begin{aligned}
S_{\max}(|A|) &= -\ln \text{Tr}\left(M^\downarrow(M^\uparrow)^{N/|A|-1}\right)^{|A|} + S_0 \\
&= -\ln(2\cosh \eta |A|) + \ln(2\cosh \alpha N) \\
&= -\ln \frac{\cosh \eta |A|}{\cosh \alpha N},
\end{aligned}
$$
$$(\text{B.11})$$

where η is set by $\cosh \eta = \sin^2 \theta \cosh \alpha \frac{N}{|A|} + \cos^2 \theta \cosh \alpha \left(\frac{N}{|A|} - 2 \right)$, as claimed in Eq. (3.11).

Now we consider the thermodynamic limit $(N \to +\infty)$ of $S_{\min}(|A|)$ and $S_{\max}(|A|)$. We note that

$$
\begin{aligned}
\frac{\cosh \alpha (N - 2|A|)}{\cosh \alpha N} &= \frac{\cosh \alpha N \cosh 2\alpha |A| - \sinh \alpha N \sinh 2\alpha |A|}{\cosh \alpha N} \\
&= \cosh 2\alpha |A| - \tanh \alpha N \sinh 2\alpha |A| \\
&\xrightarrow{N \to +\infty} \cosh 2\alpha |A| - \sinh 2\alpha |A| = e^{-2\alpha |A|},
\end{aligned} \tag{B.12}
$$

therefore $S_{\min}(|A|) = -\ln(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha |A|})$ as in Eq. (3.12). Similarly we have

$$
\begin{aligned}
\cosh \eta &= \left(\sin^2 \theta + \cos^2 \theta \frac{\cosh \left(\frac{\alpha N}{|A|} - 2\alpha \right)}{\cosh \frac{\alpha N}{|A|}} \right) \cosh \frac{\alpha N}{|A|} \\
&\xrightarrow{N \to +\infty} (\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}) \cosh \frac{\alpha N}{|A|}.
\end{aligned} \tag{B.13}
$$

Take the inverse cosh function on both sides,

$$
\begin{aligned}
\eta &= \operatorname{arccosh} \left((\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}) \cosh \frac{\alpha N}{|A|} \right) \\
&\xrightarrow{N \to +\infty} \frac{\alpha N}{|A|} + \ln(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}),
\end{aligned} \tag{B.14}
$$

therefore

$$
\begin{aligned}
S_{\max}(|A|) &= -\ln \frac{\cosh \eta |A|}{\cosh \alpha N} \\
&= -\ln \frac{\cosh(\alpha N + |A| \ln(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}))}{\cosh \alpha N} \\
&\xrightarrow{N \to +\infty} -\ln e^{|A| \ln(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha})} \\
&= -|A| \ln(\sin^2 \theta + \cos^2 \theta \, e^{-2\alpha}).
\end{aligned} \tag{B.15}
$$

as in Eq. (3.12).

B.3 Dynamics of MPS Parameters

In this section, we will derive the dynamic equation for MPS parameters and explain the numerical details in solving the equation. Our starting point is the imaginary-time Schrödinger equation $-\partial_t |W_\Psi\rangle = H_{\text{EF}} |W_\Psi\rangle$, which governs the evolution of entanglement feature state $|W_\Psi\rangle$. As we represent $|W_\Psi\rangle$ as a $D = 2$ MPS proposed in Eq. (3.6), we would like to approximate the time evolution generated by H_{EF} without leaving the variational manifold of the MPS ansatz Eq. (3.7). Let us denote the MPS parameters (α, θ) jointly as a vector q. Within the variational manifold, the entanglement feature state could only evolve in the tangent plane as $\partial_t |W_q\rangle = |\partial_a W_q\rangle \dot{q}_a$, where $\dot{q}_a \equiv \partial_t q_a$ and $\partial_a \equiv \partial_{q_a}$. We seek the optimal choice of $\dot{q}_a$ such that $-\partial_t |W_q\rangle = -|\partial_a W_q\rangle \dot{q}_a$ best approximates $H_{\text{EF}} |W_q\rangle$. The solution is given by minimizing the loss function

$$\mathscr{L}(\dot{q}) = \||\partial_a W_q\rangle \dot{q}_a + H_{\text{EF}} |W_q\rangle\|^2. \tag{B.16}$$

To define the loss function $\mathscr{L}$, we also need to specify how to take the norm of the entanglement feature state. It is desired that the inner product of entanglement feature states $(\langle W_{q'}|, |W_q\rangle)$ is such defined that H_{EF} is self-adjoint, i.e. $(\langle W_{q'}|, H_{\text{EF}} |W_q\rangle) = (\langle W_{q'}| H_{\text{EF}}^{\mathsf{T}}, |W_q\rangle)$. Since $H_{\text{EF}} \neq H_{\text{EF}}^{\mathsf{T}}$ itself is not transpose symmetric, the inner product must involve a non-trivial metric, which turns out to be given by the following operator

$$W_{\mathbb{1}}^{-1} = (\tanh \delta \sinh \delta)^N e^{-\delta \sum_i X_i}, \tag{B.17}$$

where $\delta = \frac{1}{2} \ln \frac{d+1}{d-1}$ is fixed by the qudit dimension d of the quantum system. Therefore the norm of $|W_q\rangle$ should be defined as $\||W_q\rangle\| = \langle W_q | W_{\mathbb{1}}^{-1} |W_\Psi\rangle$. So the loss function can be expanded as a quadratic form of $\dot{q}$,

$$\mathscr{L}(\dot{q}) = \dot{q}_a G_{ab} \dot{q}_b + 2 h_a \dot{q}_a + c, \tag{B.18}$$

with the coefficients given by

$$
\begin{aligned}
G_{ab} &= \langle \partial_a W_q | W_{\mathbb{1}}^{-1} | \partial_b W_q \rangle, \\
h_a &= \langle \partial_a W_q | W_{\mathbb{1}}^{-1} H_{\mathrm{EF}} | W_q \rangle, \\
c &= \langle W_q | H_{\mathrm{EF}}^{\mathsf{T}} W_{\mathbb{1}}^{-1} H_{\mathrm{EF}} | W_q \rangle.
\end{aligned}
\tag{B.19}
$$

The minimum of $\mathcal{L}(\dot{q})$ is determined by

$$
G_{ab}\dot{q}_b = -h_a,
\tag{B.20}
$$

which gives the dynamic equation in Eq. (3.28). It can be formally written as $\dot{q} = -G^{-1}h$, but G can become singular in the thermodynamic limit $N \to \infty$, which requires more detailed treatment.

Now we explain how to evaluate G and h given the MPS parameters. Given the MPS $|W_q\rangle$

$$
|W_q\rangle = \quad \text{} \quad ,
\tag{B.21}
$$

G and h can be represented as the following tensor networks,

$$
\tag{B.22}
$$

where M denotes the MPS tensor given in Eq. (3.7) and $(H_{\mathrm{EF}})_{jk}$ denotes the term in H_{EF} on the $\langle jk \rangle$ link which reads $\frac{g}{2}(1 - Z_j Z_k)e^{-\delta(X_j + X_k) - \beta X_j X_k}$. The derivatives of the MPS tensor M are

192

given by

$$\partial_\alpha M^\sigma = \sinh\alpha\, I + \cosh\alpha(\sin\theta\, X + \sigma\cos\theta\, Z),$$

$$\partial_\theta M^\sigma = \cosh\alpha\, I + \sinh\alpha(\cos\theta\, X - \sigma\sin\theta\, Z). \tag{B.23}$$

To evaluate these tensor networks, we introduce the transfer operator

$$T = \quad \diamond W_{\mathbb{1}}^{-1}. \tag{B.24}$$

Let $|\tau\rangle$ be the leading eigenvector of T with the eigenvalue 1 (if the leading eigenvalue of T is not 1, we rescale $W_{\mathbb{1}}^{-1}$ to make it 1), such that

$$\tag{B.25}$$

We denote the pseudo inverse of $(1 - T)$ as $\Pi = (1 - T)^{-1}$, such that in the large N limit, the ladder diagram reads

$$\sum_{n=0}^{N} T^n = \quad \cdots \quad = \quad \boxed{\Pi} + N \quad . \tag{B.26}$$

With these preparations, we can show that G and h scales with the system size N in the following manner

$$G = N G^{(1)} + N^2 G^{(2)},$$

$$h = N h^{(1)} + N^2 h^{(2)}, \tag{B.27}$$

193

with $G^{(1,2)}$ and $h^{(1,2)}$ given by the following tensor networks

$$(B.28)$$

According to the dynamic equation Eq. (B.20), time-derivatives of MPS parameters are determined by $\dot{q} = -G^{-1}h$. To calculate the inverse of G, we note that G is a 2×2 real symmetric matrix, so G^{-1} can be expressed as

$$G^{-1} = \frac{2JGJ}{\operatorname{Tr} JGJG}, \qquad (B.29)$$

where $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Further using the form in Eq. (B.27), we have

$$
\begin{aligned}
\dot{q} &= -\frac{2JGJh}{\operatorname{Tr} JGJG} \\
&= -\frac{2\left(N^2 JG^{(1)}Jh^{(1)} + N^3 JG^{(2)}Jh^{(1)} + N^3 JG^{(1)}Jh^{(2)} + N^4 JG^{(2)}Jh^{(2)}\right)}{N^2 \operatorname{Tr} JG^{(1)}JG^{(1)} + 2N^3 \operatorname{Tr} JG^{(1)}JG^{(2)} + N^4 \operatorname{Tr} JG^{(2)}JG^{(2)}}.
\end{aligned}
\qquad (B.30)
$$

An important observation is that $JG^{(2)}Jh^{(2)} = 0$ and $JG^{(2)}JG^{(2)} = 0$, because $G^{(2)}$ and $h^{(2)}$ has the structure of $G^{(2)} = |\gamma\rangle\langle\gamma|$ and $h^{(2)} = |\gamma\rangle\eta$ where $|\gamma\rangle$ is a two-component vector and η is a real number, such that the vanishing $\langle\gamma|J|\gamma\rangle = 0$ (due to the antisymmetric nature of J) results in the

194

vanish of $JG^{(2)}Jh^{(2)}$ and $JG^{(2)}JG^{(2)}$. Then both the numerator and denominator of Eq. (B.30) is dominated by the N^3 term in the $N \to \infty$ limit. Therefore, we can evaluate the time derivative $\dot{q}$ from

$$\dot{q} = -\frac{JG^{(2)}Jh^{(1)} + JG^{(1)}Jh^{(2)}}{\text{Tr}\,JG^{(1)}JG^{(2)}}. \tag{B.31}$$

By iteratively updating $q \to q + \dot{q}\,\mathrm{d}t$ and calculating $\dot{q}$ from the diagrams in Eq. (B.28), we can obtain the time-evolution of the MPS parameters numerically.

B.4   Calculating OTOC

Here we explain how the OTOC is calculated. We start from the expansion in Eq. (3.41)

$$\text{OTOC}(i,j;t) = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!\,d^{N+2}} \langle i|H_{\text{EF}}^k W_{\mathbb{1}} P|j\rangle, \tag{B.32}$$

where $W_{\mathbb{1}} = \prod_i (d^2 + dX_i)$, $P = \prod_i X_i$, and $|i\rangle$ denotes the Ising basis state with a down-spin only at site-i and up-spins elsewhere. The entanglement feature Hamiltonian takes the form of Eq. (3.24),

$$H_{\text{EF}} = \sum_{\langle ij \rangle} \frac{1 - Z_i Z_j}{2} (u - v(X_i + X_j) + wX_iX_j), \tag{B.33}$$

with parameters u, v, w given by Eq. (3.25),

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{g\cosh\beta}{d^2 - 1} \begin{pmatrix} d^2 - \tanh\beta \\ d - d\tanh\beta \\ 1 - d^2\tanh\beta \end{pmatrix}. \tag{B.34}$$

We can combine the d^{-N} factor and the operator $W_{\mathbb{1}}P$ together to define

$$F \equiv \frac{1}{d^N} W_{\mathbb{1}} P = \prod_i (1 + dX_i). \tag{B.35}$$

Such that Eq. (B.32) can be simplified a little bit,

$$\text{OTOC}(i, j; t) = \frac{1}{d^2} \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \langle i | H_{\text{EF}}^k F | j \rangle.$$

(B.36)

We will be able to evaluate $\langle i | H_{\text{EF}}^k F | j \rangle$ for leading orders in H_{EF}, which will provide the OTOC in the short-time limit (the expansion can be thought as controlled by the small parameter t). The task to evaluate $\langle i | H_{\text{EF}}^k F | j \rangle$ can be considered as how to connect the $|j\rangle$ state (a single down-spin at site-j) to the $\langle i|$ state (a single down-spin at site-i) by applying the operator F followed by a sequence of H_{EF}. The net effect is to move the down-spin from site-j to site-i on a background of all up-spins.
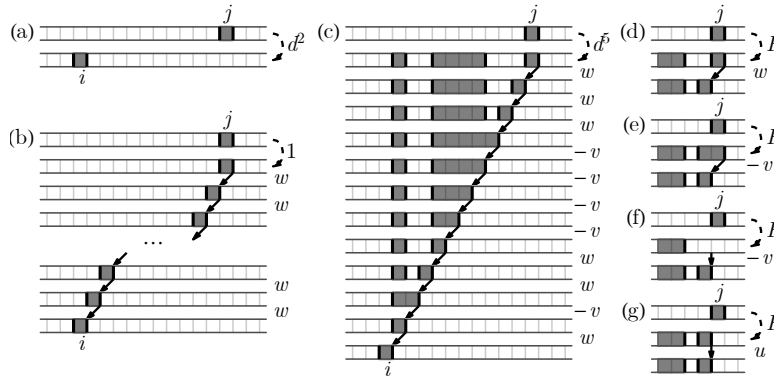
Figure B.3. Examples of entanglement region dynamics in calculating $\langle i | H_{\text{EF}}^k F | j \rangle$.

To warm up, let us start with the 0th order term $\langle i | F | j \rangle$ (i.e. $k = 0$). From Eq. (B.35), we can see that F is a non-local operator, which sprinkles spin flips with a coefficient d. It can be expanded as polynomials of X_i operators. To connect $|j\rangle$ and $|i\rangle$ states, F needs to remove the down-spin at site-j and create the down-spin at site-i, as shown in Fig. B.3(a). This corresponds to spin-flip operations at both sites, described by $X_i X_j$. The coefficient of the $X_i X_j$ term in the expansion of F is d^2 (each X operator contributes a factor d). So we have

$$\langle i | F | j \rangle = d^2.$$

(B.37)

Suppose the sites i and j are spatially separated by the distance $x = |i - j|$, it turns out that the next leading contribution is at the order of $k = x$ as $\langle i | H_{\mathrm{EF}}^x F | j \rangle$, because all the lower order terms $\langle i | H_{\mathrm{EF}}^k F | j \rangle = 0$ vanish for $0 < k < x$. This has to do with the specific algebraic relation between H_{EF} and F. As elaborated in Ref. [128], the entanglement feature Hamiltonian must satisfy the following defining properties: $H_{\mathrm{EF}} W_{\mathbb{1}} = W_{\mathbb{1}} H_{\mathrm{EF}}^{\mathsf{T}}$ and $H_{\mathrm{EF}} P = P H_{\mathrm{EF}}$. Based on the definition of F in Eq. (B.35), we have $H_{\mathrm{EF}} F = F H_{\mathrm{EF}}^{\mathsf{T}}$, therefore

$$\langle i | H_{\mathrm{EF}}^k F | j \rangle = \langle i | F H_{\mathrm{EF}}^{\mathsf{T}k} | j \rangle. \tag{B.38}$$

Since each term in H_{EF} carries a projection operator $\frac{1 - Z_i Z_j}{2}$, Eq. (B.38) implies that the left-most H_{EF} must act on an entanglement cut in the $\langle i |$ state and the right-most H_{EF} must act on an entanglement cut in the $| j \rangle$ state. Similar restrictions applies to all the intermediate actions of H_{EF}. On one hand, H_{EF} must act on the entanglement cut. On the other hand, as a local operator, each application of H_{EF} can only move/manipulate the entanglement cut locally. To connect $| i \rangle$ and $| j \rangle$ states, whose entanglement cuts are separated by at least the distance of x, the most efficient strategy is to "ride on the cut", i.e. the successive application of H_{EF} will have to keep pushing the entanglement cut from j to i and always acts on the "front cut", so as to consume the least number of steps and to make the leading order contribution in the OTOC.

As the F operator sprinkles spin-flips to the initial state $| j \rangle$, there will be multiple entanglement cuts in the resulting state $F | j \rangle$ in general. The subsequent action of H_{EF}^k will have to clear up these entanglement cuts and bring the state to $| i \rangle$. Since H_{EF} is a sum of local operators, it can only move/manipulate the entanglement cut locally. All the allowed processes are listed in Fig. 3.5. Crucially, the pair annihilation process is prohibited, meaning that the only way to reduce the number of entanglement cuts is the triple fusion process, which require us to first bring a pair of entanglement cuts close to the third one. Therefore the most efficient way (taking the least power of H_{EF}) to take $F | j \rangle$ to $| i \rangle$ is to sweep the right-most entanglement cut from site-j to site-i (assuming $j > i$). For example, as illustrated in Fig. B.3(b), suppose F

197

does not act on $|j\rangle$ (which happens with weight 1), the subsequent H_{EF} has to move the pair of entanglement cuts from site-j to site-i step by step, which amounts to a sequence of w-moves. Any additional entanglement regions between sites i and j can be eliminated by the sweeping protocol, as illustrated in Fig. B.3(c). The rule is that when the right-most entanglement cut is adjacent to another cut, they move together to the left as a pair via the pair hopping process (w-move). Otherwise, the right-most entanglement cut will move to the left by itself (v-move). The cut hopping process will shrink the current entanglement region bounded by the moving entanglement cut, until the right entanglement cut meets its left partner and becomes a pair again. However, if F generates additional entanglement regions outside the range of i to j, one will have to take additional steps to eliminate those entanglement regions, which will only contributes to higher order expansions of the OTOC. So we will not consider those cases for now, as we are interested in the leading order contribution.

Given the above protocol, we can see from Fig. B.3(c), each v-move corresponds to a spin-flip introduced by F, so the $(-v)$ amplitude will always accompanied by a factor d (coming from flipping a spin with operator F). The remaining steps will be implemented by w-move with weight w. If we sum over all possibilities, it seems that the inner product $\langle i|H_{\mathrm{EF}}^x F|j\rangle$ should be given by

$$\langle i|H_{\mathrm{EF}}^x F|j\rangle \overset{?}{=} (w - vd)^x. \tag{B.39}$$

This answer is almost correct except for a small caveat at the initial step. The initial move of the right-most entanglement cut can be caused by both H_{EF} and F. Fig. B.3(d,e) show the H_{EF} driven initial moves (where F does nothing to the spin at site-j). Fig. B.3(d,e) show the F driven initial move, where a spin-flip is acted on site-j to move the right-most entanglement cut. For subsequence moves, Fig. B.3(d,e) will not be available, because F operator can only be applied once at the very beginning. Due to the additional contribution from Fig. B.3(d,e) in the initial step, the weight associate to the initial step is modified from $(w - vd)$ to $(w - 2vd + ud^2)$.

Therefore the correct answer is given by

$$\langle i|H_{\text{EF}}^x F|j\rangle = (w - 2vd + ud^2)(w - vd)^{x-1}. \qquad (\text{B.40})$$

According to setting of the parameters u, v, w in Eq. (B.34), we have $w - vd = -g\cosh\beta$ and $w - 2vd + ud^2 = -g\cosh\beta(1 - d^2)$, hence

$$\langle i|H_{\text{EF}}^x F|j\rangle = -(d^2 - 1)(-g\cosh\beta)^x. \qquad (\text{B.41})$$

We can further evaluate the next order term $\langle i|H_{\text{EF}}^{x+1} F|j\rangle$ following the similar strategy. The result is

$$\langle i|H_{\text{EF}}^{x+1} F|j\rangle = (xd^2(1 + \tanh^2\beta) + 2(d^2 - x - 1)\tanh\beta)(-g\cosh\beta)^{x+1}. \qquad (\text{B.42})$$

Substitute the results Eq. (B.37), Eq. (B.41), Eq. (B.42) into Eq. (B.36), we obtain the OTOC to the leading orders in time

$$
\begin{aligned}
\text{OTOC}(x,t) = 1 - (1 - d^{-2})\frac{(tg\cosh\beta)^x}{x!} \\
+ (x(1 + \tanh^2\beta) + 2(1 - (x+1)d^{-2})\tanh\beta)\frac{(tg\cosh\beta)^{x+1}}{(x+1)!} + \mathcal{O}(t^{x+2}). \quad (\text{B.43})
\end{aligned}
$$

199

Appendix C

Shallow Circuit Classical Shadows

C.1 MPS-representation of a finite-depth stabilizer states

Here, we outline an algorithm for constructing the MPS representation of a stabilizer state, given the generators g_i of the corresponding stabilizer group $G = \langle g_1 \cdots g_n \rangle$. In the classical shadow protocol, we start with a trivial product state $|b\rangle = \otimes_{i=1}^n |b_i\rangle$ which represents a measurement outcome in the computational basis, and evolve by a two-local Clifford circuit. The initial stabilizers therefore can transform into more complicated Pauli group elements.

$$(-1)^{b_i} Z_i \to g_i \tag{C.1}$$

For a Clifford circuit of depth $L > 0$, the size of g_i is at most $2L$. The state ρ is then a sum over stabilizer group elements. Similarly, we can also represent the state as an MPS in the Pauli-basis, described by coefficients $c_{\vec{s}}$ where $\vec{s} \in \{0 \cdots 3\}^n$ denotes a Pauli string.

$$\rho = \frac{1}{2^n} \sum_{g \in G} g = \frac{1}{2^n} \sum_{\vec{b} \in \Omega^n} \prod_{i=1}^n g_i^{b_i} = \frac{1}{2^n} \sum_{\vec{s}} c_{\vec{s}} P^{\vec{s}} \tag{C.2}$$

For example, the GHZ state on two sites has two generators, $g_1 = ZZ, g_2 = XX$, and four nonzero components in the Pauli basis, corresponding to II, XX, YY, ZZ. Therefore, there are four nonzero components for c, namely $c_{00} = c_{11} = c_{33} = -c_{22} = 1/4$, and all other components are zero.

We describe a procedure to determine $c_{\vec{s}}$ as an MPS, given a set of generators $g_1 \cdots g_n$ and corresponding phases $(-1)^{s_1} \cdots (-1)^{s_n}$. The idea is that the local MPS tensor at site x should transmit through its virtual indices which of the stabilizers that have support on x are included in the product (i.e. have $b_i = 1$). First, we need to determine a contiguous set of sites that contain the support of each generator. Let us call this the *extent* of an operator. Given the extent of each generator, let $P_i = \mathbb{I} + (-1)^{s_i} g_i$ be an projection operator (or state in the Pauli basis) defined on the extent. Schematically, the local MPS tensor $[P_i^j]$ at site j is

$$[P_i^j] = \begin{cases} (\mathbb{I}, (-1)^{s_i} g_i^j) & j = 1 \\[2mm] \begin{pmatrix} \mathbb{I} & 0 \\ 0 & g_i^j \end{pmatrix} & 1 < j < k \\[2mm] (\mathbb{I}, g_i^j)^T & j = k \end{cases} \tag{C.3}$$

where k is the size of the extent, and g_i^j is the Pauli operator at site j in g_i. Given the MPS P_i, we only need one more ingredient to determine the MPS representation of the stabilizer state: Pauli algebra fusion tensor $V_{a,b}^c$. This tensor implements the product of Pauli operators, and evaluates to the correct phase associated to the product e.g.

$$V_{a,0}^c = \delta_{a,c}, V_{1,2}^3 = i = -V_{2,1}^3, \text{etc.} \tag{C.4}$$

Then, to determine the MPS tensor for the stabilizer state at site j, first collect all the generators whose extent is on j. Then, combine their MPS along the physical dimension using the fusion tensor. If there is more than two generators with support at j, then we can use multiple fusion tensors.

Appendix D

Classical Shadows and the Measurement-Induced Phase Transition

D.0.1 Prior and Posterior Ensembles of Classical Snapshots

Assume the Krause operator (projection operator) $K_{\boldsymbol{b}|\mathscr{C}}$ describes a projective measurement implemented by a quantum circuit $\mathscr{C}$ (with specific gate and observable choices) and resulted in the measurement outcomes $\boldsymbol{b}$, such that the probability to observe $\boldsymbol{b}$ on a state ρ is given by

$$p(\boldsymbol{b}|\rho,\mathscr{C}) = \text{Tr}(K_{\boldsymbol{b}|\mathscr{C}}\rho K^{\dagger}_{\boldsymbol{b}|\mathscr{C}}). \tag{D.1}$$

If we have no knowledge about the state ρ, we should assume $\rho = \mathbb{1}/(\text{Tr}\,\mathbb{1})$ to be maximally mixed, where $\mathbb{1}$ stands for the identity operator in the Hilbert space and $\text{Tr}\,\mathbb{1}$ is effectively the Hilbert space dimension. In this limit, Eq. (D.1) reduces to

$$p(\boldsymbol{b}|\mathscr{C}) := p(\boldsymbol{b}|\rho = \tfrac{\mathbb{1}}{\text{Tr}\,\mathbb{1}},\mathscr{C}) = \frac{1}{\text{Tr}\,\mathbb{1}}\,\text{Tr}(K_{\boldsymbol{b}|\mathscr{C}}K^{\dagger}_{\boldsymbol{b}|\mathscr{C}}). \tag{D.2}$$

The operator $K^{\dagger}_{\boldsymbol{b}|\mathscr{C}}K_{\boldsymbol{b}|\mathscr{C}}$ is a Hermitian and positive semi-definite operator, which motivates us to further normalize it to make it a state (a density matrix)

$$\sigma_{\boldsymbol{b}|C} = \frac{K^{\dagger}_{\boldsymbol{b}|\mathscr{C}}K_{\boldsymbol{b}|\mathscr{C}}}{\text{Tr}(K^{\dagger}_{\boldsymbol{b}|\mathscr{C}}K_{\boldsymbol{b}|\mathscr{C}})}. \tag{D.3}$$

We will call $\sigma_{\boldsymbol{b}|C}$ a *classical snapshot* state. It has an important property that

$$\text{Tr}(\sigma_{\boldsymbol{b}|C}\rho) = \frac{\text{Tr}(K^\dagger_{\boldsymbol{b}|\mathscr{C}}K_{\boldsymbol{b}|\mathscr{C}}\rho)}{\text{Tr}(K^\dagger_{\boldsymbol{b}|\mathscr{C}}K_{\boldsymbol{b}|\mathscr{C}})} = \frac{1}{\text{Tr}\,\mathbb{1}}\frac{p(\boldsymbol{b}|\rho,\mathscr{C})}{p(\boldsymbol{b}|\mathscr{C})}. \tag{D.4}$$

Assume that the measurement circuit $\mathscr{C}$ is drawn from some random ensemble with probability $p(\mathscr{C})$, we can define two random state ensembles for the classical snapshots:

- the *prior* snapshot ensemble (with no knowledge about ρ)

$$\mathscr{E}_\sigma = \{\sigma_{\boldsymbol{b}|\mathscr{C}}|\,(\boldsymbol{b},\mathscr{C})\sim p(\boldsymbol{b},\mathscr{C}) := p(\boldsymbol{b}|\mathscr{C})p(\mathscr{C})\}, \tag{D.5}$$

- the *posterior* snapshot ensemble (given the knowledge about ρ)

$$\mathscr{E}_{\sigma|\rho} = \{\sigma_{\boldsymbol{b}|\mathscr{C}}|\,(\boldsymbol{b},\mathscr{C})\sim p(\boldsymbol{b},\mathscr{C}|\rho) := p(\boldsymbol{b}|\rho,\mathscr{C})p(\mathscr{C})\}. \tag{D.6}$$

This means the ensemble averages are defined as

$$\begin{aligned}
\mathop{\mathbb{E}}_{\sigma\in\mathscr{E}_\sigma} f(\sigma) &:= \sum_{\boldsymbol{b},\mathscr{C}} f(\sigma_{\boldsymbol{b}|\mathscr{C}})p(\boldsymbol{b}|\mathscr{C})p(\mathscr{C}), \\
\mathop{\mathbb{E}}_{\sigma\in\mathscr{E}_{\sigma|\rho}} f(\sigma) &:= \sum_{\boldsymbol{b},\mathscr{C}} f(\sigma_{\boldsymbol{b}|\mathscr{C}})p(\boldsymbol{b}|\rho,\mathscr{C})p(\mathscr{C}),
\end{aligned} \tag{D.7}$$

where $f(\sigma)$ stands for any function of σ, $p(\boldsymbol{b}|\mathscr{C})$ and $p(\boldsymbol{b}|\rho,\mathscr{C})$ are given by Eq. (D.2) and Eq. (D.1) respectively. Then Eq. (D.4) implies that the posterior ensemble average can be expressed as a prior ensemble average as

$$\mathop{\mathbb{E}}_{\sigma\in\mathscr{E}_{\sigma|\rho}} f(\sigma) = (\text{Tr}\,\mathbb{1})\mathop{\mathbb{E}}_{\sigma\in\mathscr{E}_\sigma} f(\sigma)\,\text{Tr}(\sigma\rho). \tag{D.8}$$

D.0.2 Review of Classical Shadow Tomography

Classical shadow tomography is an efficient approach to extracting information about an unknown quantum state ρ by repeated randomized measurements. The idea of randomized measurement is to sample the measurement circuit $\mathscr{C}$ (both its structure and its gate choices) from a probability distribution $p(\mathscr{C})$, perform the measurement on the state ρ and collect the measurement outcomes $\boldsymbol{b}$. The randomized measurement effectively converts the quantum state ρ into a collection of classical snapshots in the posterior ensemble $\mathscr{E}_{\sigma|\rho}$.

The data acquisition procedure can be formulated as a quantum channel $\mathscr{M}$, called the *measurement channel*, which maps the initial quantum state ρ to the expectation of classical snapshot states over the posterior snapshot ensemble

$$\mathscr{M}(\rho) := \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_{\sigma|\rho}} \sigma = (\mathrm{Tr}\,\mathbb{1}) \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_\sigma} \sigma \, \mathrm{Tr}(\sigma\rho). \tag{D.9}$$

Here we have used Eq. (D.8). Suppose the randomized measurement scheme is topographically complete, the measurement channel $\mathscr{M}$ will be invertible. Its inverse is called the *reconstruction map*, denoted as $\mathscr{M}^{-1}$ (albeit $\mathscr{M}^{-1}$ might not be a physical channel), such that the quantum state ρ can be reconstructed from the classical snapshots by $\rho = \mathbb{E}_{\sigma \in \mathscr{E}_{\sigma|\rho}} \mathscr{M}^{-1}(\sigma)$. This also provides the means to predict the expectation value of any observable O on the state ρ as

$$\langle O \rangle := \mathrm{Tr}(O\rho) = \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_{\sigma|\rho}} \mathrm{Tr}(\mathscr{M}^{-1}(O)\sigma). \tag{D.10}$$

Note that $\mathrm{Tr}(O\mathscr{M}^{-1}(\sigma)) = \mathrm{Tr}(\mathscr{M}^{-1}(O)\sigma)$ due to the self-adjoint property of the measurement channel $\mathscr{M}$ as well as the reconstruction map $\mathscr{M}^{-1}$.

In practice, the expectation $\mathbb{E}_{\sigma \in \mathscr{E}_{\sigma|\rho}}$ is often estimated by the median-of-means over a finite number of classical snapshots collected from experiments. Due to the statistical fluctuation of finite samples, the estimation value of an observable O will fluctuate around its true expectation value $\langle O \rangle$ with a typical variance that scales with the sample number M as $\mathrm{var}\, O \sim \|O\|_{\mathscr{E}_\sigma}^2 / M$

following the law of large numbers. The coefficient $\|O\|_{\mathscr{E}_\sigma}^2$ is the *locally-scrambled shadow norm*, defined as

$$\|O\|_{\mathscr{E}_\sigma}^2 := \underset{\sigma \in \mathscr{E}_\sigma}{\mathbb{E}}\, (\mathrm{Tr}\,\mathscr{M}^{-1}(O)\sigma)^2 = \frac{\mathrm{Tr}(O\mathscr{M}^{-1}(O))}{\mathrm{Tr}\,\mathbb{1}}. \tag{D.11}$$

By definition, the identity operator $O = \mathbb{1}$ always have unit shadow norm, i.e., $\|\mathbb{1}\|_{\mathscr{E}_\sigma}^2 \equiv 1$. For traceless observable O, the locally-scrambled shadow norm $\|O\|_{\mathscr{E}_\sigma}^2$ quantifies the number M of samples needed to control the estimation variances $\mathrm{var}\,O \lesssim \delta^2$ below a desired level set by a small δ, as $M \sim \|O\|_{\mathscr{E}_\sigma}^2/\delta^2$. Therefore, the locally-scrambled shadow norm measures the *sample complexity* for classical shadow tomography to predict the observable O based on the randomized measurement scheme characterized by $\mathscr{E}_\sigma$.

D.0.3   Locally Scrambled Ensembles and Pauli Basis Approach

From the aforementioned general formulation, it is evident that the measurement channel $\mathscr{M}$ and its inverse $\mathscr{M}^{-1}$ hold a central role in classical shadow tomography. Computing these for generic randomized measurement schemes is a complex task, and no polynomially scalable algorithm currently exists. Nevertheless, progress has been made in the context of locally scrambled (or Pauli-twirled) measurements, which are randomized measurements insensitive to local basis choices.

More specifically, let $V \in \mathrm{U}(2)^N$ be a product of single-qubit unitary operators, decomposed as $V = \bigotimes_i V_i$, with $V_i \in \mathrm{U}(2)$ on every qubit i. The product operator V represents independent local basis transformations. Let $\mathscr{U}_V(\rho) := V\rho V^\dagger$ be the unitary channel that implements the unitary transformation V. A random state ensemble $\mathscr{E}_\rho = \{\rho | \rho \sim p(\rho)\}$ is said to be locally scrambled, if and only if $p(\rho) = p(\mathscr{U}_V(\rho))$ for any $V \in \mathrm{U}(2)^N$. A random channel ensemble $\mathscr{E}_\mathscr{K} = \{\mathscr{K} | \mathscr{K} \sim p(\mathscr{K})\}$ is considered locally scrambled, if and only if $p(\mathscr{K}) = p(\mathscr{U}_V \circ \mathscr{K} \circ \mathscr{U}_{V'})$ for any $V, V' \in \mathrm{U}(2)^N$. A randomized measurement scheme is locally scrambled if its corresponding prior snapshot ensemble $\mathscr{E}_\sigma$ (as a random state ensemble) is lo-

cally scrambled. For locally scrambled randomized measurements, the associated $\mathcal{M}$ and $\mathcal{M}^{-1}$ exhibit simple forms in the Pauli basis and can be computed efficiently. This conclusion remains valid when the scrambling condition is relaxed to $V \in C_1^N$, where C_1 denotes the single-qubit Clifford group, which is applicable to the cases of random Clifford gates with random Pauli measurements.

To elucidate this approach, it is more convenient to employ the Choi representation, wherein each quantum operator O is viewed as a super-state $|O\rangle\rangle$:

$$|O\rangle\rangle := \frac{1}{\sqrt{\operatorname{Tr}\mathbb{1}}} \sum_a |a\rangle \otimes O|a\rangle, \tag{D.12}$$

where a labels a complete set of orthonormal basis in the Hilbert space $\mathscr{H}$ (of ket states). All Hermitian operators in $\mathscr{H}$ span an operator space, denoted $\mathscr{L}(\mathscr{H})$, in which inner product between two operators $A, B \in \mathscr{L}(\mathscr{H})$ is defined as

$$\langle\langle A|B\rangle\rangle := \frac{\operatorname{Tr}A^\dagger B}{\operatorname{Tr}\mathbb{1}}. \tag{D.13}$$

Let $P = \prod_i P_i$ represent a (multi-qubit) Pauli operator, constructed as a product of single-qubit Pauli operators, with $P_i \in I, X, Y, Z$ acting on the ith qubit. All Pauli operators form a complete set of orthonormal basis for the Hermitian operator space $\mathscr{L}(\mathscr{H})$, as $\langle\langle P|P'\rangle\rangle = \delta_{PP'}$. Correspondingly, for any quantum channel $\mathscr{K}$ that maps a state ρ to another state $\rho' = \mathscr{K}(\rho)$, there exists a corresponding super-operator $\hat{\mathscr{K}} \in \operatorname{Hom}(\mathscr{L}(\mathscr{H}), \mathscr{L}(\mathscr{H}))$ that maps their Choi representations accordingly, as $|\rho'\rangle\rangle = \hat{\mathscr{K}}|\rho\rangle\rangle$.

Utilizing the Choi representation of quantum channel, any locally scrambled measurement channel $\mathcal{M}$ defined in Eq. (D.9) can be written as a super-operator $\hat{\mathcal{M}}$ that acts on any state ρ as

$$\hat{\mathcal{M}}|\rho\rangle\rangle = (\operatorname{Tr}\mathbb{1})^2 \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_\sigma} |\sigma\rangle\rangle\langle\langle\sigma|\rho\rangle\rangle, \tag{D.14}$$

which indicates that $\hat{\mathcal{M}} = (\mathrm{Tr}\,\mathbb{1})^2 \, \mathbb{E}_{\sigma \in \mathcal{E}_\sigma} |\sigma\rangle\rangle\langle\langle\sigma|$. Expand in Pauli basis, we generally have

$$\hat{\mathcal{M}} = (\mathrm{Tr}\,\mathbb{1})^2 \sum_{P,P'} |P\rangle\rangle \left(\mathop{\mathbb{E}}_{\sigma \in \mathcal{E}_\sigma} \langle\langle P|\sigma\rangle\rangle\langle\langle\sigma|P'\rangle\rangle \right) \langle\langle P'|. \tag{D.15}$$

However, if we assume $\mathcal{E}_\sigma$ to be a locally scrambled state ensemble, the local scrambling property requires

$$\mathop{\mathbb{E}}_{\sigma \in \mathcal{E}_\sigma} \langle\langle P|\sigma\rangle\rangle\langle\langle\sigma|P'\rangle\rangle = \delta_{PP'} \mathop{\mathbb{E}}_{\sigma \in \mathcal{E}_\sigma} \langle\langle P|\sigma\rangle\rangle\langle\langle\sigma|P\rangle\rangle = \delta_{PP'} \mathop{\mathbb{E}}_{\sigma \in \mathcal{E}_\sigma} \left(\frac{\mathrm{Tr}\,P\sigma}{\mathrm{Tr}\,\mathbb{1}} \right)^2. \tag{D.16}$$

Therefore, the measurement channel Eq. (D.15) becomes diagonal in the Pauli basis:

$$\hat{\mathcal{M}} = \sum_P |P\rangle\rangle w_{\mathcal{E}_\sigma}(P) \langle\langle P|, \tag{D.17}$$

where $w_{\mathcal{E}_\sigma}(P)$ denotes the average *Pauli weight* of classical snapshots σ in prior snapshot ensemble $\mathcal{E}_\sigma$, defined as

$$w_{\mathcal{E}_\sigma}(P) := \mathop{\mathbb{E}}_{\sigma \in \mathcal{E}_\sigma} (\mathrm{Tr}\,P\sigma)^2. \tag{D.18}$$

This definition captures an essential statistical feature of the randomized measurement scheme: $w_{\mathcal{E}_\sigma}(P)$ is the probability for the Pauli operator P to appear diagonal in the measurement basis and be directly observed in a single random realization of the measurement protocol.

Since $\hat{\mathcal{M}}$ is diagonal in the Pauli basis, its inverse is simply given by

$$\hat{\mathcal{M}}^{-1} = \sum_P |P\rangle\rangle \frac{1}{w_{\mathcal{E}_\sigma}(P)} \langle\langle P|, \tag{D.19}$$

or, more explicitly, in terms of the reconstruction map on any operator O as

$$\mathcal{M}^{-1}(O) = \sum_P \frac{\mathrm{Tr}(OP)}{w_{\mathcal{E}_\sigma}(P)\,\mathrm{Tr}\,\mathbb{1}} P. \tag{D.20}$$

Plugging Eq. (D.20) into Eq. (D.10) and Eq. (D.11) enables us to estimate the expectation value

$\langle O \rangle$ and study the shadow norm $\|O\|_{\mathscr{E}_\sigma}^2$ for any observable O. In particular, for Pauli observable P, we have

$$\langle P \rangle = \frac{1}{w_{\mathscr{E}_\sigma}(P)} \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_{\sigma|\rho}} \mathrm{Tr}(P\sigma),$$

$$\|P\|_{\mathscr{E}_\sigma}^2 = \frac{1}{w_{\mathscr{E}_\sigma}(P)}. \tag{D.21}$$

This essentially allows us to decode the quantum information in the initial state ρ from the measurement outcomes gathered from a hybrid quantum circuit and to investigate the sample complexity for decoding various observables at different measurement rates.

D.0.4 Evolution of Pauli Weights through Locally Scrambled Channels

The central challenge now is to compute the average Pauli weight $w_{\mathscr{E}_\sigma}(P)$ of the prior snapshot ensemble $\mathscr{E}_\sigma$. In general, this remains a difficult problem. However, suppose the state ensemble $\mathscr{E}_\sigma$ is constructed by applying locally scrambled elementary quantum channels to locally scrambled simple initial states. In that case, its Pauli weight can be calculated using Markov dynamics, which is a more tractable approach.

To derive the dynamics of Pauli weights under locally scrambled quantum dynamics, we first introduce a set of *region basis* states $|A\rangle\!\rangle$ in the doubled Hermitian operator space $\mathscr{L}(\mathscr{H})^2$, defined by

$$|A\rangle\!\rangle = \frac{1}{\sqrt{|\mathscr{P}_A|}} \sum_{P \in \mathscr{P}_A} |P\rangle\!\rangle \otimes |P\rangle\!\rangle, \tag{D.22}$$

where A denotes a subset of qubits and $\mathscr{P}_A = \{P \,|\, \mathrm{supp}\, P = A\}$ denotes all Pauli operators supported exactly in region A. $|\mathscr{P}_A| = 3^{|A|}$ is the cardinality of $\mathscr{P}_A$ and $|A|$ denotes the number of qubits in A. The significance of $|A\rangle\!\rangle$ lies in its invariance under doubled local basis transformation, i.e., $\hat{\mathscr{U}}_V \otimes \hat{\mathscr{U}}_V |A\rangle\!\rangle = |A\rangle\!\rangle$ for any $V \in \mathrm{U}(2)^N$. In fact, all states $|A\rangle\!\rangle$ form a complete set of orthonormal basis that spans the invariant subspace of $\mathscr{L}(\mathscr{H})^2$ under the doubled local scrambling $\hat{\mathscr{U}}_V \otimes \hat{\mathscr{U}}_V$.

Given a locally scrambled state ensemble $\mathcal{E}_\rho$ and a locally scrambled channel ensemble $\mathcal{E}_{\mathcal{K}}$, by definition, for any local basis transformation $V = \bigotimes_i V_i \in \mathrm{U}(2)^N$ (or its corresponding unitary channel $\mathcal{U}_V(\rho) = V\rho V^\dagger$), we have the following symmetry requirements

$$\hat{\mathcal{U}}_V^{\otimes 2} \underset{\rho \in \mathcal{E}_\rho}{\mathbb{E}} |\rho\rangle\!\rangle^{\otimes 2} = \underset{\rho \in \mathcal{E}_\rho}{\mathbb{E}} |\rho\rangle\!\rangle^{\otimes 2},$$

$$\hat{\mathcal{U}}_V^{\otimes 2}\left(\underset{\mathcal{K} \in \mathcal{E}_{\mathcal{K}}}{\mathbb{E}} \hat{\mathcal{K}}^{\otimes 2}\right)\hat{\mathcal{U}}_{V^\dagger}^{\otimes 2} = \underset{\mathcal{K} \in \mathcal{E}_{\mathcal{K}}}{\mathbb{E}} \hat{\mathcal{K}}^{\otimes 2}. \tag{D.23}$$

These manifest the locally scrambled properties on the 2nd-moment level. Given these symmetry requirements, the 2nd-moment locally scrambled random states (or channels) can and only need to be represented in the symmetric subspace spanned by the region basis as

$$\underset{\rho \in \mathcal{E}_\rho}{\mathbb{E}} |\rho\rangle\!\rangle^{\otimes 2} = \frac{1}{(\mathrm{Tr}\,\mathbb{1})^2} \sum_A \bar{w}_{\mathcal{E}_\rho}(A)|A\rangle\!\rangle,$$

$$\underset{\mathcal{K} \in \mathcal{E}_{\mathcal{K}}}{\mathbb{E}} \hat{\mathcal{K}}^{\otimes 2} = \sum_{A,A'} |A\rangle\!\rangle \bar{w}_{\mathcal{E}_{\mathcal{K}}}(A,A')\langle\!\langle A'|, \tag{D.24}$$

where the linear combination coefficients $\bar{w}_\rho(A)$ and $\bar{w}_{\mathcal{K}}(A,A')$ are defined as

$$\bar{w}_{\mathcal{E}_\rho}(A) := (\mathrm{Tr}\,\mathbb{1})^2 \underset{\rho \in \mathcal{E}_\rho}{\mathbb{E}} \langle\!\langle A||\rho\rangle\!\rangle^{\otimes 2},$$

$$\bar{w}_{\mathcal{E}_{\mathcal{K}}}(A,A') := \underset{\mathcal{K} \in \mathcal{E}_{\mathcal{K}}}{\mathbb{E}} \langle\!\langle A|\hat{\mathcal{K}}^{\otimes 2}|A'\rangle\!\rangle, \tag{D.25}$$

which will be called the regional Pauli weights of the locally scrambled ensembles.

Locally scrambled quantum dynamics involve passing locally scrambled random states $\rho \in \mathcal{E}_\rho$ through locally scrambled random channels $\mathcal{K} \in \mathcal{E}_{\mathcal{K}}$ such that the resulting states $\mathcal{K}(\rho)$ still form a locally scrambled random state ensemble $\mathcal{E}_{\mathcal{K}(\rho)}$ defined as:

$$\mathcal{E}_\rho \to \mathcal{E}_{\mathcal{K}(\rho)} := \{\mathcal{K}(\rho)|\rho \in \mathcal{E}_\rho, \mathcal{K} \in \mathcal{E}_{\mathcal{K}}\}. \tag{D.26}$$

Under such evolution, the 2nd-moment evolves as

$$\underset{\mathcal{K}(\rho)\in\mathscr{E}_{\mathcal{K}(\rho)}}{\mathbb{E}} |\mathcal{K}(\rho)\rangle\!\rangle^{\otimes 2} = \underset{\mathcal{K}\in\mathscr{E}_{\mathcal{K}}}{\mathbb{E}} \hat{\mathcal{K}}^{\otimes 2} \underset{\rho\in\mathscr{E}_{\rho}}{\mathbb{E}} |\rho\rangle\!\rangle^{\otimes 2}. \tag{D.27}$$

By decomposing these 2nd-moment state and channel expectations in the region basis according to Eq. (D.24), we obtain the following evolution equation for the regional Pauli weights:

$$\bar{w}_{\mathscr{E}_{\mathcal{K}(\rho)}}(A) = \sum_{A'} \bar{w}_{\mathscr{E}_{\mathcal{K}}}(A,A')\bar{w}_{\mathscr{E}_{\rho}}(A'). \tag{D.28}$$

Suppose a locally scrambled state ensemble is obtained from random quantum dynamics comprised of locally scrambled simple channels, with computable regional Pauli weights. In that case, we can use Eq. (D.28) to trace the evolution of the regional Pauli weights of the state ensemble step-by-step, eventually inferring the regional Pauli weights of the final state ensemble. This approach enables us to systematically calculate the Pauli weights of the prior snapshot ensemble $\mathscr{E}_{\sigma}$, which is the key for classical shadow tomography.

Finally, the regional Pauli weights and the ordinary Pauli weights are connected by the following relations, given P and P' are Pauli operators supported on regions A and A', respectively,

$$\begin{aligned}
w_{\mathscr{E}_{\rho}}(P) &= \frac{\bar{w}_{\mathscr{E}_{\rho}}(A)}{\sqrt{3}^{|A|}} = \underset{\rho\in\mathscr{E}_{\rho}}{\mathbb{E}} (\text{Tr}\, P\rho)^2, \\
w_{\mathscr{E}_{\mathcal{K}}}(P,P') &= \frac{\bar{w}_{\mathscr{E}_{\mathcal{K}}}(A,A')}{\sqrt{3}^{|A|+|A'|}} = \underset{\mathcal{K}\in\mathscr{E}_{\mathcal{K}}}{\mathbb{E}} \left(\frac{\text{Tr}(P\mathcal{K}(P'))}{\text{Tr}\,\mathbb{1}}\right)^2.
\end{aligned} \tag{D.29}$$

Switching back to the Pauli basis, Eq. (D.28) reduces to a similar form

$$w_{\mathscr{E}_{\mathcal{K}(\rho)}}(P) = \sum_{P'} w_{\mathscr{E}_{\mathcal{K}}}(P,P')w_{\mathscr{E}_{\rho}}(P'). \tag{D.30}$$

Based on these formulae, we can compute the Pauli weights for a few locally scrambled states and channels, which will be useful for analyzing the hybrid quantum circuit classical shadow

tomography.

- Locally scrambled state ensembles $\mathcal{E}_\rho$

  - Pure random product states $\mathcal{E}_\rho = \{\bigotimes_i |\psi_i\rangle\langle\psi_i| \mid |\psi_i\rangle \in \mathcal{H}_i\}$

$$w_{\mathcal{E}_\rho}(P) = \frac{1}{3^{|\operatorname{supp}P|}}. \tag{D.31}$$

  - Pure Page state over N qubits $\mathcal{E}_\rho = \{|\psi\rangle\langle\psi| \mid |\psi\rangle \in \mathcal{H} = \bigotimes_i \mathcal{H}_i\}$

$$w_{\mathcal{E}_\rho}(P) = \begin{cases} 1 & P = \mathbb{1}, \\ \frac{1}{2^N+1} & \text{otherwise.} \end{cases} \tag{D.32}$$

  - Maximally mixed state $\mathcal{E}_\rho = \{\mathbb{1}/(\operatorname{Tr}\mathbb{1})\}$

$$w_{\mathcal{E}_\rho}(P) = \delta_{P,\mathbb{1}} = \begin{cases} 1 & P = \mathbb{1}, \\ 0 & \text{otherwise.} \end{cases} \tag{D.33}$$

- Locally scrambled channel ensembles $\mathcal{E}_\mathcal{K}$

  - Local scrambling channel $\mathcal{E}_\mathcal{K} = \{\mathcal{U}_V \mid \mathcal{U}_V(\rho) = V\rho V^\dagger, V \in \mathrm{U}(2)^N\}$

$$w_{\mathcal{E}_\mathcal{K}}(P,P') = \frac{\delta_{\operatorname{supp}P,\operatorname{supp}P'}}{3^{|\operatorname{supp}P|}}. \tag{D.34}$$

  - Global scrambling channel among N qubits:

$$w_{\mathcal{E}_\mathcal{K}}(P,P') = \delta_{P,\mathbb{1}}\delta_{P',\mathbb{1}} + \frac{(1-\delta_{P,\mathbb{1}})(1-\delta_{P',\mathbb{1}})}{4^N - 1}, \tag{D.35}$$

  where $\mathcal{E}_\mathcal{K} = \{\mathcal{U}_U \mid \mathcal{U}_U(\rho) = U\rho U^\dagger, U \in \mathrm{U}(2^N)\}$.

– Local random projective measurement:

$$w_{\mathscr{E}_{\mathscr{K}}}(P,P') = \frac{1}{3^{|\operatorname{supp}P| + |\operatorname{supp}P'|}}, \tag{D.36}$$

where $\mathscr{E}_{\mathscr{K}} = \{\mathscr{M}|\mathscr{M}(\rho_i) = |\psi_i\rangle\langle\phi_i|\rho_i|\phi_i\rangle\langle\psi_i|, |\phi_i\rangle, |\psi_i\rangle \in \mathscr{H}_i\}$.

D.0.5 Application to the Prior Snapshot Ensemble

According to Eq. (D.3), the classical snapshot state σ takes the form of

$$\sigma = \frac{K^\dagger K}{\operatorname{Tr}(K^\dagger K)}. \tag{D.37}$$

The state can be interpreted as mapping the maximal mixed state $\rho_{\mathbb{1}} := \mathbb{1}/(\operatorname{Tr}\mathbb{1})$ by a quantum channel $\mathscr{K}$, followed by a normalization,

$$\sigma = \frac{\mathscr{K}(\rho_{\mathbb{1}})}{\operatorname{Tr}\mathscr{K}(\rho_{\mathbb{1}})}, \tag{D.38}$$

where the quantum channel $\mathscr{K}$ is defined by the Krause operator as $\mathscr{K}(\rho) = K^\dagger\rho K$. The random state ensemble of σ can be considered as derived from the random channel ensemble of $\mathscr{K}$,

$$\mathscr{E}_\sigma = \left\{\frac{\mathscr{K}(\rho_{\mathbb{1}})}{\operatorname{Tr}\mathscr{K}(\rho_{\mathbb{1}})}\middle| \mathscr{K} \in \mathscr{E}_{\mathscr{K}}\right\} \tag{D.39}$$

.

The maximally mixed state itself forms a locally scrambled state ensemble $\mathscr{E}_{\rho_{\mathbb{1}}} = \{\rho_{\mathbb{1}}\}$, whose Pauli weight $w_{\mathscr{E}_{\rho_{\mathbb{1}}}}$ is given by Eq. (D.33). Suppose the random channel ensemble $\mathscr{E}_{\mathscr{K}}$ is also locally scrambled, Eq. (D.30) can be applied to compute the Pauli weight of the composed state ensemble $\mathscr{E}_{\mathscr{K}(\rho_{\mathbb{1}})}$,

$$w_{\mathscr{E}_{\mathscr{K}(\rho_{\mathbb{1}})}}(P) = \sum_{P'}w_{\mathscr{E}_{\mathscr{K}}}(P,P')w_{\mathscr{E}_{\rho_{\mathbb{1}}}}(P'). \tag{D.40}$$

By Eq. (D.29), $w_{\mathscr{E}_{\mathscr{K}(\rho_1)}}$ is defined as

$$w_{\mathscr{E}_{\mathscr{K}(\rho_1)}} = \underset{\mathscr{K} \in \mathscr{E}_{\mathscr{K}}}{\mathbb{E}} (\operatorname{Tr} P \mathscr{K}(\rho_1))^2. \tag{D.41}$$

However, what we wish to calculate is the Pauli weight of the classical snapshot ensemble, defined as

$$w_{\mathscr{E}_{\sigma}} = \underset{\mathscr{K} \in \mathscr{E}_{\mathscr{K}}}{\mathbb{E}} \left(\operatorname{Tr} \left(P \frac{\mathscr{K}(\rho_1)}{\operatorname{Tr} \mathscr{K}(\rho_1)} \right) \right)^2 = \underset{\mathscr{K} \in \mathscr{E}_{\mathscr{K}}}{\mathbb{E}} \frac{(\operatorname{Tr} P \mathscr{K}(\rho_1))^2}{(\operatorname{Tr} \mathscr{K}(\rho_1))^2}. \tag{D.42}$$

This ensemble average is difficult to calculate due to the potential correlation between its numerator and denominator. A common strategy is to approximate the average of ratios by the ratio of averages. Using Eq. (D.41), we have

$$w_{\mathscr{E}_{\sigma}}(P) \simeq \frac{\mathbb{E}_{\mathscr{K} \in \mathscr{E}_{\mathscr{K}}}(\operatorname{Tr} P \mathscr{K}(\rho_1))^2}{\mathbb{E}_{\mathscr{K} \in \mathscr{E}_{\mathscr{K}}}(\operatorname{Tr} \mathscr{K}(\rho_1))^2} = \frac{w_{\mathscr{E}_{\mathscr{K}(\rho_1)}}(P)}{w_{\mathscr{E}_{\mathscr{K}(\rho_1)}}(\mathbb{1})}. \tag{D.43}$$

Such that the Pauli weight of the classical snapshot ensemble can be approximately estimated as the ratio of two Pauli weights that we can compute by the operator dynamics Eq. (D.40). The approximation is expected to be accurate away from the transition in the purity.

D.1 Quantum Statistical Mechanical Picture

D.1.1 Pauli Weight and Entanglement Feature

Given the prior snapshot ensemble $\mathscr{E}_{\sigma}$, we aim to compute the Pauli weight $w_{\mathscr{E}_{\sigma}}(P)$, which plays a central role in classical shadow tomography reconstruction and shadow norm estimation. The Pauli weight is defined as

$$w_{\mathscr{E}_{\sigma}}(P) := \underset{\sigma \in \mathscr{E}_{\sigma}}{\mathbb{E}} (\operatorname{Tr} P \sigma)^2, \tag{D.44}$$

for any Pauli operator $P = \prod_i P_i$ and $P_i \in \{I, X, Y, Z\}$ in qubit systems. If $\mathscr{E}_\sigma$ is a locally scrambled ensemble, its 2nd moment if fully captured by the (2nd Rényi) entanglement feature, defined as

$$W_{\mathscr{E}_\sigma}(A) := \mathop{\mathbb{E}}_{\sigma \in \mathscr{E}_\sigma} \mathrm{Tr}(\sigma^{\otimes 2} \mathsf{SWAP}_A), \tag{D.45}$$

where A denotes a subset of totally N qubits in the system, which can also be encoded as a bit string $A \in \{0, 1\}^{\times N}$ such that the qubit marked by 1 belongs to the subset. SWAP_A denotes the swap operator in the region A between the two replicas of σ. The Pauli weight and the entanglement feature are related by a linear transformation

$$w_{\mathscr{E}_\sigma}(P)|_{\mathrm{supp}\, P = A} = \left(-\frac{1}{3}\right)^{|A|} \sum_{B \subseteq A} (-2)^{|B|} W_{\mathscr{E}_\sigma}(B), \tag{D.46}$$

where $|A|$ denotes the size (cardinality) of the set A.
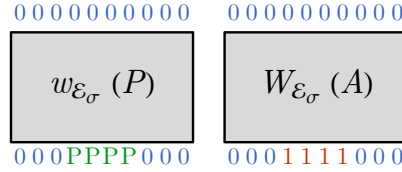
Figure D.1. Statistical mechanical picture of the Pauli weight $w_{\mathscr{E}_\sigma}$ and the entanglement feature $W_{\mathscr{E}_\sigma}$.

This relation can be simply expressed using the entanglement feature state, which is a fictitious quantum state that encodes the entanglement feature over all possible subsystems A. Let $|A\rangle$ be a set of orthonormal bit-string basis (i.e., assuming $\langle A|B \rangle = \delta_{AB}$), we can define the entanglement feature state $|W_{\mathscr{E}_\sigma}\rangle$ for the ensemble $\mathscr{E}_\sigma$ as

$$|W_{\mathscr{E}_\sigma}\rangle = \sum_A W_{\mathscr{E}_\sigma}(A)|A\rangle, \tag{D.47}$$

then both the Pauli weight and the entanglement feature admits simple representation as

$$w_{\mathscr{E}_\sigma}(P) = \langle P | W_{\mathscr{E}_\sigma} \rangle,$$
$$W_{\mathscr{E}_\sigma}(A) = \langle A | W_{\mathscr{E}_\sigma} \rangle,$$

(D.48)

where the state $|P\rangle$ is defined by

$$|P\rangle := \prod_{i \in \text{supp}\, P} \frac{2X_i - I_i}{3} |\mathbf{0}\rangle,$$

(D.49)

with $|\mathbf{0}\rangle := \bigotimes_i |0\rangle_i$ being the all-0 state (the basis state of empty region $A = \emptyset$) and $X = |0\rangle\langle 1| + |1\rangle\langle 0|$, $I = |0\rangle\langle 0| + |1\rangle\langle 1|$ are on-site operators in the entanglement feature Hilbert space. Physically, $|0\rangle_i$ (or $|1\rangle_i$) corresponds to the identity (or swap) boundary condition on the qubit i in defining the 2nd-moment computation. The state $|P\rangle$ defines the boundary condition for Pauli weight, corresponding to imposing the Pauli boundary condition within the Pauli operator support.

D.1.2 Entanglement Feature State in Measurement-Induced Transition

Suppose the prior snapshot state is given by $\sigma = K^\dagger K / \text{Tr}(K^\dagger K)$ with K being the Kraus operator that describes a quantum channel of the hybrid quantum circuit with measurement rate p.

- $p = 0$ limit, $\mathscr{E}_\sigma(p=0)$ only contains a maximally mixed state, whose entanglement feature state is

$$|W_{\mathscr{E}_\sigma(p=0)}\rangle = \frac{1}{4^N} W_{\mathbb{1}} |\mathbf{0}\rangle,$$

(D.50)

where $W_{\mathbb{1}}$ is the entanglement feature operator of the identity quantum channel, given by

$$W_{\mathbb{1}} = 2^N \prod_i (2I_i + X_i).$$

(D.51)

- $p = 1$ limit, $\mathscr{E}_\sigma(p=1)$ is an ensemble of random product state, meaning that the entanglement entropy vanishes for all regions, so $\forall A : W_{\mathscr{E}_\sigma}(A) = 1$, the entanglement feature state can be written as

$$|W_{\mathscr{E}_\sigma(p=1)}\rangle = 2^{N/2}|+\rangle, \tag{D.52}$$

where $|+\rangle := \otimes_i |+\rangle_i$ is the all-plus state with $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ being the eigenstate of $X = +1$. Given that $|+\rangle$ is also an eigenstate of $W_\mathbb{1}$, it does not hurt to rewrite Eq. (D.52) as

$$|W_{\mathscr{E}_\sigma(p=1)}\rangle = \left(\frac{1}{3\sqrt{2}}\right)^N W_\mathbb{1}|+\rangle. \tag{D.53}$$

Given the two limits Eq. (D.51) and Eq. (D.53), we can propose a variational ansatz for the entanglement feature state

$$|W_{\mathscr{E}_\sigma(p)}\rangle \propto W_\mathbb{1}|\Psi_p\rangle, \tag{D.54}$$

where $|\Psi_p\rangle$ should interpolate between the ferromagnetic state $|0\rangle$ and the paramagnetic state $|+\rangle$ from $p = 0$ to $p = 1$. A simple way to realize such an interpolation is to consider $|\Psi_p\rangle$ being the ground state of a transverse field Ising model

$$H = -J\sum_i Z_i Z_{i+1} - h\sum_i X_i, \tag{D.55}$$

where the h/J ratio is supposed to depend on the measurement rate in such a way that

$$\begin{cases} h/J = 0 & \text{when } p = 0, \\ h/J \to \infty & \text{when } p = 1. \end{cases} \tag{D.56}$$

Within this Ising model description, the measurement-induced entanglement transition happens at p_c, corresponding to $h/J = 1$ (at the Ising critical point). Although it should be emphasized that this is only a "mean-field" description of the entanglement transition, and it is known that the measurement-induced criticality is not in the Ising universality class. Nevertheless, this

216

description provides us with ways to phenomenologically model and describe the Pauli weight for the prior snapshots of hybrid shadow tomography away from the critical point.

Given the variational ansatz of the entanglement feature state $|W_{\mathscr{E}_\sigma(p)}\rangle$ in Eq. (D.54), the Pauli weight can be evaluated as

$$
\begin{aligned}
w_{\mathscr{E}_\sigma(p)}(P) &= \frac{\langle P|W_{\mathscr{E}_\sigma(p)}\rangle}{\langle \mathbf{0}|W_{\mathscr{E}_\sigma(p)}\rangle} = \frac{\langle \mathbf{0}|(\prod_{i\in\mathrm{supp}\,P} \frac{2X_i-I_i}{3})W_{\mathbb{1}}|\Psi_p\rangle}{\langle \mathbf{0}|W_{\mathbb{1}}|\Psi_p\rangle} \\
&= \frac{\langle \mathbf{0}|\prod_{i\in\mathrm{supp}\,P}X_i\prod_{i\notin\mathrm{supp}\,P}(2I_i+X_i)|\Psi_p\rangle}{\langle \mathbf{0}|\prod_i(2I_i+X_i)|\Psi_p\rangle},
\end{aligned}
\tag{D.57}
$$

where the denominator is introduced to normalize the Pauli weight (such that $\forall p: w_{\mathscr{E}_\sigma(p)}(I) = 1$) as the ansatz state $|W_{\mathscr{E}_\sigma(p)}\rangle$ was unnormalized. Introduce a binary indicator

$$
\theta_{i\notin P} := \begin{cases} 0 & i \in \mathrm{supp}\,P, \\ 1 & i \notin \mathrm{supp}\,P, \end{cases}
\tag{D.58}
$$

Eq. (D.57) can be written in a more compact form

$$
w_{\mathscr{E}_\sigma(p)}(P) = \frac{\langle \mathbf{0}|\prod_i(2\theta_{i\notin P}I_i+X_i)|\Psi_p\rangle}{\langle \mathbf{0}|\prod_i(2I_i+X_i)|\Psi_p\rangle}.
\tag{D.59}
$$

D.1.3   Volume-Law Phase ($p < p_c$)

At the $p = 0$ limit, the entanglement feature state is given by Eq. (D.51), corresponding to $|\Psi_{p=0}\rangle = |\mathbf{0}\rangle$ as the ground state of the Ising model Eq. (D.55) at its ferromagnetic fixed point ($h/J = 0$). Given that $|\Psi_{p=0}\rangle = |\mathbf{0}\rangle$ is a product state, the Pauli weight Eq. (D.57) can be evaluated at each site independently,

$$
w_{\mathscr{E}_\sigma(p=0)}(P) = \prod_{i\in\mathrm{supp}\,P} \frac{\langle 0|X|0\rangle}{\langle 0|2I+X|0\rangle} = 0.
\tag{D.60}
$$

The Pauli weight vanishes for all P because without any measurement ($p = 0$) there is no way to infer any information about any non-trivial Pauli observable.

To move away from this extreme limit, we can turn on the transverse field term h in the Ising model Eq. (D.55) and use perturbation theory to estimate the corrected ground state $|\Psi_p\rangle$ in the $h/J \ll 1$ regime. To the 1st order of h/J, we have

$$|\Psi_p\rangle \simeq \left(1 + \frac{h}{4J} \sum_i X_i\right)|\mathbf{0}\rangle \simeq e^{\frac{h}{4J} \sum_i X_i}|\mathbf{0}\rangle = \prod_i e^{\frac{h}{4J} X_i}|\mathbf{0}\rangle, \tag{D.61}$$

where we exponentiate the leading order correction. With this, the Pauli weight becomes

$$w_{\mathcal{E}_\sigma}(P) = \prod_{i \in \mathrm{supp}\, P} \frac{\langle 0|X e^{\frac{h}{4J} X}|0\rangle}{\langle 0|(2I + X)e^{\frac{h}{4J} X}|0\rangle} = \left(\frac{1}{1 + 2\coth(\frac{h}{4J})}\right)^{|\mathrm{supp}\, P|}. \tag{D.62}$$

Assuming $h/J \propto p$ is linear in p, the above result implies that the shadow norm should scale with the Pauli operator size $k = |\mathrm{supp}\, P|$ as

$$\|P\|_{\mathcal{E}_\sigma}^2 = \frac{1}{w_{\mathcal{E}_\sigma}(P)} = (1 + 2\coth(cp))^k = \beta^k, \tag{D.63}$$

where c is some unknown coefficient. So the base $\beta = 1 + 2\coth(cp)$ will diverge as $\beta \simeq 1/p$ as the measurement rate $p \to 0$ approaches zero, which is consistent with our numerical result in the main text.

D.1.4 Area-Law Phase ($p > p_c$)

At the $p = 1$ limit, the entanglement feature state is given by Eq. (D.53), corresponding to $|\Psi_{p=1}\rangle = |+\rangle$ as the ground state of the Ising model Eq. (D.55) at is paramagnetic fixed point ($J = 0$). Given that $|\Psi_{p=1}\rangle = |+\rangle$ is a project state, the Pauli weight Eq. (D.57) can be evaluated at each site independently,

$$w_{\mathcal{E}_\sigma(p=1)}(P) = \prod_{i \in \mathrm{supp}\, P} \frac{\langle 0|X|+\rangle}{\langle 0|2I + X|+\rangle} = \prod_{i \in \mathrm{supp}\, P} \frac{\langle 0|+\rangle}{3\langle 0|+\rangle} = \frac{1}{3^{|\mathrm{supp}\, P|}}, \tag{D.64}$$

given that $X|+\rangle = |+\rangle$. This reproduces the known result of Pauli weight for the random product state ensemble.

To move away from this extreme limit, we can turn on the Ising coupling J in the Ising model Eq. (D.55) and use perturbation theory to estimate the corrected ground state $|\Psi_p\rangle$ in the $J/h \ll 1$ regime. To the 1st order of J/h, we have

$$|\Psi_p\rangle \simeq \left(1 + \frac{J}{4h}\sum_i Z_i Z_{i+1}\right)|+\rangle. \tag{D.65}$$

With this, we can evaluate the inner product

$$
\begin{aligned}
\langle 0|\prod_i (2\theta_{i\notin P} I_i + X_i)|\Psi_p\rangle &\simeq \langle 0|\prod_i (2\theta_{i\notin P} I_i + X_i)|+\rangle + \frac{J}{4h}\sum_j \langle 0|\prod_i (2\theta_{i\notin P} I_i + X_i) Z_j Z_{j+1}|+\rangle \\
&= \left(1 + \frac{J}{4h}\sum_j \frac{2\theta_{j\notin P} - 1}{2\theta_{j\notin P} + 1}\frac{2\theta_{j+1\notin P} - 1}{2\theta_{j+1\notin P} + 1}\right)\prod_i (2\theta_{i\notin P} + 1)\langle 0|+\rangle \\
&\simeq e^{\frac{J}{4h}\sum_j \phi_{j,P}}\prod_i (2\theta_{i\notin P} + 1)\langle 0|+\rangle \\
&= \prod_i e^{\frac{J}{4h}\phi_{i,P}}(2\theta_{i\notin P} + 1)\langle 0|+\rangle,
\end{aligned}
\tag{D.66}
$$

where we have introduced the symbol

$$\phi_{i,P} = \frac{2\theta_{j\notin P} - 1}{2\theta_{j\notin P} + 1}\frac{2\theta_{j+1\notin P} - 1}{2\theta_{j+1\notin P} + 1} = \begin{cases} 1 & i + \frac{1}{2} \in \operatorname{supp} P, \\ -\frac{1}{3} & i + \frac{1}{2} \in \partial \operatorname{supp} P, \\ \frac{1}{9} & i + \frac{1}{2} \notin \operatorname{supp} P, \end{cases} \tag{D.67}$$

and exponentiate the perturbation (given $J/H \ll 1$). Given the result Eq. (D.66), we can evaluate

the Pauli weight using Eq. (D.59),

$$
\begin{aligned}
w_{\mathscr{E}_{\sigma}(p)}(P) &= \frac{\prod_i e^{\frac{J}{4h}\phi_{i,P}}(2\theta_{i\notin P}+1)\langle\mathbf{0}|+\rangle}{\prod_i(3e^{\frac{J}{4h}\frac{1}{9}})\langle\mathbf{0}|+\rangle} \\
&\simeq \prod_{i\in\operatorname{supp}P}\frac{1}{3}e^{\frac{J}{4h}(1-\frac{1}{9})} \\
&= \left(\frac{1}{3e^{-\frac{2J}{9h}}}\right)^{|\operatorname{supp}P|}.
\end{aligned}
\tag{D.68}
$$

Assuming $J/h \propto (1-p)$ is linear in $(1-p)$, the above result implies that the shadow norm should scale with the Pauli operator size $k = |\operatorname{supp}P|$ as

$$
\|P\|_{\mathscr{E}_{\sigma}}^2 = \frac{1}{w_{\mathscr{E}_{\sigma}}(P)} = (3e^{-c'(1-p)})^k = \beta^k,
\tag{D.69}
$$

where c' is some unknown coefficient. The base $\beta = 3e^{-c'(1-p)} \simeq 3(1-c'(1-p))$ will decrease from 3 as p deviates from 1, which is also consistent with our numerical result in the main text.

D.1.5   Entanglement Transition ($p = p_c$)

At $p = p_c$, the hybrid quantum circuit undergoes measurement-induced entanglement transition. Although the transition is not described by the Ising CFT, it is instructive to gain a qualitative understanding about the transition using the Ising analogy. In the Ising analogy, the entanglement transition corresponds to the Ising critical point in the Ising model Eq. (D.55).

Suppose the ground state $|\Psi_p\rangle$ is now described by the Ising CFT. Using the Kramers-Wannier duality: $X_i \leftrightarrow \tilde{Z}_{i-1/2}\tilde{Z}_{i+1/2}$, the string operator $\prod_{i\in A}X_i$ for a contiguous region A can be mapped to the product of the dual Ising operator $\prod_{j\in\partial A}\tilde{Z}_j$, such that

$$
\langle\mathbf{0}|\prod_{i\in A}X_i|\Psi_p\rangle = \langle\mathbf{0}|\prod_{j\in\partial A}\tilde{Z}_j|\Psi_p\rangle \sim |A|^{-2\Delta}\langle\mathbf{0}|\Psi_p\rangle,
\tag{D.70}
$$

where Δ corresponds to the scaling dimension of the dual Ising operator in the boundary CFT.

Using this interpretation, we can expand

$$\langle\mathbf{0}|\prod_{i\in A}(2I_i+X_i)|\Psi_p\rangle = 2^{|A|}\sum_{B\subseteq A}\frac{1}{2^{|B|}}\langle\mathbf{0}|\prod_{i\in B}X_i|\Psi_p\rangle$$
$$= 2^{|A|}\sum_{B\subseteq A}\frac{1}{2^{|B|}}\langle\mathbf{0}|\prod_{j\in\partial B}\tilde{Z}_j|\Psi_p\rangle. \tag{D.71}$$

We can give a loose bound for the correlation by

$$\delta_{B\emptyset}\leq\frac{\langle\mathbf{0}|\prod_{j\in\partial B}\tilde{Z}_j|\Psi_p\rangle}{\langle\mathbf{0}|\Psi_p\rangle}\leq 1, \tag{D.72}$$

therefore

$$2^{|A|}=2^{|A|}\sum_{B\subseteq A}\frac{1}{2^{|B|}}\delta_{B\emptyset}\leq\frac{\langle\mathbf{0}|\prod_{i\in A}(2I_i+X_i)|\Psi_p\rangle}{\langle\mathbf{0}|\Psi_p\rangle}\leq 2^{|A|}\sum_{B\subseteq A}\frac{1}{2^{|B|}}=2^{|A|}\left(1+\frac{1}{2}\right)^{|A|}=3^{|A|}. \tag{D.73}$$

Based on Eq. (D.57), the Pauli weight can be estimated by

$$w_{\mathscr{E}_\sigma(p)}(P)\simeq\frac{\langle\mathbf{0}|\prod_{i\in\mathrm{supp}\,P}X_i|\Psi_p\rangle}{\langle\mathbf{0}|\prod_{i\in\mathrm{supp}\,P}(2I_i+X_i)|\Psi_p\rangle}, \tag{D.74}$$

which, given Eq. (D.70) and Eq. (D.73), can be bounded by

$$2^{-k}k^{-2\Delta}\geq w_{\mathscr{E}_\sigma(p)}(P)\geq 3^{-k}k^{-2\Delta}, \tag{D.75}$$

where $k=|\mathrm{supp}\,P|$ is the size of the Pauli operator. Thus we conclude that the Pauli weight at $p=p_c$ should take the form of

$$w_{\mathscr{E}_\sigma(p)}(P)=\beta^{-k}k^{-2\Delta}, \tag{D.76}$$

with a base $2 \leq \beta \leq 3$. Correspondingly, the shadow norm takes the form of

$$\|P\|_{\mathscr{E}_\sigma}^2 = \frac{1}{w_{\mathscr{E}_\sigma}(P)} = \beta^k k^{2\Delta}, \tag{D.77}$$

a form that is consistent with our numerical result in the main text. Our fitting in the main text shows that $\beta \approx 2.2$ is within the bound.

D.1.6 Summary of Results

In conclusion, the quantum statistical mechanical model presented in this appendix shows that the shadow norm takes the following form

$$\|P\|_{\mathscr{E}_\sigma}^2 = \begin{cases} \beta_{\text{vol}}^k & p < p_c, \\ \beta_{\text{min}}^k k^{2\Delta} & p = p_c, \\ \beta_{\text{are}}^k & p = p_c. \end{cases} \tag{D.78}$$

- In the volume law phase, the base $\beta_{\text{vol}} \sim 1/p$ diverges as $p \to 0$.

- In the area law phase, the base $\beta_{\text{are}} \sim 3e^{-c'(1-p)}$ approaches 3 from below as $p \to 1$.

- Given that β increases in both phases as we go away from the critical point, the entanglement transition should have the minimal β, denoted as β_{min}, and (loosely) bounded by $2 \leq \beta_{\text{min}} \leq 3$.

- At the critical point, there is a power law correction $k^{2\Delta}$, with universal exponent Δ.

D.2 Toy Models

D.2.1 Area-Law Phase ($p > p_c$)

When the measurement rate p is greater than the critical value p_c, the classical snapshot σ is in the area-law phase. The measurement circuit can be modeled by Fig. D.2(a), whose corresponding classical snapshot state σ is a product of n-qubit random stabilizer state. Within

each block, the entanglement is maximal. Between different blocks, there is no entanglement (product state). Therefore, the block size n parameterizes the typical range of local entanglement in the classical snapshot state σ.

Figure D.2. (a) A toy model for the randomized measurement in the area-law phase. Each green block represents a n-qubit random Clifford gate. (b) The dependence of the shadow norm scaling base β on the block size n.

The limit $p = 1$ corresponds to $n = 1$, that every qubit gets measured immediately before any inter-qubit scrambling. As the measurement rate p reduces (but still in the area-law phase $p_c < p < 1$), the system will be scrambled within the local entanglement range before gets probed by the measurement, but the scrambling range is still finite. This situation can be modeled by $n > 1$. We expect n to increase effectively as we p reduces from 1.

Consider a Pauli observable P whose support happens to cover m of the n-qubit blocks. On one hand, the operator size of P is $|\operatorname{supp} P| = mn$. On the other hand, its Pauli weight is given by

$$w_{\mathcal{E}_\sigma}(P) = \left(\frac{2^n - 1}{4^n - 1}\right)^m = \frac{1}{(2^n + 1)^m}. \tag{D.79}$$

This result can be understood as follows. The Pauli weight can be interpreted as the probability that a Pauli observable P gets transformed by the random Clifford gates into a diagonal operator in the measurement basis, such that it can be directly probed by the measurement. Within each n-qubit block, the random Clifford gate scrambles any particular non-identity Pauli observable to one of all $(4^n - 1)$ non-identity Pauli observables, among which only $(2^n - 1)$ are diagonal (as

Pauli strings of I and Z) in the measurement basis. Hence the probability of P to be diagonalized within each block is $(2^n - 1)/(4^n - 1) = 1/(2^n + 1)$. To diagonalize P across all the m blocks, the probability multiplies to $1/(2^n + 1)^m$, as concluded in Eq. (D.79). As a result, the shadow norm of P is

$$\|P\|_{\mathscr{E}_\sigma}^2 = \frac{1}{w_{\mathscr{E}_\sigma}(P)} = (2^n + 1)^m. \tag{D.80}$$

Assuming the shadow norm scales with the operator size as $\|P\|_{\mathscr{E}_\sigma}^2 = \beta^{|\operatorname{supp} P|}$, the base β can be extracted as

$$\log \beta = \lim_{|\operatorname{supp} P| \to \infty} \frac{\log \|P\|_{\mathscr{E}_\sigma}^2}{|\operatorname{supp} P|} = \lim_{m \to \infty} \frac{m \log(2^n + 1)}{mn} = \frac{\log(2^n + 1)}{n}. \tag{D.81}$$

The dependence of β on n is shown in Fig. D.2(b). As the measurement rate p decreases from 1, the scrambling range n increases, and β will eventually decrease from 3 to 2 (in this toy model's ideal case).

D.2.2 Volume-Law Phase ($p < p_c$)

When the measurement rate p is less than the critical value p_c, the classical snapshot σ is in the volume-law phase with local error correction encoding. A model to describe such states is to consider random stabilizer states further encoded by random stabilizer codes. The corresponding measurement circuit will be like Fig. D.3(a). The system contains N physical qubits, grouped by n-qubit blocks. Within each block, the n physical qubit is first decoded into a logical qubit and $(n - 1)$ syndrome qubits. The syndrome qubits are measured, and the logical qubits are further scrambled before being finally measured. We assume $N \gg n \gg 1$.

Consider an entanglement region A that covers exactly m of the n-qubit blocks (assuming $mn \ll N$), the entanglement entropy of σ in such region scales as $S_\sigma(A) = m \log 2$, while the region size is $|A| = mn$. So σ is a volume-law state with the volume-law coefficient

$$f = \frac{S_\sigma(A)}{|A| \log 2} = \frac{1}{n}. \tag{D.82}$$

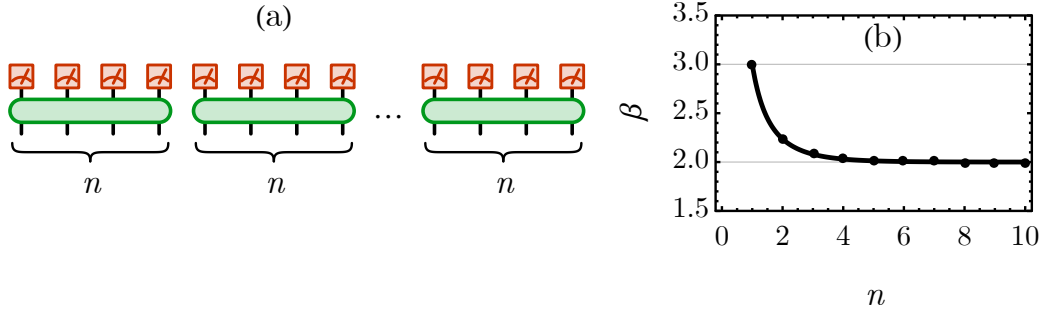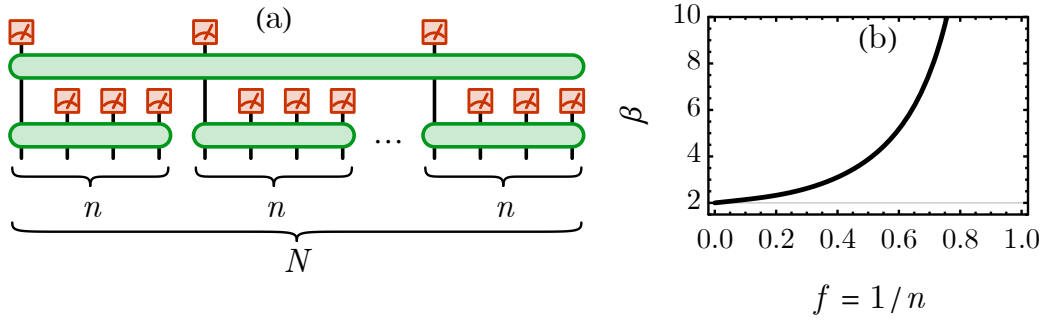**Figure D.3.** (a) A toy model for the randomized measurement in the volume-law phase. Each green block represents a random Clifford gate. The lower layer gates serve as the encoding gates of random stabilizer codes in each n-qubit block. The upper layer gate scrambles all the N/n logical bits. (b) The dependence of the shadow norm scaling base β on the volume-law coefficient f.

This $1/n$ ratio is also the rate of the local error correction code (the ratio between the numbers of logical v.s. physical qubits). As the measurement rate p increases from 0 to p_c, the volume-law coefficient f decreases from 1 to 0, corresponding to n increasing from 1 to ∞.

Consider a Pauli observable P whose support happens to cover m of the n-qubit blocks. On one hand, the operator size of P is $|\operatorname{supp} P| = mn$. On the other hand, its Pauli weight is given by

$$w_{\mathscr{E}_\sigma}(P) = q^m + \varepsilon(r^m - q^m), \quad \text{with}$$

$$
\begin{aligned}
\varepsilon &= \frac{2^{N/n} - 1}{4^{N/n} - 1} = \frac{1}{2^{N/n} + 1}, \\
q &= \frac{2^{n-1} - 1}{4^n - 1}, \\
r &= \frac{4 \times 2^{n-1} - 1}{4^n - 1}.
\end{aligned}
\tag{D.83}
$$

This result can be understood as follows. The Pauli weight can be interpreted as the probability that a Pauli observable P gets transformed by the random Clifford gates into a diagonal operator in the measurement basis, such that it can be directly probed by the measurement. There are two possible scenarios:

225

- P gets transformed to an operator that is non-identity and diagonal in the syndrome subspace and identity in the logical subspace. The probability for this to happen is q^m. Because in each block, any particular non-identity Pauli observable will be transformed to one of the $(4^n - 1)$ non-identity Pauli observable, among which only $(2^{n-1} - 1)$ are syndrome subspace diagonal and logical subspace identity. They are of the following form:

$$I \otimes \begin{bmatrix} I \\ Z \end{bmatrix}^{\otimes(n-1)} \quad \text{excluding } I^{\otimes n}. \tag{D.84}$$

So the probability for this to happen is $q = (2^{n-1} - 1)/(4^n - 1)$ in each block and q^m over m blocks. In this case, P will be directly measured by the syndrome qubit measurements. After the measurement, it collapses to an identity operator on logical qubits, whose measurement outcome can be determined with probability 1. Therefore, this scenario has a total contribution of $q^m \times 1$ to the Pauli weight, corresponding to the first term in Eq. (D.83).

- P gets transformed to an operator that is non-identity and diagonal in both the syndrome and logical subspaces. The probability for this to happen is $\varepsilon(r^m - q^m)$. Within each block, any particular non-identity Pauli observable will be transformed to one of the $(4^n - 1)$ non-identity Pauli observables, among which only $(4 \times 2^{n-1} - 1)$ are diagonal in the syndrome subspace regardless of its action in the logical subspace. They are of the following form

$$\begin{bmatrix} I \\ X \\ Y \\ Z \end{bmatrix} \otimes \begin{bmatrix} I \\ Z \end{bmatrix}^{\otimes(n-1)} \quad \text{excluding } I^{\otimes n}. \tag{D.85}$$

So the probability for this to happen is $r = (4 \times 2^{n-1} - 1)/(4^n - 1)$ and r^m over m blocks. After the syndrome qubit measurements, the operator will collapse to a Pauli string in the

logical subspace of the following form

$$
\begin{bmatrix} I \\ X \\ Y \\ Z \end{bmatrix}^{\otimes m}.
\tag{D.86}
$$

This has not excluded the possibility of $I^{\otimes m}$, which should be excluded to avoid double counting the same scenario discussed previously. Thankfully, we already know that the probability for P to become a logical identity operator after the syndrome qubit measurement is q^m. So the probability for P to be syndrome diagonal and logical non-identity is $(r^m - q^m)$. Under the global scrambling of N/n logical qubits by the upper layer gate, the observable P gets further mapped to one of the $(4^{N/n} - 1)$ non-identity Pauli operators in the logical subspace, among which only $(2^{N/n} - 1)$ are diagonal and can be directly probed by the logical qubit measurements. So this further multiplies the probability by $\varepsilon = (2^{N/n} - 1)/(4^{N/n} - 1)$, resulting in a contribution of $\varepsilon(r^m - q^m)$ to the Pauli weight, corresponding to the second term in Eq. (D.83).

According to Eq. (D.83), $\varepsilon \to 0$ in the thermodynamic limit $N \gg n$, so the second term can be neglected, and the Pauli weight is dominated by the first term

$$
w_{\mathscr{E}_\sigma}(P) \stackrel{N \gg n}{=} q^m = \left(\frac{2^{n-1} - 1}{4^n - 1} \right)^m.
\tag{D.87}
$$

In conclusion, the shadow norm of P in this toy model is given by

$$
\|P\|_{\mathscr{E}_\sigma}^2 = \frac{1}{w_{\mathscr{E}_\sigma}(P)} \stackrel{N \gg n}{=} \left(\frac{4^n - 1}{2^{n-1} - 1} \right)^m.
\tag{D.88}
$$

Assuming the shadow norm scales with the operator size as $\|P\|_{\mathscr{E}_\sigma}^2 = \beta^{|\operatorname{supp} P|}$, the base β can be

extracted as

$$\log \beta = \lim_{|\mathrm{supp}\,P| \to \infty} \frac{\log \|P\|_{\mathscr{E}_\sigma}^2}{|\mathrm{supp}\,P|} = \lim_{m \to \infty} \frac{1}{mn} m \log\left(\frac{4^n - 1}{2^{n-1} - 1}\right) = \frac{1}{n}\log\left(\frac{4^n - 1}{2^{n-1} - 1}\right). \qquad \text{(D.89)}$$

Given that n is related to the volume law coefficient $f = 1/n$ by Eq. (D.82), the base β can also be expressed in terms of f as

$$\log \beta = f \log\left(\frac{4^{1/f} - 1}{2^{1/f-1} - 1}\right) \stackrel{f \to 0}{=} (1 + f)\log 2. \qquad \text{(D.90)}$$

The dependence of β on f is shown in Fig. D.3(b). As the measurement strength increases from 0 to p_c, the volume-law coefficient f decreases from 1 to 0, and β decreases from ∞ to 2. Near the measurement induced phase transition (as $f \to 0$), β increases with the volume-law coefficient f as $\beta = 2^{1+f}$.

Bibliography

[1] Scott Aaronson. Shadow Tomography of Quantum States. *arXiv e-prints*, page arXiv:1711.01053, November 2017.

[2] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[3] Scott Aaronson and Guy N. Rothblum. Gentle Measurement of Quantum States and Differential Privacy. *arXiv e-prints*, page arXiv:1904.08747, April 2019.

[4] Dmitry A. Abanin, Ehud Altman, Immanuel Bloch, and Maksym Serbyn. Colloquium: Many-body localization, thermalization, and entanglement. *Rev. Mod. Phys.*, 91:021001, May 2019.

[5] A. A. Akhtar and Yi-Zhuang You. Multiregion entanglement in locally scrambled quantum dynamics. *Physical Review B*, 102(13):134203, October 2020.

[6] Ahmed A. Akhtar, Hong-Ye Hu, and Yi-Zhuang You. Scalable and Flexible Classical Shadow Tomography with Tensor Networks. *Quantum*, 7:1026, June 2023.

[7] Vincenzo Alba. Entanglement negativity and conformal field theory: a monte carlo study. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(05):P05013, may 2013.

[8] Vincenzo Alba and Pasquale Calabrese. Entanglement and thermodynamics after a quantum quench in integrable systems. *Proceedings of the National Academy of Science*, 114(30):7947–7951, Jul 2017.

[9] Alba, Vincenzo and Calabrese, Pasquale. Quantum information dynamics in multipartite integrable systems. *EPL*, 126(6):60001, 2019.

[10] Huzihiro Araki and Elliott H. Lieb. Entropy inequalities. *Communications in Mathematical Physics*, 18(2):160–170, Jun 1970.

[11] Mirko Arienzo, Markus Heinrich, Ingo Roth, and Martin Kliesch. Closed-form analytic expressions for shadow estimation with brickwork circuits, 2022.

[12] Mirko Arienzo, Markus Heinrich, Ingo Roth, and Martin Kliesch. Closed-form analytic expressions for shadow estimation with brickwork circuits. *arXiv e-prints*, page arXiv:2211.09835, November 2022.

[13] Yimu Bao, Soonwon Choi, and Ehud Altman. Theory of the Phase Transition in Random Unitary Circuits with Measurements. *arXiv e-prints*, page arXiv:1908.04305, Aug 2019.

[14] Yimu Bao, Soonwon Choi, and Ehud Altman. Theory of the phase transition in random unitary circuits with measurements. *Physical Review B*, 101(10):104301, March 2020.

[15] Yimu Bao, Soonwon Choi, and Ehud Altman. Theory of the phase transition in random unitary circuits with measurements. *Physical Review B*, 101(10):104301, March 2020.

[16] Yimu Bao, Soonwon Choi, and Ehud Altman. Symmetry enriched phases of quantum circuits. *Annals of Physics*, 435:168618, December 2021.

[17] Jens H. Bardarson, Frank Pollmann, and Joel E. Moore. Unbounded Growth of Entanglement in Models of Many-Body Localization. *Physical Review Letters*, 109(1):017202, Jul 2012.

[18] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.

[19] Bruno Bertini, Pavel Kos, and Tomaž Prosen. Entanglement Spreading in a Minimal Model of Maximal Many-Body Quantum Chaos. *Physical Review X*, 9(2):021033, Apr 2019.

[20] Christian Bertoni, Jonas Haferkamp, Marcel Hinsche, Marios Ioannou, Jens Eisert, and Hakop Pashayan. Shallow shadows: Expectation estimation using low-depth random Clifford circuits. *arXiv e-prints*, page arXiv:2209.12924, September 2022.

[21] Christian Bertoni, Jonas Haferkamp, Marcel Hinsche, Marios Ioannou, Jens Eisert, and Hakop Pashayan. Shallow shadows: Expectation estimation using low-depth random Clifford circuits. *arXiv e-prints*, page arXiv:2209.12924, September 2022.

[22] Eugenio Bianchi, Lucas Hackl, Mario Kieburg, Marcos Rigol, and Lev Vidmar. Volume-law entanglement entropy of typical pure quantum states. *PRX Quantum*, 3(3), jul 2022.

[23] O. Bohigas, M. J. Giannoni, and C. Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.*, 52:1–4, Jan 1984.

[24] A. Bohrdt, C. B. Mendl, M. Endres, and M. Knap. Scrambling and thermalization in a diffusive quantum many-body system. *New Journal of Physics*, 19(6):063001, June 2017.

[25] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin. Quantum source of entropy for black holes. *Phys. Rev. D*, 34:373–383, Jul 1986.

[26] Fernando G. S. L. Brandão and Michał Horodecki. Exponential decay of correlations implies area law. *Communications in Mathematical Physics*, 333(2):761–798, nov 2014.

[27] Kaifeng Bu, Dax Enshan Koh, Roy J. Garcia, and Arthur Jaffe. Classical shadows with Pauli-invariant unitary ensembles. *arXiv e-prints*, page arXiv:2202.03272, February 2022.

[28] Kaifeng Bu, Dax Enshan Koh, Roy J. Garcia, and Arthur Jaffe. Classical shadows with pauli-invariant unitary ensembles, 2022.

[29] Pasquale Calabrese and John Cardy. Evolution of entanglement entropy in one-dimensional systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(4):04010, Apr 2005.

[30] Pasquale Calabrese and John Cardy. Entanglement entropy and conformal field theory. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504005, dec 2009.

[31] Pasquale Calabrese, John Cardy, and Erik Tonni. Entanglement negativity in quantum field theory. *Phys. Rev. Lett.*, 109:130502, Sep 2012.

[32] Pasquale Calabrese, John Cardy, and Erik Tonni. Entanglement negativity in extended systems: a field theoretical approach. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(02):P02008, feb 2013.

[33] Carlton M. Caves, Christopher A. Fuchs, and Rüdiger Schack. Unknown quantum states: The quantum de Finetti representation. *Journal of Mathematical Physics*, 43(9):4537–4559, September 2002.

[34] Amos Chan, Andrea De Luca, and J. T. Chalker. Solution of a Minimal Model for Many-Body Quantum Chaos. *Physical Review X*, 8(4):041019, Oct 2018.

[35] Amos Chan, Andrea De Luca, and J. T. Chalker. Solution of a Minimal Model for Many-Body Quantum Chaos. *Physical Review X*, 8(4):041019, October 2018.

[36] Amos Chan, Andrea De Luca, and J. T. Chalker. Spectral Statistics in Spatially Extended Chaotic Quantum Many-Body Systems. *Physical Review Letters*, 121(6):060601, Aug 2018.

[37] Amos Chan, Andrea De Luca, and J. T. Chalker. Spectral statistics in spatially extended chaotic quantum many-body systems. *Physical Review Letters*, 121(6), aug 2018.

[38] Amos Chan, Rahul M. Nandkishore, Michael Pretko, and Graeme Smith. Unitary-

projective entanglement dynamics. *Physical Review B*, 99(22):224307, Jun 2019.

[39] A. Chandran, C. R. Laumann, and V. Oganesyan. Finite size scaling bounds on many-body localized phase transitions. *arXiv e-prints*, page arXiv:1509.04285, Sep 2015.

[40] Senrui Chen, Wenjun Yu, Pei Zeng, and Steven T. Flammia. Robust shadow estimation. *PRX Quantum*, 2:030348, Sep 2021.

[41] Xiao Chen and Tianci Zhou. Quantum chaos dynamics in long-range power law interaction systems. *Physical Review B*, 100(6):064305, Aug 2019.

[42] Xiao Chen and Tianci Zhou. Quantum chaos dynamics in long-range power law interaction systems. *Physical Review B*, 100(6):064305, August 2019.

[43] Xiao Chen and Tianci Zhou. Quantum chaos dynamics in long-range power law interaction systems. *Physical Review B*, 100(6):064305, August 2019.

[44] Soonwon Choi, Yimu Bao, Xiao-Liang Qi, and Ehud Altman. Quantum error correction and entanglement phase transition in random unitary circuits with projective measurements. *arXiv e-prints*, page arXiv:1903.05124, Mar 2019.

[45] Soonwon Choi, Yimu Bao, Xiao-Liang Qi, and Ehud Altman. Quantum Error Correction in Scrambling Dynamics and Measurement-Induced Phase Transition. *Physical Review Letters*, 125(3):030505, July 2020.

[46] Soonwon Choi, Yimu Bao, Xiao-Liang Qi, and Ehud Altman. Quantum Error Correction in Scrambling Dynamics and Measurement-Induced Phase Transition. *Physical Review Letters*, 125(3):030505, July 2020.

[47] Chia-Min Chung, Vincenzo Alba, Lars Bonnes, Pochung Chen, and Andreas M. Läuchli. Entanglement negativity via the replica trick: A quantum monte carlo approach. *Phys. Rev. B*, 90:064401, Aug 2014.

[48] Ze-Pei Cian, Hossein Dehghani, Andreas Elben, Benoît Vermersch, Guanyu Zhu, Maissam Barkeshli, Peter Zoller, and Mohammad Hafezi. Many-body chern number from statistical correlations of randomized measurements. *Phys. Rev. Lett.*, 126:050501, Feb 2021.

[49] J. Ignacio Cirac, David Pé rez-García, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4), dec 2021.

[50] J. Ignacio Cirac, David Pé rez-García, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4), dec 2021.

[51] Benoit Collins. Moments and cumulants of polynomial random variables on unitary groups, the itzykson-zuber integral and free probability, 2002.

[52] Benoît Collins and Piotr Śniady. Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group. *Communications in Mathematical Physics*, 264(3):773–795, Jun 2006.

[53] Josiah Couch, Stefan Eccles, Phuc Nguyen, Brian Swingle, and Shenglong Xu. Speed of quantum information spreading in chaotic systems. *Phys. Rev. B*, 102:045114, Jul 2020.

[54] G. M. D'Ariano and P. Perinotti. Optimal Data Processing for Quantum Measurements. *Physical Review Letters*, 98(2):020403, January 2007.

[55] Giuseppe De Tomasi, Daniel Hetterich, Pablo Sala, and Frank Pollmann. Dynamics of strongly interacting systems: From Fock-space fragmentation to many-body localization. *Physical Review B*, 100(21):214313, December 2019.

[56] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[57] J. Eisert, M. Friesdorf, and C. Gogolin. Quantum many-body systems out of equilibrium. *Nature Physics*, 11(2):124–130, Feb 2015.

[58] Andreas Elben, Steven T. Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller. The randomized measurement toolbox. *arXiv e-prints*, page arXiv:2203.11374, March 2022.

[59] Andreas Elben, Richard Kueng, Hsin-Yuan Robert Huang, Rick van Bijnen, Christian Kokail, Marcello Dalmonte, Pasquale Calabrese, Barbara Kraus, John Preskill, Peter Zoller, and Benoît Vermersch. Mixed-State Entanglement from Local Randomized Measurements. *Physical Review Letters*, 125(20):200501, November 2020.

[60] Dax Enshan Koh and Sabee Grewal. Classical Shadows With Noise. *Quantum*, 6:776, August 2022.

[61] Ruihua Fan, Sagar Vijay, Ashvin Vishwanath, and Yi-Zhuang You. Self-organized error correction in random unitary circuits with measurement. 2020.

[62] Ruihua Fan, Sagar Vijay, Ashvin Vishwanath, and Yi-Zhuang You. Self-organized error correction in random unitary circuits with measurement, Feb 2020.

[63] Ruihua Fan, Sagar Vijay, Ashvin Vishwanath, and Yi-Zhuang You. Self-organized error correction in random unitary circuits with measurement. *Physical Review B*, 103(17):174309, May 2021.

[64] Ruihua Fan, Pengfei Zhang, Huitao Shen, and Hui Zhai. Out-of-time-order correlation for many-body localization. *Science Bulletin*, 62(10):707 – 711, May 2017.

[65] Matthew P. A. Fisher, Vedika Khemani, Adam Nahum, and Sagar Vijay. Random Quantum Circuits. *Annual Review of Condensed Matter Physics*, 14:335–379, March 2023.

[66] Steven T. Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, September 2012.

[67] Martin Gärttner, Justin G. Bohnet, Arghavan Safavi-Naini, Michael L. Wall, John J. Bollinger, and Ana Maria Rey. Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet. *Nature Physics*, 13(8):781–786, Aug 2017.

[68] Chenhua Geng, Hong-Ye Hu, and Yijian Zou. Differentiable programming of isometric tensor networks. *Machine Learning: Science and Technology*, 3(1):015020, jan 2022.

[69] Hrant Gharibyan, Masanori Hanada, Stephen H. Shenker, and Masaki Tezuka. Onset of random matrix behavior in scrambling systems. *Journal of High Energy Physics*, 2018(7):124, Jul 2018.

[70] Hrant Gharibyan, Masanori Hanada, Stephen H. Shenker, and Masaki Tezuka. Onset of random matrix behavior in scrambling systems. *Journal of High Energy Physics*, 2018(7):124, July 2018.

[71] Hrant Gharibyan, Masanori Hanada, Brian Swingle, and Masaki Tezuka. Quantum Lyapunov spectrum. *Journal of High Energy Physics*, 2019(4):82, Apr 2019.

[72] E. N. Gilbert and J. Riordan. Symmetry types of periodic sequences. *Illinois J. Math.*, 5:657–665, 1961.

[73] Sarang Gopalakrishnan, David A. Huse, Vedika Khemani, and Romain Vasseur. Hydrodynamics of operator spreading and quasiparticle diffusion in interacting integrable systems. *Physical Review B*, 98(22):220303, Dec 2018.

[74] Shimpei Goto and Ippei Danshita. Performance of the time-dependent variational principle for matrix product states in the long-time evolution of a pure state. *Physical Review B*, 99(5):054307, February 2019.

[75] D. Gottesman. The Heisenberg Representation of Quantum Computers. *eprint arXiv:quant-ph/9807006*, July 1998.

[76] Daniel Gottesman. *Stabilizer codes and quantum error correction*. PhD thesis, California

Institute of Technology, January 1997.

[77] Daniel Gottesman. The heisenberg representation of quantum computers. 1998.

[78] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going Beyond Bell's Theorem. *arXiv e-prints*, page arXiv:0712.0921, December 2007.

[79] Tarun Grover and Matthew P. A. Fisher. Entanglement and the sign structure of quantum states. *Physical Review A*, 92(4), oct 2015.

[80] Tarun Grover and Matthew P. A. Fisher. Entanglement and the sign structure of quantum states. *Physical Review A*, 92(4):042308, October 2015.

[81] Michael J. Gullans and David A. Huse. Dynamical Purification Phase Transition Induced by Quantum Measurements. *Physical Review X*, 10(4):041020, October 2020.

[82] Madalin Guta, Jonas Kahn, Richard Kueng, and Joel A. Tropp. Fast state tomography with optimal error bounds. *arXiv e-prints*, page arXiv:1809.11162, September 2018.

[83] Madalin Guta, Jonas Kahn, Richard Kueng, and Joel A. Tropp. Fast state tomography with optimal error bounds. *Journal of Physics A: Mathematical and Theoretical*, 53(20):204001, April 2020.

[84] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *arXiv e-prints*, page arXiv:1508.01797, August 2015.

[85] Charles Hadfield, Sergey Bravyi, Rudy Raymond, and Antonio Mezzacapo. Measurements of Quantum Hamiltonians with Locally-Biased Classical Shadows. *arXiv e-prints*, page arXiv:2006.15788, June 2020.

[86] Charles Hadfield, Sergey Bravyi, Rudy Raymond, and Antonio Mezzacapo. Measurements of Quantum Hamiltonians with Locally-Biased Classical Shadows. *Communications in Mathematical Physics*, 391(3):951–967, May 2022.

[87] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, 107:070601, Aug 2011.

[88] Jutho Haegeman, Christian Lubich, Ivan Oseledets, Bart Vandereycken, and Frank Verstraete. Unifying time evolution and optimization with matrix product states. *Phys. Rev. B*, 94:165116, Oct 2016.

[89] Guang Hao Low. Classical shadows of fermions with particle number symmetry. *arXiv*

e-prints, page arXiv:2208.08964, August 2022.

[90] Thomas Hartman and Juan Maldacena. Time evolution of entanglement entropy from black hole interiors. *Journal of High Energy Physics*, 2013:14, May 2013.

[91] M B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, aug 2007.

[92] Markus Hauru, Maarten Van Damme, and Jutho Haegeman. Riemannian optimization of isometric tensor networks. *SciPost Phys.*, 10:40, 2021.

[93] Patrick Hayden, Sepehr Nezami, Xiao-Liang Qi, Nathaniel Thomas, Michael Walter, and Zhao Yang. Holographic duality from random tensor networks. *Journal of High Energy Physics*, 2016(11):9, Nov 2016.

[94] Wen Wei Ho and Dmitry A. Abanin. Entanglement dynamics in quantum many-body systems. *Physical Review B*, 95(9):094302, Mar 2017.

[95] Wen Wei Ho and Dmitry A. Abanin. Entanglement dynamics in quantum many-body systems. *Physical Review B*, 95(9):094302, March 2017.

[96] Pavan Hosur, Xiao-Liang Qi, Daniel A. Roberts, and Beni Yoshida. Chaos in quantum channels. *Journal of High Energy Physics*, 2016(2):4, Feb 2016.

[97] Timothy H. Hsieh and Liang Fu. Bulk Entanglement Spectrum Reveals Quantum Criticality within a Topological State. *Physical Review Letters*, 113(10):106801, Sep 2014.

[98] Hong-Ye Hu, Soonwon Choi, and Yi-Zhuang You. Classical Shadow Tomography with Locally Scrambled Quantum Dynamics. *arXiv e-prints*, page arXiv:2107.04817, July 2021.

[99] Hong-Ye Hu, Ryan LaRose, Yi-Zhuang You, Eleanor Rieffel, and Zhihui Wang. Logical shadow tomography: Efficient estimation of error-mitigated observables. *arXiv e-prints*, page arXiv:2203.07263, March 2022.

[100] Hong-Ye Hu and Yi-Zhuang You. Hamiltonian-driven shadow tomography of quantum states. *Physical Review Research*, 4(1):013054, January 2022.

[101] Hongye Hu. *Efficient representation and learning of quantum many-body states*. PhD thesis, UC San Diego, 2022.

[102] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, June 2020.

[103] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, June 2020.

[104] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Efficient Estimation of Pauli Observables by Derandomization. *Physical Review Letters*, 127(3):030503, July 2021.

[105] Marko Žnidarič. Optimal two-qubit gate for generation of random bipartite entanglement. *Phys. Rev. A*, 76:012318, Jul 2007.

[106] Matteo Ippoliti, Michael J. Gullans, Sarang Gopalakrishnan, David A. Huse, and Vedika Khemani. Entanglement phase transitions in measurement-only dynamics. *Phys. Rev. X*, 11:011030, Feb 2021.

[107] Matteo Ippoliti and Vedika Khemani. Learnability transitions in monitored quantum dynamics via eavesdropper's classical shadows. *arXiv e-prints*, page arXiv:2307.15011, July 2023.

[108] Matteo Ippoliti, Yaodong Li, Tibor Rakovszky, and Vedika Khemani. Operator relaxation and the optimal depth of classical shadows, 2023.

[109] Matteo Ippoliti, Yaodong Li, Tibor Rakovszky, and Vedika Khemani. Operator Relaxation and the Optimal Depth of Classical Shadows. *Physical Review Letters*, 130(23):230403, June 2023.

[110] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Physical Review A*, 64(5):052312, November 2001.

[111] Chao-Ming Jian, Yi-Zhuang You, Romain Vasseur, and Andreas W. W. Ludwig. Measurement-induced criticality in random quantum circuits. *arXiv e-prints*, page arXiv:1908.08051, Aug 2019.

[112] Chao-Ming Jian, Yi-Zhuang You, Romain Vasseur, and Andreas W. W. Ludwig. Measurement-induced criticality in random quantum circuits. *Physical Review B*, 101(10), mar 2020.

[113] Chao-Ming Jian, Yi-Zhuang You, Romain Vasseur, and Andreas W. W. Ludwig. Measurement-induced criticality in random quantum circuits. *Physical Review B*, 101(10):104302, March 2020.

[114] Cheryne Jonay, David A. Huse, and Adam Nahum. Coarse-grained dynamics of operator and state entanglement. *arXiv e-prints*, page arXiv:1803.00089, Feb 2018.

[115] Adam M. Kaufman, M. Eric Tai, Alexander Lukin, Matthew Rispoli, Robert Schittko,

Philipp M. Preiss, and Markus Greiner. Quantum thermalization through entanglement in an isolated many-body system. *Science*, 353(6301):794–800, Aug 2016.

[116] Vedika Khemani, David A. Huse, and Adam Nahum. Velocity-dependent Lyapunov exponents in many-body quantum, semiclassical, and classical chaos. *Physical Review B*, 98(14):144304, Oct 2018.

[117] Vedika Khemani, Ashvin Vishwanath, and David A. Huse. Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws. *Phys. Rev. X*, 8:031057, Sep 2018.

[118] Vedika Khemani, Ashvin Vishwanath, and David A. Huse. Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws. *Physical Review X*, 8(3):031057, July 2018.

[119] Hyungwon Kim and David A. Huse. Ballistic spreading of entanglement in a diffusive nonintegrable system. *Phys. Rev. Lett.*, 111:127205, Sep 2013.

[120] Jonas A. Kjäll, Jens H. Bardarson, and Frank Pollmann. Many-Body Localization in a Disordered Quantum Ising Chain. *Physical Review Letters*, 113(10):107204, Sep 2014.

[121] Benedikt Kloss, Yevgeny Bar Lev, and David Reichman. Time-dependent variational principle in matrix-product state manifolds: Pitfalls and potential. *Phys. Rev. B*, 97:024307, Jan 2018.

[122] Dax Enshan Koh and Sabee Grewal. Classical Shadows With Noise. *Quantum*, 6:776, August 2022.

[123] Jin Ming Koh, Shi-Ning Sun, Mario Motta, and Austin J. Minnich. Measurement-induced entanglement phase transition on a superconducting quantum processor with mid-circuit readout. *Nature Physics*, jun 2023.

[124] Jonah Kudler-Flam, Masahiro Nozaki, Shinsei Ryu, and Mao Tian Tan. Quantum vs. classical information: operator negativity as a probe of scrambling. *arXiv e-prints*, page arXiv:1906.07639, Jun 2019.

[125] Ivan Kukuljan, Sašo Grozdanov, and Tomaž Prosen. Weak quantum chaos. *Physical Review B*, 96(6):060301, August 2017.

[126] Ilya Kull, Andras Molnar, Erez Zohar, and J. Ignacio Cirac. Classification of matrix product states with a local (gauge) symmetry. *Annals of Physics*, 386:199–241, November 2017.

[127] Wei-Ting Kuo, A. A. Akhtar, Daniel P. Arovas, and Yi-Zhuang You. Markovian entangle-

ment dynamics under locally scrambled quantum evolution. *Physical Review B*, 101(22), jun 2020.

[128] Wei-Ting Kuo, A. A. Akhtar, Daniel P. Arovas, and Yi-Zhuang You. Markovian entanglement dynamics under locally scrambled quantum evolution. *Phys. Rev. B*, 101:224202, Jun 2020.

[129] Wei-Ting Kuo, A. A. Akhtar, Daniel P. Arovas, and Yi-Zhuang You. Markovian entanglement dynamics under locally scrambled quantum evolution. *Physical Review B*, 101(22):224202, June 2020.

[130] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the fast scrambling conjecture. *Journal of High Energy Physics*, 2013:22, Apr 2013.

[131] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the fast scrambling conjecture. *Journal of High Energy Physics*, 2013:22, April 2013.

[132] Jong Yeon Lee, Junyeong Ahn, Hengyun Zhou, and Ashvin Vishwanath. Topological Correspondence between Hermitian and Non-Hermitian Systems: Anomalous Dynamics. *arXiv e-prints*, page arXiv:1906.08782, Jun 2019.

[133] Yuri D Lensky and Xiao-Liang Qi. Chaos and High Temperature Pure State Thermalization. *arXiv.org*, May 2018.

[134] Eyal Leviatan, Frank Pollmann, Jens H. Bardarson, David A. Huse, and Ehud Altman. Quantum thermalization dynamics with Matrix-Product States. *arXiv e-prints*, page arXiv:1702.08894, February 2017.

[135] Ryan Levy, Di Luo, and Bryan K. Clark. Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers. *arXiv e-prints*, page arXiv:2110.02965, October 2021.

[136] Ryan Levy, Di Luo, and Bryan K. Clark. Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers. *arXiv e-prints*, page arXiv:2110.02965, October 2021.

[137] Jun Li, Ruihua Fan, Hengyan Wang, Bingtian Ye, Bei Zeng, Hui Zhai, Xinhua Peng, and Jiangfeng Du. Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator. *Phys. Rev. X*, 7:031011, Jul 2017.

[138] Yaodong Li, Xiao Chen, and Matthew P. A. Fisher. Quantum Zeno effect and the many-body entanglement transition. *Physical Review B*, 98(20):205136, Nov 2018.

[139] Yaodong Li, Xiao Chen, and Matthew P. A. Fisher. Quantum Zeno effect and the many-body entanglement transition. *Physical Review B*, 98(20):205136, November 2018.

[140] Yaodong Li, Xiao Chen, and Matthew P. A. Fisher. Measurement-driven entanglement transition in hybrid quantum circuits. *arXiv e-prints*, page arXiv:1901.08092, Jan 2019.

[141] Yaodong Li, Xiao Chen, and Matthew P. A. Fisher. Measurement-driven entanglement transition in hybrid quantum circuits. *Physical Review B*, 100(13):134306, October 2019.

[142] Yaodong Li, Yijian Zou, Paolo Glorioso, Ehud Altman, and Matthew P. A. Fisher. Cross Entropy Benchmark for Measurement-Induced Phase Transitions. *Physical Review Letters*, 130(22):220404, June 2023.

[143] Hong Liu and S. Josephine Suh. Entanglement tsunami: Universal scaling in holographic thermalization. *Phys. Rev. Lett.*, 112:011601, Jan 2014.

[144] Hong Liu and S. Josephine Suh. Entanglement Tsunami: Universal Scaling in Holographic Thermalization. *Physical Review Letters*, 112(1):011601, Jan 2014.

[145] Zi-Wen Liu, Seth Lloyd, Elton Zhu, and Huangjun Zhu. Entanglement, quantum randomness, and complexity beyond scrambling. *Journal of High Energy Physics*, 2018(7):41, Jul 2018.

[146] Tsung-Cheng Lu, Timothy H. Hsieh, and Tarun Grover. Detecting topological order at finite temperature using entanglement negativity. 2019.

[147] David J. Luitz, Nicolas Laflorencie, and Fabien Alet. Many-body localization edge in the random-field Heisenberg chain. *Physical Review B*, 91(8):081103, Feb 2015.

[148] Márk Mezei. Membrane theory of entanglement dynamics from holography. *Physical Review D*, 98(10):106025, Nov 2018.

[149] Márk Mezei and Douglas Stanford. On entanglement spreading in chaotic systems. *Journal of High Energy Physics*, 2017(5):65, May 2017.

[150] R. Moessner and S. L. Sondhi. Equilibration and order in quantum floquet matter. *Nature Physics*, 13:424 EP –, 04 2017.

[151] Martin Müller-Lennert, Fré déric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12), dec 2013.

[152] Adam Nahum, Sthitadhi Roy, Brian Skinner, and Jonathan Ruhman. Measurement and entanglement phase transitions in all-to-all quantum circuits, on quantum trees, and in

Landau-Ginsburg theory. *arXiv e-prints*, page arXiv:2009.11311, September 2020.

[153] Adam Nahum, Jonathan Ruhman, and David A. Huse. Dynamics of entanglement and transport in one-dimensional systems with quenched randomness. *Phys. Rev. B*, 98:035118, Jul 2018.

[154] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum entanglement growth under random unitary dynamics. *Physical Review X*, 7(3), jul 2017.

[155] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum entanglement growth under random unitary dynamics. *Phys. Rev. X*, 7:031016, Jul 2017.

[156] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum Entanglement Growth under Random Unitary Dynamics. *Physical Review X*, 7(3):031016, July 2017.

[157] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum Entanglement Growth under Random Unitary Dynamics. *Physical Review X*, 7(3):031016, July 2017.

[158] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator spreading in random unitary circuits. *Phys. Rev. X*, 8:021014, Apr 2018.

[159] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator Spreading in Random Unitary Circuits. *Physical Review X*, 8(2):021014, April 2018.

[160] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator Spreading in Random Unitary Circuits. *Physical Review X*, 8(2):021014, April 2018.

[161] Rahul Nandkishore and David A. Huse. Many-body localization and thermalization in quantum statistical mechanics. *Annual Review of Condensed Matter Physics*, 6(1):15–38, mar 2015.

[162] Rahul Nandkishore and David A. Huse. Many-Body Localization and Thermalization in Quantum Statistical Mechanics. *Annual Review of Condensed Matter Physics*, 6:15–38, Mar 2015.

[163] Laimei Nie, Masahiro Nozaki, Shinsei Ryu, and Mao Tian Tan. Signature of quantum chaos in operator entanglement in 2d CFTs. *arXiv e-prints*, page arXiv:1812.00013, Nov 2018.

[164] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[165] Simone Notarnicola, Andreas Elben, Thierry Lahaye, Antoine Browaeys, Simone Montangero, and Benoit Vermersch. A randomized measurement toolbox for rydberg quantum technologies, 2021.

[166] Ryan O'Donnell and John Wright. Efficient quantum tomography. *arXiv e-prints*, page arXiv:1508.01907, August 2015.

[167] M. Ohliger, V. Nesme, and J. Eisert. Efficient and feasible state tomography of quantum many-body systems. *New Journal of Physics*, 15(1):015024, January 2013.

[168] M. Ohliger, V. Nesme, and J. Eisert. Efficient and feasible state tomography of quantum many-body systems. *New Journal of Physics*, 15(1):015024, January 2013.

[169] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio. Generic entanglement can be generated efficiently. *Phys. Rev. Lett.*, 98:130502, Mar 2007.

[170] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117 – 158, 2014.

[171] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, oct 2014.

[172] Don N. Page. Average entropy of a subsystem. *Phys. Rev. Lett.*, 71:1291–1294, Aug 1993.

[173] Shriya Pai, Michael Pretko, and Rahul M. Nandkishore. Localization in Fractonic Random Circuits. *Physical Review X*, 9(2):021003, Apr 2019.

[174] Marco Paini and Amir Kalev. An approximate description of quantum states. *arXiv e-prints*, page arXiv:1910.10543, October 2019.

[175] Daniel E. Parker, Xiangyu Cao, Alexander Avdoshkin, Thomas Scaffidi, and Ehud Altman. A Universal Operator Growth Hypothesis. *arXiv e-prints*, page arXiv:1812.08657, Dec 2018.

[176] Daniel E. Parker, Xiangyu Cao, Alexander Avdoshkin, Thomas Scaffidi, and Ehud Altman. A Universal Operator Growth Hypothesis. *Physical Review X*, 9(4):041017, October 2019.

[177] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. Curran Associates Inc., Red Hook, NY, USA, 2019.

[178] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. Entropy, entanglement, and area: Analytical results for harmonic lattice systems. *Physical Review Letters*, 94(6), feb 2005.

[179] Hoi Chun Po, Lukasz Fidkowski, Takahiro Morimoto, Andrew C. Potter, and Ashvin Vishwanath. Chiral Floquet Phases of Many-Body Localized Bosons. *Physical Review X*, 6(4):041070, Oct 2016.

[180] Frank Pollmann, Ari M. Turner, Erez Berg, and Masaki Oshikawa. Entanglement spectrum of a topological phase in one dimension. *Physical Review B*, 81(6):064439, February 2010.

[181] Pedro Ponte, Anushya Chandran, Z. Papić, and Dmitry A. Abanin. Periodically driven ergodic and many-body localized quantum systems. *Annals of Physics*, 353:196 – 204, 2015.

[182] Andrew C. Potter and Romain Vasseur. Entanglement dynamics in hybrid quantum circuits. *arXiv e-prints*, page arXiv:2111.08018, November 2021.

[183] Tomaž Prosen and Iztok Pižorn. Operator space entanglement entropy in a transverse Ising chain. *Physical Review A*, 76(3):032316, September 2007.

[184] X.-L. Qi, Z. Yang, and Y.-Z. You. Holographic coherent states from random tensor networks. *Journal of High Energy Physics*, 8:60, August 2017.

[185] Xiao-Liang Qi, Emily J. Davis, Avikar Periwal, and Monika Schleier-Smith. Measuring operator size growth in quantum quench experiments. *arXiv e-prints*, page arXiv:1906.00524, Jun 2019.

[186] Xiao-Liang Qi, Emily J. Davis, Avikar Periwal, and Monika Schleier-Smith. Measuring operator size growth in quantum quench experiments. *arXiv e-prints*, page arXiv:1906.00524, June 2019.

[187] Tibor Rakovszky, Frank Pollmann, and C. W. von Keyserlingk. Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation. *Phys. Rev. X*, 8:031058, Sep 2018.

[188] Tibor Rakovszky, Frank Pollmann, and C. W. von Keyserlingk. Diffusive Hydrodynamics of Out-of-Time-Ordered Correlators with Charge Conservation. *Physical Review X*, 8(3):031058, July 2018.

[189] Tibor Rakovszky, C. W. von Keyserlingk, and Frank Pollmann. Entanglement growth after inhomogenous quenches. *arXiv e-prints*, page arXiv:1907.00869, Jul 2019.

[190] Michael Redhead. *Incompleteness, nonlocality, and realism: A prolegomenon to the*

243

philosophy of Quantum Mechanics. Clarendon Press, 2002.

[191] Daniel A. Rowlands and Austen Lamacraft. Noisy coupled qubits: Operator spreading and the fredrickson-andersen model. *Phys. Rev. B*, 98:195125, Nov 2018.

[192] Stefan H. Sack, Raimel A. Medina, Alexios A. Michailidis, Richard Kueng, and Maksym Serbyn. Avoiding barren plateaus using classical shadows. *PRX Quantum*, 3:020365, Jun 2022.

[193] M. Sanz, M. M. Wolf, D. Pérez-García, and J. I. Cirac. Matrix product states: Symmetries and two-body Hamiltonians. *Physical Review A*, 79(4):042308, April 2009.

[194] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326:96–192, January 2011.

[195] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, jan 2011.

[196] Alireza Seif, Ze-Pei Cian, Sisi Zhou, Senrui Chen, and Liang Jiang. Shadow Distillation: Quantum Error Mitigation with Classical Shadows for Near-Term Quantum Processors. *arXiv e-prints*, page arXiv:2203.07309, March 2022.

[197] Ramamurti Shankar. *Principles of quantum mechanics*. Plenum, New York, NY, 1980.

[198] Brian Skinner, Jonathan Ruhman, and Adam Nahum. Measurement-induced phase transitions in the dynamics of entanglement. *Physical Review X*, 9(3), jul 2019.

[199] Brian Skinner, Jonathan Ruhman, and Adam Nahum. Measurement-Induced Phase Transitions in the Dynamics of Entanglement. *Physical Review X*, 9(3):031009, Jul 2019.

[200] Brian Skinner, Jonathan Ruhman, and Adam Nahum. Measurement-Induced Phase Transitions in the Dynamics of Entanglement. *Physical Review X*, 9(3):031009, July 2019.

[201] Leonard Susskind, George Hrabovsky, and Art Friedman. *The Theoretical Minimum*. Basic Books, 2013.

[202] M. Szyniszewski, A. Romito, and H. Schomerus. Entanglement transition from variable-strength weak measurements. *Phys. Rev. B*, 100:064204, Aug 2019.

[203] Marcin Szyniszewski, Alessandro Romito, and Henning Schomerus. Entanglement transition from variable-strength weak measurements. *arXiv e-prints*, page arXiv:1903.05452, Mar 2019.

[204] Marcin Szyniszewski and Henning Schomerus. Random-matrix perspective on many-body

entanglement with a finite localization length, Jul 2020.

[205] Qicheng Tang and W. Zhu. Measurement-induced phase transition: A case study in the non-integrable model by density-matrix renormalization group calculations. *arXiv e-prints*, page arXiv:1908.11253, Aug 2019.

[206] Romain Vasseur and Joel E. Moore. Nonequilibrium quantum dynamics and transport: from integrability to many-body localization. *Journal of Statistical Mechanics: Theory and Experiment*, 6:064010, Jun 2016.

[207] Romain Vasseur, Andrew C. Potter, Yi-Zhuang You, and Andreas W. W. Ludwig. Entanglement Transitions from Holographic Random Tensor Networks. *arXiv e-prints*, page arXiv:1807.07082, Jul 2018.

[208] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Phys. Rev. Lett.*, 93:207204, Nov 2004.

[209] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, Mar 2008.

[210] Guifré Vidal. Efficient simulation of one-dimensional quantum many-body systems. *Phys. Rev. Lett.*, 93:040502, Jul 2004.

[211] Lev Vidmar and Marcos Rigol. Entanglement entropy of eigenstates of quantum chaotic hamiltonians. *Physical Review Letters*, 119(22), nov 2017.

[212] Sagar Vijay. Measurement-driven phase transition within a volume-law entangled phase. 2020.

[213] Sagar Vijay and Ashvin Vishwanath. Finite-Temperature Scrambling of a Random Hamiltonian. *arXiv.org*, March 2018.

[214] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. Operator hydrodynamics, otocs, and entanglement growth in systems without conservation laws. *Phys. Rev. X*, 8:021013, Apr 2018.

[215] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws. *Physical Review X*, 8(2):021013, April 2018.

[216] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws.

Physical Review X, 8(2), apr 2018.

[217] David Wallace. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford, GB: Oxford University Press, 2012.

[218] Xiaoguang Wang, Shohini Ghose, Barry C. Sanders, and Bambi Hu. Entanglement as a signature of quantum chaos. *Phys. Rev. E*, 70:016217, Jul 2004.

[219] Don Weingarten. Asymptotic behavior of group integrals in the limit of infinite rank. *Journal of Mathematical Physics*, 19(5):999–1001, 1978.

[220] Zack Weinstein, Shane P. Kelly, Jamir Marino, and Ehud Altman. Scrambling Transition in a Radiative Random Unitary Circuit. *arXiv e-prints*, page arXiv:2210.14242, October 2022.

[221] Xiao-Gang Wen. Topological order: From long-range entangled quantum matter to a unified origin of light and electrons. *ISRN Condensed Matter Physics*, 2013:1–20, mar 2013.

[222] Eugene P. Wigner. Random matrices in physics. *SIAM Review*, 9(1):1–23, 1967.

[223] Kai-Hsin Wu, Tsung-Cheng Lu, Chia-Min Chung, Ying-Jer Kao, and Tarun Grover. Entanglement renyi negativity across a finite temperature transition: a monte carlo study. 2019.

[224] Shenglong Xu and Brian Swingle. Locality, Quantum Fluctuations, and Scrambling. *arXiv e-prints*, page arXiv:1805.05376, May 2018.

[225] Shenglong Xu and Brian Swingle. Locality, Quantum Fluctuations, and Scrambling. *Physical Review X*, 9(3):031048, July 2019.

[226] Shenglong Xu and Brian Swingle. Locality, Quantum Fluctuations, and Scrambling. *Physical Review X*, 9(3):031048, July 2019.

[227] Zhi-Cheng Yang, Fangli Liu, Alexey V. Gorshkov, and Thomas Iadecola. Hilbert-space fragmentation from strict confinement. *Phys. Rev. Lett.*, 124:207602, May 2020.

[228] Yi-Zhuang You and Yingfei Gu. Entanglement features of random Hamiltonian dynamics. *Physical Review B*, 98(1):014309, Jul 2018.

[229] Yi-Zhuang You and Yingfei Gu. Entanglement features of random Hamiltonian dynamics. *Physical Review B*, 98(1):014309, July 2018.

[230] Yi-Zhuang You, Zhao Yang, and Xiao-Liang Qi. Machine learning spatial geometry from

entanglement features. *Phys. Rev. B*, 97:045153, Jan 2018.

[231] Yi-Zhuang You, Zhao Yang, and Xiao-Liang Qi. Machine learning spatial geometry from entanglement features. *Physical Review B*, 97(4):045153, February 2018.

[232] Aidan Zabalo, Michael J. Gullans, Justin H. Wilson, Sarang Gopalakrishnan, David A. Huse, and J. H. Pixley. Critical properties of the measurement-induced transition in random quantum circuits. *Physical Review B*, 101(6):060301, February 2020.

[233] Andrew Zhao, Nicholas C. Rubin, and Akimasa Miyake. Fermionic partial tomography via classical shadows. *Phys. Rev. Lett.*, 127:110504, Sep 2021.

[234] Tianci Zhou and Xiao Chen. Operator dynamics in a Brownian quantum circuit. *Physical Review E*, 99(5):052212, May 2019.

[235] Tianci Zhou and Xiao Chen. Operator dynamics in a Brownian quantum circuit. *Physical Review E*, 99(5):052212, May 2019.

[236] Tianci Zhou and Xiao Chen. Operator dynamics in a Brownian quantum circuit. *Physical Review E*, 99(5):052212, May 2019.

[237] Tianci Zhou and David J. Luitz. Operator entanglement entropy of the time evolution operator in chaotic systems. *Physical Review B*, 95(9):094206, Mar 2017.

[238] Tianci Zhou and Adam Nahum. Emergent statistical mechanics of entanglement in random unitary circuits. *arXiv e-prints*, page arXiv:1804.09737, Apr 2018.

[239] Tianci Zhou and Adam Nahum. Emergent statistical mechanics of entanglement in random unitary circuits. *Phys. Rev. B*, 99:174205, May 2019.

[240] Tianci Zhou and Adam Nahum. Emergent statistical mechanics of entanglement in random unitary circuits. *Physical Review B*, 99(17):174205, May 2019.

[241] Tianci Zhou and Adam Nahum. Emergent statistical mechanics of entanglement in random unitary circuits. *Physical Review B*, 99(17):174205, May 2019.

[242] Marko Znidaric. Exact convergence times for generation of random bipartite entanglement. *Phys. Rev. A*, 78:032324, Sep 2008.

[243] Marko Žnidarič. Entanglement growth in diffusive systems. *Communications Physics*, 3(1), jun 2020.

[244] Michael Zwolak and Guifré Vidal. Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm. *Phys. Rev.*

Lett., 93:207205, Nov 2004.

[245] Karol Życzkowski, Paweł Horodecki, Michał Horodecki, and Ryszard Horodecki. Dynamics of quantum entanglement. *Physical Review A*, 65(1):012101, Jan 2002.

[246] Karol Ż yczkowski, Paweł Horodecki, Anna Sanpera, and Maciej Lewenstein. Volume of the set of separable states. *Physical Review A*, 58(2):883–892, aug 1998.