Measured Quantum-State Stochastic Processes

By

ARIADNA ELENA VENEGAS-LI

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

James P. Crutchfield

John B. Rundle

Richard T. Scalettar

Committee in Charge

2023

To my parents, Sandra and Walley.

Contents

Measured Quantum-State Stochastic Processes

Abstract

A century of concerted effort brought new levels of physical understanding and engineering capability to quantum physics that promise major advances in fundamental theory and technology applications. A case in point, quantum computation has the potential to become one of humanity's most innovative and disruptive technologies. Advances there are pushing physics to probe quantum phenomena in systems that are increasingly complex, and in more detailed ways than ever before. It is now clearer than ever that further progress will require constructively working with noise, error, and environmental interactions. A key part of this is to understand the properties of time series of quantum states emitted by a quantum system—the main subject of this work.

This dissertation sets out to study time-series of quantum states emitted by a quantum system. The fundamental objects in this are Quantum-State Stochastic Processes (QSSPs)—sequences of stochastically generated quantum states. In particular, the focus is on the interaction of a classical observer—via quantum measurement—with these objects, and the informational and statistical characterization of the classical stochastic processes that result from that interaction.

As a classical observer uses a measurement protocol to observe a Quantum-State Stochastic Process, the outcomes form a time series. Individual time series are realizations of a stochastic process over the measurements' classical outcomes. This dissertation studies the dependence of that stochastic process of measurement outcomes on both the QSSP and the measurement protocol. In particular, it demonstrates that regardless of the measurement protocol—and for several specific protocols explicitly discussed in this work—the output classical stochastic process is generically highly complex in two specific senses: (i) it is generically unpredictable, to a degree that depends on the measurement choice, and (ii) achieving optimal prediction for these stochastic processes will generically require an infinite amount of memory.

Inspired by the study of classical stochastic processes, this dissertation uses and adapts the theory of *computational mechanics* and hidden Markov models to understand and categorize these measured stochastic processes. Specifically, we identify the mechanism underlying their complicatedness as generator nonunifilarity—the degeneracy between sequences of generator states and sequences

of measurement outcomes. This makes it possible to quantitatively explore the influence that measurement choice has on a quantum process' degrees of randomness and structural complexity using recently introduced methods from ergodic theory. This dissertation provides the explicit metrics and associated algorithms for that quantification. It demonstrates that under certain conditions, dependence of these metrics on the choice of measurements is smooth. In addition, these metrics are used to design informationally-optimal measurements of time-series of quantum states.

Most approaches to quantum stochastic dynamics focus on the evolution of the quantum state of a particular system—for example, quantum collision models. In contrast, the results here lay the groundwork for the study of time series of quantum states emitted by a quantum system, which have so far been largely unexplored. Additionally, they provide a first description of what classical interaction with these time series yields.

Acknowledgments

I am immensely grateful to all the people that have made this long and surprising journey possible. I would like to first thank my advisor, Jim Crutchfield, for his support, guidance, and vision. In times of self-doubt Jim believed in me and pushed me to go on, but he also allowed me to do things in my own time. His curiosity inspires me and his logical rigor has shown me how to be a better scientist. He has also provided me with excellent tips of places to see in California.

Thanks to Jim, I also got to be part of a stimulating research community that has helped shaped the way I think. In particular I want to thank: David Gier, for being an awesome and chaotic research partner and friend whose humor I can always count on. Jacob H. for his dark-humored friendship, honesty, and allways having cool facts to share. Fabio Anza, for being a great mentor and sharing with me his joy and fearlessness of asking big questions. Jinghao Liu, for our deep conversations about life, random messages and weird but excellent humor. Dany Masante, for being a great travel partner and singing Shakira oldies with me. The CSC is filled with awesome people, and I also want to thank Jeff, Greg, Adam R., Komi, Alex, Kyle, Chris, Mikhael, Adam K., Sam, Alec, Korana, Cina, John M., Ryan and all the other present and past members of the group that have made my time there enjoyable and helped me grow as a researcher. Thank you also to Lenna Crabbe-Charlesworth for her caring and support during this last year.

I next thank my family, who are at the core of who I am and what I do. Sandra, my mother and friend, who has shown me the importance of connecting with other humans and connecting with myself. Walley, my father and support, who has shown me the value of working for what you want and enjoying what you like. Thank you both for your boundless love, selflessness, and for instilling in me curiosity, confidence, compassion, and encouraging me to pursue what I want. Ruben, my brother, for showing me the way to the world out there, and for being a loving and guiding hand even when life keeps us physically far. Lara, for being a shiny little ball of joy and wonder. And to Tia Tey, Tio Pedro, Tia Gis, Tia Marlen, Tio Popo, Meli, Mo, Sebas, Dianita, Antonio, Kattia, and the rest of my family for their unconditional love and support.

Another essential part of my life are my friends, who have brought me joy and have chosen to share their time with me, without them I would not have made it through a Ph.D.. My longest time friends: Manfred, Mario and Brian, it has been a pleasure to grow alongside you, thank you for being

my *los de siempre*. Thank you also to my friends from UCR with whom I share the memories of a tropical college life and continue making memories: Felipe, Bojo, Berni, Flavio, Joshua, Totolate, Dennisse, Diana, Oscar, Cesar and Juan Pablo, who doesn't quite fit in that category but is an ever-reliable source of good humour and almost always correct opinions. Marcela, thank you for becoming a little piece of home and family. Thank you to my Davis friends for the good times and adventures. In particular Nastya, Jessie, Victoria, Megan, Jon, Abel, Arian, Sasha, Chase, Imran, Pao, Anna, Ju, Cameron, Rose, Jenna, Chris and James: all the trips, dinners, movies, happy hours, work-hangs, jokes and many ice creams kept me sane and connected.

Last but not least, thank you Jacob Lane for your companionship, encouragement, friendship and humour, for almost always knowing when to be the voice of reason and when to be the voice of crazy, for being a great partner in adventure and in crisis, and for your love.

CHAPTER 1

Introduction

1.1. Motivation

Observation and control are at the basis of our interactions with the physical world. When observing complex systems, time series of measurements have proven a useful tool to building models and gaining insight into the behavior of such systems. For instance, one can use a time series to find patterns and reconstruct the causal dynamical structure of the underlying process that generates the time series [1, 2].

Time series of controlled quantum states are essential to quantum physics, quantum information and computing, and their implementations in novel technologies. Moreover, as quantum technologies scale to larger qubit collections that evolve coherently over increasingly longer times, fault-tolerant design and error correction of quantum-state time series become increasingly necessary.

Several fault-tolerant systems and diagnostic tools have been developed assuming quantum processes with uncorrelated noise [3–7]. Unfortunately, progressing beyond those assumptions to more physically realistic non-Markovian, correlated processes has been challenging. To date, error correction for non-Markovian quantum processes can be deployed only in specific cases. Moreover, contemporary theory offers a restricted toolset for quantum process identification and control [8–10].

The following explores one reason for these challenges and limited progress. In short, there is substantially more complicated statistics and correlational structure embedded in non-Markovian quantum processes and in the classical stochastic processes that result from measurement than currently appreciated. We show how to identify the signatures of these complexities and how to constructively address the challenges they pose.

To ameliorate historical inconsistencies, along the way we give unified definitions of Markov and non-Markov processes. Said most simply, these address the role of memory in a structurally consistent way—a way that also gives access to modern multivariate information theory. Clarity in this is

essential to appreciating the structural varieties of complex quantum processes. Lasting progress in complex quantum processes depends on this clarity.

Beyond their direct relevance in quantum information tasks, time series of quantum states are output signatures of the dynamics of many physical processes. For instance, the conformational dynamics of biological macromolecules can be probed using single-molecule fluorescence resonance energy transfer (smFRET) [11, 12]. While the specific techniques vary, the general procedure in these experiments is that the macromolecules are exposed to a laser beam, are excited and emit photons. The characteristics of these photons, when detected, provide some information about the conformation of the molecule at the time of probing. As the experimental capabilities improve, the techniques for analysis of these time series of photons are still developing. Classical hidden Markov model (HMM) techniques have been used to study these systems [13, 14]. We posit that the techniques developed here will serve to more thoroughly explore these physical dynamics.

Explicating the structures embedded in quantum processes requires stepping back, to revisit a basic question: How to characterize the stochastic process that results from measuring sequential quantum states? The answer is found in a recently-introduced framework for correlated and state-dependent quantum processes [15, 16]. Notably, its toolkit relies only on classical dynamical systems. The turn to the latter is perhaps unexpected from the perspective of quantum physics, but makes sense given that the goal is to describe the classical data an experimentalist has in hand in the laboratory. Specifically, the tools rely on fundamental results that both reach back more than half a century to early ergodic and stochastic process theories, but also call on contemporary mathematics of dynamical systems—specifically, iterated function systems and their stable asymptotic invariant measures [17–19].

Thus, the objects of study here are time series of sequential quantum systems and the stochastic processes that result from sequentially measuring the state of each one.

Generating a time series of sequential quantum states usually occurs under the control of an experimental apparatus. If control over the apparatus is not perfect or if it undergoes dynamics that are unstable or not fully understood, then the time series of emitted quantum states can be profitably regarded as a stochastic process, with the stochasticity encapsulating the environmental or

dynamic effects that cannot be analytically captured. It may also be desired for a given application, such as quantum cryptography, that a quantum state process be stochastic.

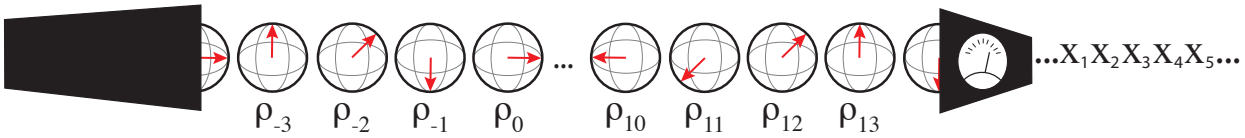


FIGURE 1.1. A general controlled quantum source (CQS) as a discrete-time quantum dynamical system (black box) that stochastically generates a time series of quantum states $\ldots \rho_{t-2}\ \rho_{t-1}\ \rho_t \ldots$ (density matrices). Measuring each state in the sequence realizes a classical stochastic process over random variables $\ldots X_1 X_2 X_3 X_4 X_5 \ldots$.

Figure 1.1 illustrates this with a black box quantum system that emits a quantum state ρ_t (density matrix) at each time t. We refer to this source as a *controlled quantum source* (CQS) and to its output as a *quantum-state stochastic process* (QSSP). It is important to note here that a distinct quantum state is emitted at each timestep and the object of study is the time series of these states. To emphasize, we are not investigating the dynamical evolution of an individual quantum state or individual quantum system.

Continuing from left to right in Figure 1.1, measurement of a quantum-state time series results in a stochastic process of classical measurement outcomes. Of that classical process we then ask:

- *Statistics*: What are the properties of this observed classical process—its randomness, correlation, memory?
- *System identification*: What properties of the underlying quantum stochastic generator hidden in the black box can be reconstructed from the observed classical process?
- *Measurement choice*: How does measurement affect the observed statistics and controller identification?

Computational mechanics [**20**–**22**] was originally introduced to constructively answer these questions for purely classical hidden processes. To do this, it extracts the "effective theory" from a time series of observations and provides measures of the time series' randomness and structure. In particular, a process' *statistical complexity* C_μ quantifies how much structure or memory is required to do optimal prediction. And, the *entropy rate* h_μ quantifies a process' intrinsic randomness—the rate of information production.

This toolset, together with the fact that measuring a quantum time series results in a classical time series, motivates our approach. We adapt these classical measures of intrinsic randomness and structure to describe the classical time series observed when applying measurement operators to a time series of quantum states. This provides a description of relevant properties of a stochastic process of quantum states—properties that have proven usefully diagnostic and descriptive in the classical setting. The following endeavors to show that they are also usefully diagnostic and descriptive for quantum processes. Moreover, our approach allows analyzing the effects that measurement choice has on the observed complexity of a quantum dynamical process.

This serves as a starting point to more fully appreciating the complicatedness of quantum dynamical processes and the role that measurement plays. Over the longer term, building on this, the goal is to characterize stochastic quantum processes beyond being Markovian or non-Markovian—memoryless or memoryful—to arrive at understanding of their informational and statistical properties. These metrics can then support more informed approaches to error correction and to potentially leveraging noise for particular tasks, just as their classical analogs have for thermodynamic computing [**23**, **24**].

1.2. Dissertation Outline

To set this work in the context of other approaches to stochastic dynamics and open quantum systems, Chapter 2 provides a brief overview of similar efforts in the field and makes a clear distinction between those efforts and the developments presented here. Once the direction of this work is put in context, Chapter 3 presents a summary of the classical physics background necessary to analyze the stochastic processes that result from measuring a quantum time series.

With both the context and the background tools available, Chapter 4 contains a formal definition for the main object of study, the Quantum-State Stochastic Processes. It then goes on define in detail the characteristics of the QSSPs that are considered in this dissertation: stationary, ergodic and separable QSSPs. This chapter also defines a particular model of QSSP controlled source—the classically-controlled quantum source—and gives a nod to the quantum-infromation quantification of QSSPs.

Chapter 5 then discusses and defines possible measurement protocols for QSSPs and focuses on a particular example. It then develops a constructive way to generate measured stochastic process

presentations in the form of hidden Markov models and discusses their general properties. In particular, it is argued and demonstrated that these models will generically lead to highly complex processes, and the specific ways in which this is true are made explicit by introducing metrics of information generation and storage in these stochastic processes. Using these tools we introduce three example processes that serve to illustrate the three types of observable behavior when measuring a QSSP generated by a classically-controlled quantum source.

Having introduced QSSPs, their measurement protocols, and metrics to quantify and classify measured QSSPs, Chapter 6 focuses on understanding how these metrics depend on the choice of measurement. By working through typical-case examples, it is shown that both the intrinsic randomness and the memory structure of the measured stochastic processes vary smoothly with respect to measurement parameters. The general characteristics and behaviors are discussed, and notions of informationally-optimal measurements are introduced.

The conclusions for this work and future directions are presented in Chapter 7.

CHAPTER 2

Alternative Approaches to Quantum Processes

To start, let's locate our development in the context of similar but distinct approaches to the study of open quantum systems, dynamics and processes.

The study of quantum stochastic processes dates back to the 1960s and 70s [25–29]. Despite the long tradition, the field has been riddled with multiple foundational challenges that have stalled its progress. That said, the last two decades have seen an increase in concerted efforts to advance our understanding of the field, which still has a long way to go. Perhaps the most common setting for quantum processes investigates a single quantum system evolving in time. In this, stochasticity in the system's state evolution arises from its interaction with an environment. Within this setting stochasticity in temporal evolution can also arise from inherent nonlinear dynamics or repeated measurement and other state mappings [30, 31].

Historically, though, there is a longstanding effort to characterize stochasticity in quantum dynamics as a means to manage quantum noise [32, 33]. Much of the machinery developed to describe quantum stochastic phenomena arose from open quantum systems, seeded around quantum master equations and relying heavily on assuming process Markovity [34, 35]. More recently, an effort emerged to understand, detect, and quantify non-Markovity, and many examples of specific non-Markovian quantum dynamics have been analyzed in detail [8, 9, 36–40].

From the perspective of Markov's original concept of "complex chains" [41, 42] and the modern theory of (classical) Markov processes, though, these approaches rely on an unnecessarily-varied set of Markovity definitions. One consequence is that, most basically, they do not always agree on what process memory and correlation are. This also makes comparing results across distinct investigations and approaches challenging. To address this, we introduce a unified definition of memory that we apply to classical and measured quantum processes. Section 3.2 discusses this in more detail.

Directly importing the concept of classical stochastic processes to the quantum domain has proved challenging, as References [43, 44] discuss in detail. While there are several causes, the primary difficulty is that the Kolmogorov Extension Theorem—specifying event-sequence probabilities and measures—breaks down when the random variables are quantal and are generated by sequential quantum mechanisms.

Recently, process tensors were introduced to treat such quantum stochastic processes [43, 45–48]. They deliver a gateway to probe quantum stochasticity beyond the binary distinction of Markov versus non-Markov. Modeling correlations beyond merely Markovian will allow analyzing a broader class of quantum processes. That said, the endeavor is new.

While there are parallels with process-tensor representations, the following focuses on a different type of dynamical process—different also from individual quantum-system temporal evolution. It considers a sequence of distinct entities, in which the quantum state of each is a random variable at each time step. In this, there is a rough parallel to quantum spin chains. Notably, these variables can be correlated by the physical mechanism that sequentially generates them. It distinguishes itself from the previously mentioned settings, in which the sequential evolution of a single physical entity is tracked, and the random variable at each timestep is the quantum state of a single physical system.

Consider a physical system that emits qubits, each in a state that is noisy or stochastic; for instance, photons emanating from a blinking quantum dot [49, 50], or any other single-photon source, a technology that is actively being developed for technological applications [51–55]. The difference between these processes and the more-familiar evolving single-system dynamics is that, at each time step, the quantum state of the emitted qubit is its own random variable. That is, each successive random variable lives in its own Hilbert space. The associated random variable takes on a specific qubit state, and measuring these states does not interfere with the quantum process generator. In short, as the quantum source emits quantum states the dimension of the product Hilbert space grows. As a consequence, addressing time-asymptotic properties requires working with an infinite product state space. These changes the kind of investigation we pursue. We investigate time series of quantum states emitted by a physical system analyzing the multivariate statistical and

informational properties of the system's dynamics, in a similar spirit to the classical development in Reference [56].

Specifically, References [15] and [16]—which form the backbone of this dissertation—recently introduced and developed a framework for such Quantum-State Stochastic Processes. There, the problem of quantum processes is cast in a way that is amenable to directly applying the tools of classical stochastic processes to characterize the informational and structural properties of Quantum-State Stochastic Processes. It was shown that a qubit time series, when observed through projective measurements, generically results in a highly complex classical stochastic process. *Highly complex* here means that the observed process has positive entropy rate and requires an infinite number of temporal features to optimally predict future outcomes. It was also demonstrated that measurement choice—the manner in which an observer interacts with a qubit stream—can drive a quantum process to appear more or less complex.

To accomplish this, we adapted the metrics of computational mechanics to describe randomness and structure in the measured processes. The following Chapter introduces the concepts necessary for this adaptation, which will then be used to study measured QSSPs in the following chapters.

CHAPTER 3

Classical Background

This chapter introduces the classical physics required to describe and characterize measured quantum state stochastic processes.

3.1. Classical Stochastic Processes

A classical stochastic process $\overleftrightarrow{X}$ is a bi-infinite series of indexed observables produced by a system and is defined by the probability measure over the random variables corresponding to the observables: $\Pr(\ldots X_{t-1}, X_t, X_{t+1} \ldots = \ldots x_{t-1}, x_t, x_{t+1} \ldots)$. In this, the random variables corresponding to the observables are denoted with capital letters $\ldots X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2} \ldots$ and their realizations are denoted by lowercase letters $\ldots x_{t-2}, x_{t-1}, x_t, x_{t+1}, x_{t+2} \ldots$. The x_t values are drawn from a discrete alphabet $\mathcal{A}$. The label t in the indexing is chosen to evoke the traditional time-indexing of stochastic processes in which the random variables are sequential measurements of a physical system. Although more general stochastic processes exist, throughout this work we will limit our study to *discrete time stochastic processes*—those in which the domain of the indexing variable is discrete.

Considering families of finite distributions proves useful in the study of stochastic processes. That is, distributions of the type $\Pr(X_{t:t+l})$ over a contiguous block of outcomes of the form $x_{t:t+l} \in \mathcal{A}^l$. Explicitly, we denote blocks of contiguous random variables of length l as $X_{t:t+l} = X_t, X_{t+1}, \ldots X_{t+l-1}$, and blocks of contiguous outcomes as $x_{t:t+l} = x_t, x_{t+1}, \ldots x_{t+l-1}$. In both cases the left index is inclusive, and the left index is exclusive.

We restrict our focus to studying stationary stochastic processes.

DEFINITION 1. *A stochastic process is said to be* stationary *if its probability measure is invariant over time(or index)-shifts:*

$$\Pr(X_{t:t+l} = x_{t:t+l}) = \Pr(X_{0:l} = x_{0:l}) \ , \tag{3.1}$$

for all t and l. That is, the joint distribution for blocks of length l is time (or index)-translation invariant.

Independent identically distributed (I.I.D.) processes provide the simplest example of stationary stochastic processes. In these, the random variables X_t are drawn from the same probability distribtion over outcomes $x_t \in \mathcal{A}$ and are independent of each other—that is, the outcome of random variable at time t is not affected by the outcomes of any other random variable. For this reason, we say IID processes are *memoryless*.

Until the early twentieth century, most stochastic processes were considered IID processes [57]. In 1907 Andrei A. Markov introduced the concept of "complex chains" [41, 42, 58, 59], and with it opened the door to the study of memoryful, structured processes with temporal correlations. Markov processes are the simplest class of stochastic process that have memory and temporal correlations. And they provide a basis to study more complex memoryful stochastic processes.

3.2. Process Markovity

In the most common definition, a *Markov process* or *Markov chain* is a stochastic process in which the outcome of a random variable depends exclusively in the outcome of the previous random variable and is independent of all other past outcomes. Mathematically, that is:

$$\Pr(X_t = x_t | X_{-\infty:t} = x_{-\infty:t}) = \Pr(X_t = x_t | X_{t-1} = x_{t-1}) \tag{3.2}$$

This can straightforwardly be extended to include any finite number of previous random variables:

DEFINITION 2. *A* Markov process *or* Markov chain *of order R is a stationary stochastic process* $\mathbf{X}$ *in which the probability distribution satisfies the following:*

$$\begin{aligned} &\Pr(X_t = x_t | X_{-\infty:t} = x_{-\infty:t}) \\ &\quad = \Pr(X_t = x_t | X_{t-R:t} = x_{t-R:t}) \\ &\quad = \Pr(X_t = x_t | X_{t-1} = x_{t-1} \dots X_{t-R} = x_{t-R}), \end{aligned} \tag{3.3}$$

for all $t \in \mathbb{Z}$ and $R \in \mathbb{N}$.

That is, the probability distribution of a particular random variable conditioned on the past depends only on the values of the previous R random variables. The number R is called the *Markov order* of the process. Another way to phrase this is that, for Markov processes, R is the answer to the question: how far in the past do I need to look in order to make an optimal prediction of what the next outcome is going to be?

From this definition, we emphasize the following:

- Memoryless or independent identically distributed (I.I.D.) processes are stochastic processes with $R = 0$. We also refer to them as Markov processes of order 0.
- Stochastic processes with $1 \leq R < \infty$ are memoryful. We refer to them as Markov processes of order R.
- Memoryful stochastic processes that do not satisfy the Markov condition in Equation (3.3) for finite R, but require R to be infinite are infinite-order Markov processes. Perhaps a bit surprisingly, but the latter are quite common [**60**]. These are the best candidates for the descriptor "non-Markov processes".

The nomenclature above is directly derived from the definition of Markov property, consistent with A. A. Markov's original motivations to study "complex chains", his phrase for memoryful stochastic processes. Memoryless is $R = 0$, Markov is $1 \leq R < \infty$, and infinite Markov or non-Markov is $R = \infty$.

A generalization of Markov Processes that allows for the study of non-Markov processes was introduced later in the twentieth century as *functions of Markov chains*. These processes are now

known also by the names of Hidden Markov Processes (HMPs) or Hidden Markov Chains (HMCs). A wide class of stationary stochastic processes can be modeled with HMPs [61–63].

3.3. Hidden Markov Processes

Hidden Markov Processes generalize Markov processes in the following way: they consist of an internal Markov process that is observed by a funciton of its internal-state sequence. Directly dealing with probability measures over bi-infinite sequences of random variables is cumbersome, more so when dealing with stochastic processes with more structure than Markov processes. For this reason, we instead use finite-state models that are capable of generating the same probability measure over sequences of outcomes of any length as those of the stochastic process they are modeling.

Definition 3 (Hidden Markov Model). *A* hidden Markov model *(HMM) is a tuple* $(\boldsymbol{\mathcal{S}}, \mathcal{A}, \{T^x\})$ *that consists of:*

(1) a set $\boldsymbol{\mathcal{S}}$ *of hidden (or internal) states* $\sigma \in \boldsymbol{\mathcal{S}}$.

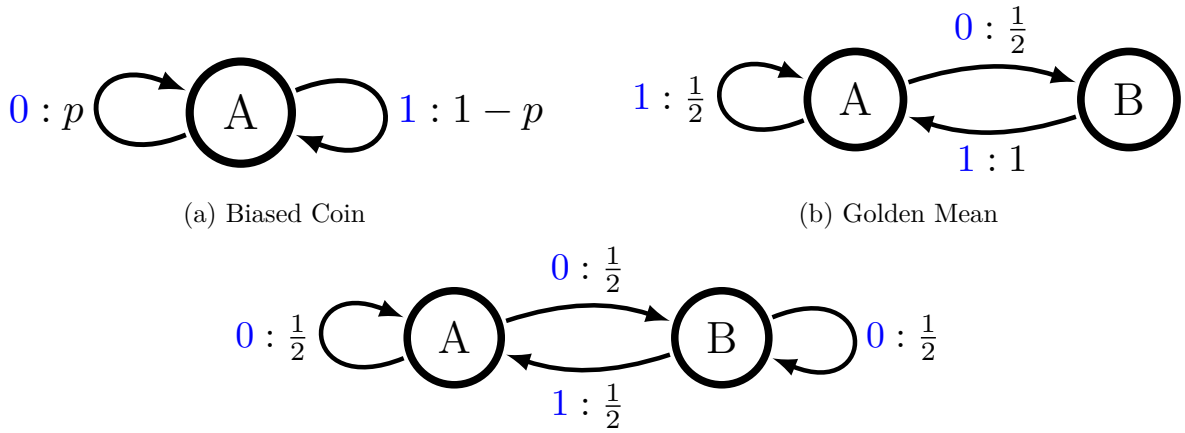
(2) a discrete alphabet $\mathcal{A}$*: a set of symbols (or observables) that the HMP emits on state-to-state transitions at each time step.*

(3) $\{T^{(x)}\}$, $x \in \mathcal{A}$ *a set of symbol-labeled transition matrices such that* $T^{(x)}_{\sigma\sigma'} = \Pr(x, \sigma'|\sigma)$ *with* $\sigma, \sigma' \in \boldsymbol{\mathcal{S}}$.

An HMM directly defines the dynamic over hidden states, $\ldots \mathcal{S}_{t-1}\mathcal{S}_t\mathcal{S}_{t+1} \ldots$, which is itself Markovian (order $R = 1$). The transition matrix $T = \sum_{x\in\mathcal{A}} T^x$ gives the probability of transitioning from one internal state to another at each time-step $T_{\sigma\sigma'} = \Pr(\mathcal{S}_t = \sigma|\mathcal{S}_{t-1} = \sigma')$. This internal dynamic is not taken to be observable, but part of the internal structure that defines the dynamic on the sequence of symbols $\ldots X_{t-1}X_tX_{t+1} \ldots$, which constitute the stochastic process that the HMM generates.

T also defines a stationary state distribution π over hidden states, such that $\pi \cdot T = \pi$. That is, π is a row vector such that $\pi_\sigma = \Pr(\sigma)$ with $\sigma \in \boldsymbol{\mathcal{S}}$.

Figure 3.1 shows three examples of HMMs. The three HMMs have the same alphabet $\mathcal{A} = \{0, 1\}$. The HMM in Figure 3.1a has one hidden state, $\boldsymbol{\mathcal{S}} = \{A\}$ and generates an IID process with bias p. The HMMs in Figure 3.1b and Figure 3.1c have two hidden states, $\boldsymbol{\mathcal{S}} = \{A, B\}$. The outgoing edges



(a) Biased Coin

(b) Golden Mean

(c) Simple Nonunifilar Source

FIGURE 3.1. Examples of HMMs for three different stochastic processes.

of each hidden state represent the symbols that can be observed or emitted by the process when the model is in that hidden state. An edge starting at state σ and ending at state σ' is labeled as $x : \Pr(x, \sigma'|\sigma)$ and corresponds with the labeled-transition matrix entry $T^{(x)}_{\sigma\sigma'}$. Edges that correspond to transitions with zero probability are omitted. At each timestep an edge is taken and a different symbol is emitted. For instance, following the HMM on Figure 3.1b: if the model is in state A it generates a 1 with probability 0.5 and remains in state A, or generates a 0 with probability 0.5 and transition to state B.

An HMM then is a model for a stochastic process consisting of the set of emitted symbol sequences and their associated probabilities. It is important to note here that, even if the dynamic over hidden states is Markovian, HMMs generate a more general class of stochastic processes that includes non-Markovian processes. While the hidden states give an explicit mechanism for producing a stochastic process, the stochastic process itself is defined only over the set of symbols $x \in \mathcal{A}$. On occasion, we refer to HMMs as *machines*, in the sense of *state machines*.

HMMs are useful in that they specify a finite mechanistic procedure to produce the correct probabilities for a stochastic process (specifically a HMP). For each HMP there is an infinite family of distinct HMMs that can model it. The simplest way to see this is to consider duplicating each hidden state with identical transitions, this would exactly preserve the statistics of the symbol

sequences and, thus, the stochastic process. The HMMs that generate a stochastic process are a process' *presentations*.

In point of fact, the development here works with *edge-labeled* HMMs. There are also *state-labeled* HMMs that emit symbols on entering a state. Both presentation classes generate the same class of stochastic processes. They are equivalent in this sense and they can be directly interconverted. Edge-labeled presentations, though, do offer computational advantages when calculating informational properties and in interpreting the functionality of their operation.

That said, not all presentations of a process are created equal. Some, for example, have a deceptively simple-looking structural property called unifilarity.

3.3.1. Presentation Unifilarity.

Definition 4 (Unifilarity). *An HMM transition is unifilar if the current internal state σ and the emitted symbol x uniquely determine the next internal state σ':* $\Pr(\sigma'|x,\sigma) = 1$*, for at most one x. That is, there is at most one transition leaving a state for each symbol.*

An HMM is unifilar *if all its transitions are.*

When an HMM is *nonunifilar* there is ambiguity in the next state for at least one transition. For example, in Figure 3.1, the HMMs in Figure 3.1a and 3.1b are unifilar, since given a hidden state and the observed symbol, there is no ambiguity in what the next hidden state is. On the other hand, the HMM in Figure 3.1c is nonunifilar: if the model is in hidden state A, and the next observed symbol is a 0, the next hidden state could be A or B with equal probability.

It is a notable fact—one motivating the distinction in the first place—that processes generated by finite unifilar HMMs are less (typically much less) complex than those that can be generated only by finite nonunifilar HMMs.

An intuitive way to see why this occurs is the following. Consider a realization of a given process. If it was emitted by a unifilar HMM the realization has a one-to-one or at most one-to-finite correspondence between observed symbols and sequences of hidden-state transitions. In contrast, a realization generated by a nonunifilar HMMs has a one-to-infinite correspondence between observed symbols and hidden state transitions. The result in this case is that the number of possible of sequences of hidden states that emit a particular sequence of observed symbols grows exponentially

with sequence length. In general, a significantly more complex hidden structure is required to optimally predict processes generated by nonunifilar HMMs.

3.3.2. Predictors and Generators. A stochastic process' unifilar presentation, up to redundancies in states or transitions, is an optimal predictor. That is, given an HMM's hidden state, the probabilities of the next observed symbols are the optimal, most informed prediction of what that next observed symbol will be.

Unifilar HMMs being process predictors contrasts with nonunifilar HMMs which are not predictors. The latter are only generators of process realizations. Moreover, their states are typically poor predictors.

One way to restate the distinction between process predictors (unifilar presentations) and process generators (nonunifilar presentations) is the following. On the one hand, for a unifilar presentation, there is a deterministic relation between the past $x_{-\infty:t}$ and the current hidden state σ_t. That is, $\sigma_t = f(x_{-\infty:t})$, where $f(\cdot)$ is a function; while many pasts $x_{-\infty:t}$ may lead to σ_t. Moreover, for all such pasts, σ_t must have the same conditional distribution $\Pr(X_{t:\infty}|\sigma_t) = \Pr(X_{t:\infty}|x_{-\infty:t})$ of future sequences given the observed past. Since we can use $\Pr(X_{t:t+1}|\sigma_t)$ to predict future observations, we say that the hidden states in a unifilar presentation are *predictive*.

In contrast, when employing a process' nonunifilar presentation to predict its future $\Pr(X_{t:t+1}|\cdot)$ requires a mixture of distributions $\Pr(X_{t:t+1}|\sigma_t)$ over the presentation's states $\{\sigma_t\}$. In this sense, nonunifilar states are not predictive. Nonunifilar presentations still *generate* the process accurately, since the states and transitions explicitly provide a probabilistic procedure for eventually emitting all of a process' realizations with the correct probabilities.

Starting with any unifilar HMM presentation for a stochastic process, one can eliminate redundancies in information about the past by merging states σ_t with identical future probability distributions $\Pr(X_{t:\infty}|\sigma_t)$. Eliminating these redundancies gives a unique minimal optimal predictive HMM for a stochastic process. This canonical presentation is called a process' ϵ-*machine* [20, 22]. The ϵ-machine's states are a process' *causal states*. They are causal is the sense that they give the minimal mechanism that allows one to exactly predict process realizations. The following section introduces these concepts in more detail.

3.4. ϵ-Machines

For ease of notation, we refer here to the past sequences of a process as $\overleftarrow{X} = X_{-\infty:t}$ and the future sequences as $\overrightarrow{X} = X_{t:\infty}$. For finite futures of length ℓ we use $\overrightarrow{X}^\ell = X_{t:t+\ell}$.

As addressed in the previous section, the hidden states in a unifilar presentation must satisfy the condition that, given an observed past sequence $\overleftarrow{x}$, all the allowed hidden states induced by that observation must have the same distribution of futures $\Pr(\overrightarrow{X}|\overleftarrow{x})$. If every observed past induces a unique allowed state in a unifilar presentation, we call that a *causal state*.

Definition 5. Causal states *are the equivalence classes of pasts determined by the equivalence relation* $\sim_\epsilon$. *The latter defines two infinite past sequences* $\overleftarrow{x}$ *and* $\overleftarrow{x}'$ *as equivalent—*$\overleftarrow{x} \sim_\epsilon \overleftarrow{x}'$*—if and only if they have the same conditional distribution of futures:*

$$\epsilon(\overleftarrow{x}) = \{\overleftarrow{x}' | \Pr(\overrightarrow{X}^\ell = \overrightarrow{x}^\ell | \overleftarrow{X} = \overleftarrow{x}) = \Pr(\overrightarrow{X}^\ell = \overrightarrow{x}^\ell | \overleftarrow{X} = \overleftarrow{x}')\} ,$$

where $\overrightarrow{x}^\ell \in \overrightarrow{X}^\ell$, $\overleftarrow{x}' \in \overleftarrow{X}$, *and* $\ell \in \mathbb{Z}^+$. *As before, we denote a causal state random variable by* $\mathcal{S}$, *a particular causal state of an HMM by* σ, *and the set of causal states by* $\boldsymbol{\mathcal{S}}$.

As Reference [21] details, a given causal state and the next observed symbol of a process determine a unique next causal state. A given causal state σ also provides a well-defined conditional probability $\Pr(\overrightarrow{X}|\sigma)$ for all possible future sequences $\overrightarrow{X}$. These two facts together mean there exists a well-defined set of labeled transition matrices $\{T^{(x)}\}$ that describe the probabilities of transition between causal states given an observed symbol x.

Definition 6 (ϵ-machine). *The causal state set* $\boldsymbol{\mathcal{S}}$, *together with its corresponding set of labeled transition matrices* $\{T^{(x)} : x \in \mathcal{A}\}$ *define a process'* ϵ-machine.

By definition, the ϵ-machine is unifilar, and therfore predictive. What is more, for any given stochastic process, its ϵ-machine is unique. It is also the process presentation with maximally accurate prediction of minimal statistical complexity. This makes the ϵ-machine a natural canonical HMM for a process.

The HMM shown in Figure 3.1a is the ϵ-machine for an IID process with bias p, that is, a stochastic process of 0s and 1s in which at each timestep a 0 is output with probability p or a 1 is output

with probability $1-p$. The HMM in Figure 3.1b is the ϵ-machine for the *Golden Mean* process, a stochastic process of all possible sequences of 0s and 1 in which there are no consecutive 0s. The HMM in Figure 3.1c is not an ϵ-machine, its nonunifilarity means that its states are not causal and that it is not predictive.

Since the ϵ-machine is a cannonical presentation for a process, it is important to highlight the following. Given a nonuifilar presentation of a stochastic process, one can use the *mixed state algorithm* to construct a unifilar presentation of the same process. From a unifilar presentation of a process one can merge states with redundant information about future symbols and thus obtain the process' ϵ-machine. This is discussed in more detail in the following section.

3.5. Mixed-State Presentations

The following introduces the *Mixed State Algorithm* (MSA) that converts a nonunifilar presentation to a unifilar presentation of a process, the process' *Mixed State Presentation* (MSP).

First, let's assume that an observer has an HMM presentation M for a process $\mathcal{P}$ that emits symbols $x \in \mathcal{A}$. Before making any observations, the observer has probabilistic knowledge of the current state $\eta_0 = \Pr(\boldsymbol{\mathcal{S}})$. We call this a *state of knowledge* or *belief distribution*. Typically, the best guess for an observer prior to observing any output of the system is $\eta_0 = \pi$, the stationary state distribution of M.

Once M generates a word $w = x_0 x_1 \ldots x_\ell$ the observer's *state of knowledge* of M's current state can be updated to $\eta(w)$. Assuming M has N hidden states, $\eta(w)$ is a row vector with N entries, each corresponding to the probability that M is in a particular hidden state, that is:

$$\eta(w)_\sigma \equiv \Pr(\mathcal{S}_\ell = \sigma | X_{0:\ell} = w, \mathcal{S}_0 \sim \pi) \ . \tag{3.4}$$

The collection of possible *states of knowledge* $\eta(w)$ forms M's set $\boldsymbol{\mathcal{R}}$ of *mixed states*:

$$\boldsymbol{\mathcal{R}} = \{\eta(w) : w \in \mathcal{A}^+, \Pr(w) > 0\} \ ,$$

where $\mathcal{A}^+$ is the set of all words with positive length.

When contructing a MSP from a presentation M with N hidden states, the mixed states η can be represented in the $N-1$-simplex, the set of row vectors such that $\eta \cdot \mathbf{1} = 1$, and $\eta_i \geq 0 \ , \ \forall i \in$

$\{1, \ldots, N\}$. Here $\mathbf{1}$ is a column vector of length N such that all entries are 1. So effectively, the condition $\eta \cdot \mathbf{1} = 1$ means that all entries of η sum to 1.

We can also define the mixed-state measure $\mu(\eta)$— the probability of being in a particular mixed state:

$$\Pr(\eta(w)) = \Pr(\mathcal{S}_\ell | X_{0:\ell} = w, \mathcal{S}_0 \sim \pi) \Pr(w) \ .$$

From this follows the probability of transitioning from $\eta(w)$ to $\eta(wx)$ on observing symbol x:

$$\Pr(\eta(wx)|\eta(w)) = \Pr(x|\mathcal{S}_\ell \sim \eta(w)) \ .$$

This defines the mixed-state dynamic $\mathcal{W}$, in terms of the original process, not in terms of a particular HMM presentation of it.

Now, with the bases of the mixed-state dynamic $\mathcal{W}$, we concentrate on how to build the MSP given any presentation M of a process.

The probability of generating symbol x when in mixed state η is:

$$\Pr(x|\eta) = \eta \cdot T^{(x)} \cdot \mathbf{1} \ , \tag{3.5}$$

where again, $\mathbf{1}$ is a column vector of 1s and $T^{(x)}$ is the x-labeled transition matrix of M. Upon seeing symbol x, the current mixed state η_t is updated:

$$\eta_{t+1}(x) = \frac{\eta_t \cdot T^{(x)}}{\Pr(x|\eta)} \ . \tag{3.6}$$

Thus, given an HMM presentation we calculate the mixed state of Equation (3.4) via:

$$\eta(w) = \frac{\pi \cdot T^{(w)}}{\pi \cdot T^{(w)} \cdot \mathbf{1}} \ .$$

And the mixed-state transition dynamic is then:

$$\begin{aligned} \Pr(\eta_{t+1}, x|\eta_t) &= \Pr(x|\eta_t) \\ &= \eta_t \cdot T^{(x)} \cdot \mathbf{1} \ , \end{aligned}$$

since Equation (3.6) says that, by construction, the MSP is unifilar. That is, the next mixed state is a function of the previous mixed state and the memitted (observed) symbol.

To build the set of mixed states $\boldsymbol{\mathcal{R}}$, one must make an initial mixed state assumption η_0. As mentioned above, this typically is the stationary state distribution π of M. Equation (3.6) then states how to construct all other mixed states and what the symbol labeled transition dynamic is between them.

Together the mixed states and their dynamic give the HMM's *mixed-state presentation* (MSP) $\mathcal{U} = \{\boldsymbol{\mathcal{R}}, \mathcal{W}\}$, a unifilar presentation for the process generated by presentation M.

When constructing a MSP from a finite HMM three things can happen: the mixed state set $\boldsymbol{\mathcal{R}}$ can be finite, countably infinite or uncountably infinite.

The first case, when $\boldsymbol{\mathcal{R}}$ is a finite set, happens when starting from a unifilar HMM or, rarely, when starting from a nonunifilar HMM presentation of a process. If $\boldsymbol{\mathcal{R}}$ is finite, the HMM recovered is the ϵ-machine of the process, since the states of the MSP are *causal*. The MSP might have additional transient states, which can be useful to describe the transients of the process when synchronizing to the ϵ-machine by observing the process, but they can be removed and the recurrent states will constitute the ϵ-machine of the process.

The second case, in which $\boldsymbol{\mathcal{R}}$ is an infinte but countable set, happens occasionally when starting from a nonunifilar presentation of a process M. For instance, when M has a *synchronizing* symbol, such as symbol 1 in the HMM given in Figure 3.1c, then $\boldsymbol{\mathcal{R}}$ is infinite but countable. Here, observation of a 1 *synchronizes* the observer to M in the sense that after observation, the machine can only be in state A. In such cases the MSP will typically have a *renewal process* structure [**64**, **65**]. In this case, after removing transient states, the MSP will be the ϵ-machine of the process.

The third case, in which $\boldsymbol{\mathcal{R}}$ is uncountably infinte is significantly more complicated and, in fact, typical when beginning from a nonunifilar presentation M of a process. In this case, the constructed presentation is predictive, in the sense discussed in the previous section. While minimality is not guaranteed, one can achieve it. This is accomplished by checking for and merging states that have the same conditional probability over future outcomes, which makes them redundant. While this is an interesting open problem, minimality or nonminimality of the MSPs constructed in this case can generally be ignored for many uses. That is, in the sense that the MSP can be considered the

ϵ-machine of the process for the purposes of quantifying informational properties of the process they represent. In the following sections we define those information properties and their quantifiers and how to compute them.

3.6. Information Properties of Stochastic Processes

To characterize and quantify stochastic processes we build on Shannon's information theory [66, 67]. The basis of which is the *Shannon entropy*, a measure of the information contained in a random variable. The Shannon entropy, which we will also call just the entropy or information, of a random variable X is defined as:

$$H[X] = -\sum_{x \in \mathcal{A}} \Pr(X = x) \log_2 \Pr(X = x) \ . \tag{3.7}$$

The above is taken with the added imposition of $0 \log_2 0 = 0$. On the one hand, $H[X]$ is maximized when $\Pr(X)$ is uniform, and each outcome x is equally likely. On the other hand, it is minimized when there is certainty in the random variable X, or $\Pr(X = x^*)$ for some outocme x^*. Intuitively, the entropy of a random variable quantifies how much information one can learn when observing a random variable or, alternatively, how much uncertainty one has when prediciting such random variable.

With this tool in hand, we set out to define the ways to define the intrisic randomness and the memory structure of a stochastic process.

3.6.1. Intrisic Randomness and Shannon Entropy Rate. The *entropy rate* of a stochastic process measures the intrinsic randomness of the process. It is defined as follows:

Definition 7. *Given a stochastic process* $\overleftrightarrow{X}$*, its* entropy rate *is given by:*

$$h_\mu = \lim_{\ell \to \infty} \frac{\mathrm{H}[X_{0:\ell}]}{\ell} \ , \tag{3.8}$$

where $\mathrm{H}[X_{0:\ell}]$ *is the Shannon block entropy [56]. That is, the Shannon entropy over consecutive-symbol blocks of length* ℓ *of the process.*

From Equation (3.8), we see that h_μ is the asymptotic average uncertainty per observed symbol of the stochastic process. Or, said differently, it quantifies how much information an observer gains asymptotically with each newly measured symbol.

Reference [17] explores in detail how to compute the entropy rate of a process generated by a given HMM. Here, we summarize.

For unifilar HMMs, Shannon [67] showed the entropy rate of the stochastic process is exactly computable in closed-form from the HMM's transition matrices and stationary state distribution π, as follows:

$$h_\mu = -\sum_{\sigma \in \mathcal{S}} \pi_\sigma \sum_{x \in \mathcal{A}} \sum_{\sigma' \in \mathcal{S}} T^{(x)}_{\sigma\sigma'} \log T^{(x)}_{\sigma\sigma'} \ . \tag{3.9}$$

As was previously mentioned, in a unifilar presentation there is a one-to-one correspondence between sequences of observer symbols $\ldots x_{t-1}x_t x_{t+1}\ldots$ and sequences of hidden states $\ldots \sigma_{t-1}\sigma_t\sigma_{t+1}\ldots$. In turn, the state-averaged transition uncertainty, weighed by the stationary state distribution of the presentation π, in Equation (3.9) above, corresponds to the observed symbol uncertainty of the process.

For nonunifilar HMMs, there is a one-to-many correspondence between sequences of observed symbols and sequences of hidden states. Thus, applying Equation (3.9) to a nonunifilar presentation of a process can, and will usually, overestimate the true entropy rate of the process.

In Reference [68], Blackwell discusses the entropy rate of *functions of Markov chains* and proposes that, in general, it can be written as an integral weighed by a measure over what we have introduced above as mixed states. Elaborating on this idea, Reference [17] showed that, in the notation of HMMs, the correct general expression for a process' entropy rate is an integral of the transition uncertainty over the mixed-state simplex $\boldsymbol{\mathcal{R}}$ weighted by the invariant measure $\mu(\eta)$:

$$h^B_\mu = -\int_{\boldsymbol{\mathcal{R}}} d\mu(\eta) \sum_{x \in \mathcal{A}} \Pr(x|\eta) \log_2 \Pr(x|\eta) \ . \tag{3.10}$$

The B superscript here is a nod to Blackwell's contribution. We can see that Equation (3.10) reduces to Equation (3.9) for the case in which the MSP of a process has a finite set of states.

Moreover, Reference [17] goes on to describe how to numerically evaluate the integral given in Equation (3.10) in a practical way. The general contractivity of the MSP dynamic $\mathcal{W}$ on the simplex and the ergodicity of the stochastic processes allow for accurate evaluation of the integral expression. This is implemented by taking an average over a time series of mixed states η_t, rather than integrating over the Blackwell measure $\mu(\eta)$. This yields the process' entropy rate:

$$\widehat{h_\mu^B} = -\lim_{\ell\to\infty} \frac{1}{\ell} \sum_{x\in\mathcal{A}} \sum_{i=0}^{\ell} \Pr(x|\eta_i) \log_2 \Pr(x|\eta_i) \ , \tag{3.11}$$

where $\Pr(x|\eta_i) = \eta(x_{0:i}) \cdot T^{(x)} \cdot \mathbf{1}$, and $x_{0:i}$ represents the first i symbols of an arbitrarily long sequence $x_{0:\ell}$ generated by the process' MSP. By tracking the long time behavior of a single realization, one can accurately estimate the entropy rate of the process given the MSP of the process.

3.6.2. Information Storage. To quantify the structure of a process' presentation M, the most straightforward measure is its number $|\boldsymbol{\mathcal{S}}|$ of hidden states. Beyond that, a more insightful metric is the amount of historical memory or information the presentation states contain. This is given by the Shannon entropy of the state distribution:

$$\mathrm{H}[\boldsymbol{\mathcal{S}}] = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \pi_\sigma \log_2 \pi_\sigma \ . \tag{3.12}$$

It quantifies how much information the hidden states store about past observations. That is, it measures how much memory a given HMM has. And, since unifilar presentations are predictors, $\mathrm{H}[\boldsymbol{\mathcal{S}}]$ is an upper bound on the amount of information one must maintain on average to optimally predict the process. So that these metrics describe actual properties of the stochastic process in question and not those of a particular HMM presentation—that, say, could have an overly-large and redundant set of states—we use a process' ϵ-machine which, as stated above, is a process' minimal optimal predictive HMM presentation [20–22].

The ϵ-machine's hidden states are a process' *causal states* since they optimally capture the process' causal structure. With them one can make the most accurate predictions of future symbols and their associated probabilities. The memory in the causal states is then the minimal amount of information from the past that must be stored in hidden states to optimally predict the future.

DEFINITION 8. *A stochastic process'* statistical complexity C_μ *is the Shannon entropy of its* ϵ*-machine's causal states* $\boldsymbol{\mathcal{S}}$*:*

$$\begin{aligned} C_\mu &= \mathrm{H}[\boldsymbol{\mathcal{S}}] \\ &= -\sum_{\sigma \in \boldsymbol{\mathcal{S}}} \Pr(S = \sigma) \log_2 \Pr(S = \sigma) \ . \end{aligned} \tag{3.13}$$

Operationally, C_μ is the minimal memory required to optimally predict the future of a stochastic process.

For stochastic processes with ϵ-machines that have a finite number of causal states, or even a countably infinite number of causal states, C_μ can be computed or numerically approximated, respectively. In those cases the statistical complexity of the process converges to a finite number. For stochastic processes with ϵ-machines that have an uncountable infinity of causal states, C_μ diverges, this indicates that they require an infinite amount of memory to optimally predict. Beyond that statement, one can still quantify the structure and memory of these processes by tracking how the memory resources required to predict the process scale with increasing the precission of the prediction [69].

To quantify the memory structure of these uncountably-infinite-state processes with divergent statistical complexity, Reference [18] introduces the concept of *statistical complexity dimension*, d_μ. d_μ quantifies the divergence rate of C_μ, it is the information dimension of the Blackwell measure $\mu(\eta)$ on the set of mixed states $\boldsymbol{\mathcal{R}}$:

$$d_\mu = -\lim_{\epsilon \to 0} \frac{H_\epsilon[\boldsymbol{\mathcal{R}}]}{\log_2 \epsilon} \ . \tag{3.14}$$

This tracks the rate at which the memory requirements for optimal prediction grow with increasing precision $-\ln \epsilon$. Specifically, $H_\epsilon[Q]$ is the Shannon entropy of the continuous-valued random variable Q coarse-grained at size ϵ. So, in particular d_μ is the information dimension of the set of mixed states coarsgrained at size ϵ. Even when C_μ diverges, d_μ allows us to quantify the rate at which it does and is therfore a measure of memory structure for stochastic proceses. In particular, for processes for which C_μ is finite, d_μ is zero.

Evaluating the statistical complexity dimension d_μ is not a simple matter, though. The procedure is presented in detail in Refs. [18, 19]. In particular, Reference [19] introduces the *ambiguity rate* h_a, which quantifies the rate at which optimal predictive models discard information by introducing uncertainty over the infinite past. The difference between the entropy rate and the ambiguity rate in a process is fundamental to determine its statistical complexity dimension, as well as the cardinality of its set of mixed states $\boldsymbol{\mathcal{R}}$.

The details of the meaning and calculation of the ambiguity rate are highly nuanced and beyond the scope of this work. Here we highlight only a few main points, which are required for the analysis of Quantum State Stochastic Processes:

(1) When $h_\mu - h_a = 0$, the ϵ-machine of the stochastic process discards information at the same rate that it generates information. In this case C_μ converges to a finite value and $d_\mu = 0$. This happens for processes with finite state ϵ-machine and, for our practical purposes, also for processes with countably-infinite-state ϵ-machines.

(2) When $h_\mu > h_a \geq 0$, the stochastic process' C_μ diverges and its ϵ-machine requires an uncountable infinity of causal states. In that case, the statistical complexity dimension can be computed from knowledge of the process' h_μ, h_a and the Lyapunov spectrum of the causal-state dynamic.

With that, we have presented all the classical tools that will be used to study measured quantum-state stochastic processes. The following chapter sets out to develop a rigurous definition of quantum-state stochastic processes and how we model them.

CHAPTER 4

Quantum State Stochastic Processes

The following introduces the main objects of study—quantum-state stochastic processes—and the information sources that generate them.

4.1. Quantum Processes

Consider a given quantum source that emits a sequence of individual quantum states. At each time step, the quantum state it emits takes on a value from a finite set. We refer to these sources as *controlled quantum sources* (CQSs) and, in their operation, they generate *quantum-state stochastic processes* (QSSPs). In Figure 1.1, the black box represents the CQS, and the sequence of output qubits are a realization of the QSSP. We will now define quantum-state stochastic processes.

Let R_t denote the random variable for the quantum state emitted at time t. The realization of R_t as a particular quantum state is $\rho_t \in \mathcal{A}_Q$, where $\mathcal{A}_Q$ is the set of quantum states in a Hilbert space that the random variable can take. The random variable for a sequence of quantum states emitted between times t and $t + \ell$ is denoted by the block random variable $R_{t:t+\ell}$ (inclusive on the left, exclusive on the right). Then $\rho_{t:t+\ell}$ denotes the realized sequence of quantum states.

Definition 9 (Homogeneous Quantum-State Stochastic Process). *Let $\mathcal{A}_Q \subseteq \mathcal{H}^d$ be the set of available quantum states in the d-dimensional Hilbert space $\mathcal{H}^d$. $\Omega = \mathcal{A}_Q^{\mathbb{Z}}$ is then the space of bi-infinite sequences over $\mathcal{A}_Q$. Consider the probability space $\mathcal{P} = (\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is the $\sigma-$algebra on the cylinder sets of Ω and P a probability measure over the cylinder sets. $R_{-\infty:\infty}$ denotes the discrete-time random-variable sequence of quantum states described by the quantum-state stochastic process $\mathcal{P}$. It comprises the sequences of random variables that take on values according to a measurable function $T_t : \Omega \to \mathcal{A}_Q^t$:*

$$R_t = T_t(R_{-\infty:\infty}) \ ,$$

for $t \in \mathbb{Z}$.

In simpler terms, a QSSP is defined by a probability measure over realizations of bi-infinite sequences of random variables, $R_{-\infty:\infty}$.

To emphasize again, we work with distinct physical objects at each time step. That is, each random variable R_t takes a value on its own Hilbert space $\mathcal{H}_t^d$. In a *homogeneous* process the Hilbert spaces are of the same type. This is in contrast to a situation in which one could track the discrete time evolution of the quantum state of a physical system, as described in Chapter 2. Think, for example, of a time series of photons, each photon emitted by the source at a certain time step, in this or that quantum state, and each of those states are in $\mathcal{H}^2$.

We focus our attention in *homogeneous* QSSPs, but emphasize that there is room for future study of *nonhomogenenous* QSSPs, in which the alphabet $\mathcal{A}_Q^t$ is time dependent and can live in a Hilbert space of different dimension than some other $\mathcal{A}_Q^{t'}$.

We will further restrict our study to *stationary* and *ergodic* QSSPs, a fairly common restriction in the study of stochastic pocesses.

Definition 10 (Stationarity). *A stationary QSSP is one in which the probability of observing a particular sequence of quantum states is independent of the time at which the observation is made. That is, the probability of an observed quantum sequence is time-translation invariant:*

$$\Pr(R_{t:t+\ell} = \rho_{t:t+\ell}) = \Pr(R_{0:\ell} = \rho_{0:\ell}) \ , \tag{4.1}$$

for all $t \in \mathbb{Z}$, $\ell \in \mathbb{Z}$, *and* $\rho_{0:\ell}$.

Definition 11 (Ergodicity). *A QSSP is* ergodic *if all long realizations of the strochastic process obey the QSSP's statistical properties. That is, given a long realization* $R_{0:n} = \rho_{0:n}$, *the probability of observing a finite realization of* $R_{0:\ell}$ *of length* $\ell \ll n$ *in* $\rho_{0:n}$ *is the same as observing that same block in multiple realizations of length* ℓ *drawn from the QSSP.*

The following considers only QSSPs satisfying these three definitions. Further, it considers QSSPs which are *separable*.

DEFINITION 12 (Separable QSSP). *A QSSP is* separable *if for all t, the output quantum state realized by R_t is pure. That is, $\rho_t \in \mathcal{A}_Q$ satisfies $\rho_t^2 = \rho_t$. Sequences of random variables in a QSSP can then be written as:*

$$R_{t:t+\ell} = R_t \otimes R_{t+1} \otimes \ldots \otimes R_{t+\ell-1} \tag{4.2}$$

and a realization is both a pure state and the tensor product of the individual pure quantum states:

$$\rho_{t:t+\ell} = \rho_t \otimes \rho_{t+1} \otimes \ldots \otimes \rho_{t+\ell-1} \ .$$

Effectively, this rules out of the present study QSSPs in which there is entanglement within the sequences of output quantum states. While the present work focuses in time series without entanglement, the time series of quantum states are still correlated and prove to be complicated. This work serves as a starting point to approach time series with entanglement in the future. For instance, considering time series in which the quantum state of the chain of quantum states can be described as a Matrix Product states. These can be generated with Hidden Markov Models [**70**], as the QSSPs described below, but also allow for a specific form of entanglement.

We note here, that since the QSSPs to be considered are made of pure qudit states at each time step, the stochastic process can also be described in ket notation. In this case, the quantum state of the qudit at time t is described by $|\Psi_t\rangle$ and takes on value $|\psi_t\rangle$. The state of a block can then be described by $|\psi_{w(\ell)}\rangle = |\psi_{t:t+\ell}\rangle$, where $w(\ell)$ is a realization of an allowed 'word' or sequence of symbols of length ℓ of the QSSP. If a particular block of size ℓ has a probability $\Pr(w(\ell))$ of being in state $|\psi_{w(\ell)}\rangle$, then a block of length ℓ of the QSSP can be described by the density matrix:

$$\rho_{0:\ell} = \sum_{w(\ell)} \Pr(w(\ell)) \, |\psi_{w(\ell)}\rangle \, \langle\psi_{w(\ell)}| \tag{4.3}$$

Where the sum goes over all allowed values of $w(\ell)$ for the QSSP. This notation, although it will not be used commonly in the present work, connects it with the development presented in Reference [**71**] and simplifies the definition of several relevant quantifiers of the quantum information content of the QSSP.

To summarize, together these definitions describe a setup in which a quantum system undergoing a certain dynamic and emits individual qudits at each time step. The latter are in pure quantum states. The quantum system that emits the qudits is the controlled qubit source (CQS). Occasionally, this is abbreviated as the *controller* to make reference to an experimenter having some level of control, perhaps through design, over qudit emission, which can be stochastic. We now propose a way to model classical controlers.

4.2. Quantum Sources and Generator Presentations

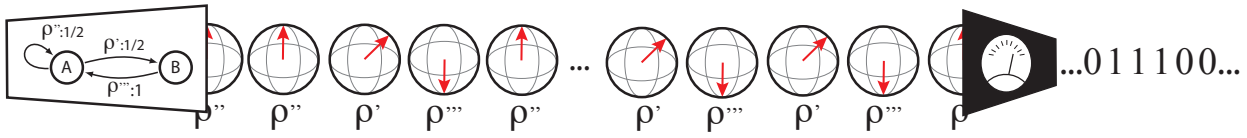


FIGURE 4.1. Classically-Controlled Quantum Source (cCQS) and its Emitted Process: (Left) As the internal hidden Markov controller operates, at each time t the emitted symbols $\mathcal{A}$ (not shown) determine the qubit's quantum state $\rho_t \in \mathcal{A}_Q$. As shown, the latter are realized as density matrices ρ', ρ'', and ρ'''. (For simplicity the controller's emitted symbols $\mathcal{A}$ are not explicitly shown, only the resulting quantum states $\rho \in \mathcal{A}_Q$.) (Middle) The resulting output quantum process is a sequence of distinct $\mathcal{H}^2$ Hilbert spaces, displayed as a series of Bloch spheres with realized state vectors displayed inside. (Right) Measuring each qubit realizes a classical stochastic process ...

In short, quantum sources generate quantum processes. Here, we concentrate on a particular implementation of CQSs—those that generate QSSPs via a classical controller with a finite memory in the form of a hidden Markov model (HMM) [**61**, **62**, **72**]. Recall that classical HMMs were described and summarized in Section 3.3. To illustrate, this means that the black box in Figure 1.1 is reified as in the white box shown in Figure 4.1: a HMM controlling what the box emits. Correspondingly, we refer to this qubit source as a *classically controlled qubit source* (cCQS).

The choice of finite-state HMM controller to model a cCQS is natural, in that HMMs are a standard representation of finite-memory stochastic processes, are widely used to model noisy classical sources and communication channels, and can be fully analyzed [**73**, **74**]. Also, looking to the future and extending the present framework, HMMs are easily extended to represent more general sources, such as those driven by quantum controllers [**75**].

DEFINITION 13 (Classically-Controlled Quantum Source). *A cCQS is a tuple* $(\boldsymbol{\mathcal{S}}, \mathcal{A}, \mathcal{A}_Q, \{\mathbb{T}^\rho\})$ *where:*

(1) $\boldsymbol{\mathcal{S}}$ *is the set of hidden states of the HMM controller.*

(2) Alphabet $\mathcal{A}$ *is a finite set of symbols in one-to-one correspondence with the set* $\mathcal{A}_Q$ *of available pure qudit quantum states. The latter are emitted by the cCQS when a symbol is encountered on a transition. To simplify notation, both a symbol and its qudit state are denoted by a density matrix* $\rho \in \mathcal{H}^d$.

(3) $\{\mathbb{T}^\rho : \rho \in \mathcal{A}_Q\}$ *is a set of quantum-state labeled transition matrices of size* $|\boldsymbol{\mathcal{S}}| \times |\boldsymbol{\mathcal{S}}|$. $\mathbb{T}^\rho_{\sigma\sigma'}$ *is the probability of transitioning from internal state* σ *to internal state* σ' *(both in* $\boldsymbol{\mathcal{S}}$*) while emitting symbol (or, effectively, quantum state)* ρ.

The labeled transition matrices $\{\mathbb{T}^\rho\}$ sum to the internal-state stochastic transition matrix over hidden states: $\mathbb{T} = \sum_{\rho \in \mathcal{A}_Q} \mathbb{T}^\rho$. This, in turn, determines the HMM's stationary internal-state distribution π as the left eigenvector of $\mathbb{T}$ with eigenvalue 1: $\pi = \pi\mathbb{T}$. π is then a vector of size $|\boldsymbol{\mathcal{S}}|$ in which the entry π_σ represents the asymptotic probability of the HMM being in internal state $\sigma \in \boldsymbol{\mathcal{S}}$.

It is important here to make a comment about the use of the word *state*, which is arguably overloaded with meaning in phyiscs. Here, the terms *hidden* or *internal states*, refer to the internal memory states of a HMM or a cCQS. This is in contrast with *quantum state* or *qudit state*.

For an example, see the cCQS in Figure 4.2, where:

(1) $\boldsymbol{\mathcal{S}} = \{A, B\}$.

(2) $\mathcal{A}_Q = \{\rho_\psi = |\psi\rangle\langle\psi|, \rho_\varphi = |\varphi\rangle\langle\varphi|\}$. Here, $|\psi\rangle$ and $|\varphi\rangle$ are pure quantum states.

(3) $\{\mathbb{T}^\rho\} = \{\mathbb{T}^{\rho_\varphi}, \mathbb{T}^{\rho_\psi}\}$, where:

$$\mathbb{T}^{\rho_\varphi} = \begin{pmatrix} 0 & 1/2 \\ 0 & 0 \end{pmatrix}$$

$$\mathbb{T}^{\rho_\psi} = \begin{pmatrix} 1/2 & 0 \\ 1 & 0 \end{pmatrix} .$$

(4) $\pi = \begin{pmatrix} 2/3 & 1/3 \end{pmatrix}$, the left eigenvector of:

$$\begin{aligned} \mathbb{T} &= \mathbb{T}^{\rho_\varphi} + \mathbb{T}^{\rho_\psi} \\ &= \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix} . \end{aligned}$$

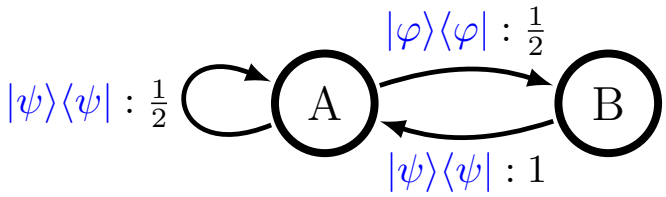


FIGURE 4.2. cCQS Operation: If the HMM controller is in state A the cCQS has equal probabilities of remaining there or transitioning to state B. If a transition to state B then occurs, the system emits a qubit in the pure state $|\varphi\rangle\langle\varphi|$. In the next time step the cCQS must transition to state A and it then emits a qubit in state $|\psi\rangle\langle\psi|$. That is, the edge labels $\rho : p$ indicate taking the state-to-state transition with probability p and emitting quantum state ρ. As the cCQS operates, a qubit is output at each time step, over time the result is a qubit time series.

The current work implements cCQSs with finite-memory HMM controllers: $|\boldsymbol{\mathcal{S}}| < \infty$. It also specifies that HMM controller transitions are *unifilar*, as defined in Section 3.3.1: the current internal hidden state and emitted quantum state uniquely determine the next internal hidden state.

Section 3.3.1 gives a more general and detailed account of unifilarity. Sections 3.3.2 and 3.6 then highlight several of its consequences. As will become clear, the distinction between unifilar and nonunifilar HMMs plays a large role in driving the complexity of quantum processes on their own and when measured.

These choices ensure that the cCQS's randomness and complexity can be directly calculated. And so, the effects of measurement on the quantum process are made most explicit. That said, the analysis can be applied directly to nonunifilar controllers, with the caveat that the controller's structure and randomness are more complicated to quantify, as was made explicit in Section 3.6.

4.3. Quantum Process Characterization

In the same way that we use classical information theory to describe classical stochastic processes, as summarized in Chapter 3.6, we can use quantum information theory to describe the properties of separable QSSPs. While the main focus of this work is the characterization of the classical stochastic processes obtained after observation of a QSSP, to illustrate the quantum-information characterization of QSSPs we highlight here a few select quantities that serve to characterize the QSSPs directly. A detailed characterization of the quantum information properties of QSSPs is developed in Reference [**71**].

While in the classical regime the basic metric for information is the Shannon entropy Equation (3.7), in the quantum regime that role is played by the *von Neumann entropy* [**76**, **77**].

Definition 14. *The von Neumann entropy S of a quantum mechanical system in a quantum state described by a density matrix ρ is given by:*

$$S(\rho) = -tr(\rho \log_2 \rho) = -\sum_i \lambda_i \log_2 \lambda_i \ , \tag{4.4}$$

with λ_i the eigenvalues of the density matrix ρ.

$S(\rho) = 0$ only when ρ describes a pure state, so $S(\rho)$ is a way to quantify the departure of a system from a pure state. As can be seen form Equation (4.4), it can also be interpreted as the Shannon entropy of the eiganvalues of the density matrix ρ describing the quantum state.

We can also consider the von Neumann entropy of blocks or sequences of length ℓ for the QSSP.

Definition 15. *The quantum block entropy or von Neumann block entropy of a block of length ℓ of a QSSP is given by:*

$$S(\ell) = -\operatorname{Tr}(\rho_{0:\ell} \log_2 \rho_{0:\ell}) \ , \tag{4.5}$$

with $\rho_{0:\ell}$ defined as in Equation (4.3).

$S(\ell) = 0$ when $\rho_{0:\ell}$ describes a pure quantum state. As stated in detail in Reference [**71**], $S(\ell)$ is a concave, nondecreasing function of ℓ. It quantifies the amount of information in a block of lenght ℓ random variables of a QSSP.

4.3.1. von Neumann Entropy Rate.

We can then track the average amount of quantum information content per symbol in a QSSP.

DEFINITION 16 (Quantum Entropy Rate). *The von Neumann entropy rate or quantum entropy rate of a QSSP is given by:*

$$s = \lim_{\ell \to \infty} \frac{S(\ell)}{\ell} \ . \tag{4.6}$$

The von Neumann entropy rate exists for all stationary quantum-state stochastic processes. Reference [71] states and proves its relationship to the classical entropy rate of the classical stochastic processes that can be obtained by measuring the QSSP. These will be the objects of our attention.

It is worth noting here that we have written a Python package, QSSP, which is available and can compute the quantities introduced above as well as manage HMMs, QSSPs and measured QSSPs as described in more detail in Appendix A.

The following chapters will consider and describe the interaction of an observer with a QSSP via measurement. To simplify the analysis, we focus on QSSPs in which the output time-series is made of qubits, the simplest carriers of quantum information. In general, the analysis to follow can be extended to qudits, but their consideration would significantly complicate the development. When relevant, comments about QSSPs consisting of qudits with $d > 2$ will be made.

CHAPTER 5

Measured Quantum-State Stochastic Processes

A central aspect of the process realized in the laboratory is how an observer interacts with the QSSP. Naturally, interactions occur via quantum measurement, but there are multiple options for both the measurement type and the measurement protocol to implement. We now set the stage for measurement protocols and define the specific measurements and protocols to be studied in this work.

5.1. Measurement Protocols

We view a measurement protocol as a communication channel between the quantum stochastic process $R_{-\infty:\infty}$ and its measured companion—a classical stochastic process $X_{-\infty:\infty}$. Denote the relationship between the QSSP random variables and those of the measured process by:

$$X_t = \mathcal{I}_t(R_t) \ . \tag{5.1}$$

Where $\mathcal{I}_t$ represents the action of a measurement on the quantum state ρ_t output by the QSSP at time t. The set $\mathcal{I} = \{\mathcal{I}_t\}$ for all t defines a measurement protocol. In a slight abuse of notation, we use the variable $\mathcal{I}_t$ to represent both the measurement channel and the particular measurement operator applied at time t. For short, we refer to the measurement protocol as just $\mathcal{I}$ and denote the relationship between the QSSP and its corresponding measured stochastic process by:

$$X_{-\infty:\infty} = \mathcal{I}(R_{-\infty:\infty}) \ . \tag{5.2}$$

To recapitulate the notation, at each time step the random variable R_t takes on the value of a particular quantum state ρ_t. That, in turn, is measured with operator $\mathcal{I}_t$. The resulting measurement outcome is denoted by the random variable X_t, which takes on a particular value $x_t \in \mathcal{A}_M$, with $\mathcal{A}_M$ the set of possible measurement outcomes.

We now define the basic measurement protocol used.

DEFINITION 17 (Single State Constant-Measurement Protocol). *As a QSSP is output, each quantum state passes through a measurement channel $\mathcal{I}_t = \mathcal{E}$, for all t. That is, the same measurement $\mathcal{E}$ is applied to each output state individually: $X_t = \mathcal{E}(R_t)$.*

Though the following employs only this protocol, we note that it is straightforward to work with measurement protocols for which $\mathcal{I}_t$ depends on time. For example, a measurement scheme in which the measurements alternate between a measurement along the z-axis and a measurement along the x-axis. Or, a potentially more useful protocol is one in which measurements are adaptive, and the $\mathcal{I}_t$s are chosen given the outputs of a subset of past measurements. One approach to the adaptive measurement protocols is described in Reference [**71**].

5.1.1. Projective Measurements. At each time-step, the observer performs a single measurement $\mathcal{E}$ on the quantum state ρ emitted by the controller. This measurement consists of a finite set of nonnegative operators $\{E_x\}$, in which the index $x \in \mathcal{A}_M$ labels the measurement outcomes. The measurement operators sum to the identity: $\sum_{x \in \mathcal{A}_M} E_x = \mathbb{I}$. The probability of measurement outcome x after measuring quantum state ρ is given by $\Pr(x|\rho) = \mathrm{Tr}(E_x \rho)$, where $\mathrm{Tr}(\cdot)$ is the trace. We call this a *positive operator-valued measurement* (POVM). A special subset of POVMs is those in which the measurement operators are mutually orthogonal, projection-valued measurements, as discussed below.

We concentrate our analysis on single-state projective measurements, in which the operators are orthogonal. This simplifies the basic framework, for now. That said, Section 6.2.1 briefly considers single-state protocols with more-general POVMs.

DEFINITION 18 (Projective Measurement on a QSSP). *A* projective measurement $\mathcal{E}$ *in a dimension d Hilbert space $\mathcal{H}^d$ consists of a set of d mutually orthogonal projectors $\{E_x\}$, with $x \in \{0, 1, \ldots d-1\}$, such that $E_x E_y = E_x \delta_{xy}$ and $\sum_x E_x = \mathbb{I}_d$. When the measurement $\mathcal{E}$ acts on a quantum system in state ρ_t, emitted by the QSSP, the outcome is x_t with probability $\Pr(x_t|\rho_t) = \mathrm{Tr}(E_{x_t} \rho_t)$. Applying a projective measurement to every state emitted by the QSSP yields a classical stochastic process over the values of x. We call the set of possible values of x the* measured process alphabet $\mathcal{A}_M$.

Notice that in a projective measurement applied to a quantum state in a Hilbert space of dimension d, there must be at most d measurement operators. In more general (non-orthogonal) POVMs, there can be more than d measurement operators.

As an example of a projective measurement, consider a measurement of a qubit consisting of two orthogonal measurement operators $\{E_0, E_1\}$. Without loss of generality these can be written as:

$$E_0 = |\psi_0\rangle \langle\psi_0|$$

$$E_1 = |\psi_1\rangle \langle\psi_1| \ .$$

Later, we refer to the set $\mathbf{E} = \{E_i : i = 0, 1, 2, \ldots\}$ of such measurement operators as the *observation basis*.

With specified $|\psi_i\rangle \in \mathcal{H}^2$ and $\langle\psi_0|\psi_1\rangle = 0$. When working with qubit projective measurements, it is convenient to parametrize them using Bloch angles as follows:

$$|\psi_0\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{5.4a}$$

$$|\psi_1\rangle = \sin\frac{\theta}{2}|0\rangle - e^{i\phi}\cos\frac{\theta}{2}|1\rangle \ . \tag{5.4b}$$

The resulting measurements are labeled 0 or 1, respectively. Let X_t denote the random variable for the outcome of measuring the state ρ_t at time t and x_t the realized value. In the case of qubit projective measurements $x_t \in \{0, 1\}$. In this way, applying projective measurement protocol to a QSSP produces a binary classical stochastic process. Knowledge of the cCQS's controller and the measurement protocol are the basic ingredients needed to analyze the mechanism that generates these classical stochastic processes—the measured processes.

5.2. Measured Processes

By Equation (5.2) the classical stochastic process that is realized by measuring the QSSP is the joint distribution $\Pr(X_{-\infty:\infty})$. The specific value x_t taken on by the random variable X_t depends both on the measurement protocol $\mathcal{I} = \{\mathcal{I}_t\}$ with measurement operators $\{E_{x_t}\}$ and the QSSP and its

HMM controller. That is, if the random variable R_t takes on the particular value ρ_t at time t, then:

$$\Pr(X_t = x | R_t = \rho_t, \mathcal{I}_t) = \mathrm{Tr}(E_{x_t} \rho_t) \ . \tag{5.5}$$

Both the measurement protocol and the QSSP can introduce correlations within the classical stochastic process. That is, even if applying a time-independent measurement protocol, such as a constant single-state measurement protocol, the correlations in $R_{-\infty:\infty}$ will yield correlations in $X_{-\infty:\infty}$. However, even if $R_{-\infty:\infty}$ is an independent, identically distributed (IID) process, a time-correlated measurement protocol $\mathcal{I}$ can yield a correlated classical stochastic process.

5.2.1. Measured Process Presentations. Importantly, in cases where the QSSP is generated by a cCQS, the measured quantum process can be modeled with a unique HMM, as the following demonstrates.

PROPOSITION 1. *Applying a projective measurement protocol $\mathcal{E}$ to a QSSP $R_{-\infty:\infty}$ generated by a finite-state cCQS M results in a measured process $X_{-\infty:\infty}$ given by a finite-state HMM.*

PROOF. *We establish this by directly constructing the HMM presentation. The HMM $\mathbb{M} = \{\boldsymbol{\mathcal{S}}, \mathcal{A}_M, \{T^x\}, \pi\}$ that generates the measured process has the following components:*

(1) The same set $\boldsymbol{\mathcal{S}}$ of internal states as the HMM that generated the QSSP.

(2) A finite alphabet $\mathcal{A}$ consisting of each possible measurement outcome.

(3) A set of labeled transition matrices $\{T^x\}$ with $x \in \mathcal{A}_M$ such that:

$$T^x = \sum_{\rho \in \mathcal{A}_Q} \mathbb{T}^\rho \Pr(x | \rho, \mathcal{E}) \tag{5.6}$$

with:

$$\Pr(x | \rho, \mathcal{E}) = \mathrm{Tr}(E_x \rho) \ . \tag{5.7}$$

(4) The same stationary distribution π as the HMM that generated the QSSP.

DEFINITION 19. *We refer to the resulting HMM as a* measured cCQS.

In this way, fixing a cCQS and a measurement basis determines a unique measured cCQS. This HMM accurately describes the classical stochastic process resulting in the lab.

One would hope to directly analyze the statistical properties of the classical process using that HMM. Or, more modestly, to better and more accurately analyze the classical process using the HMM than by simulating repeated realizations over long times to obtain statistics for estimation. The HMM, after all, exactly describes the process, being a presentation.

We demonstrate that this analysis is very far from a straightforward procedure. Moreover, the difficulties are (i) inherent and (ii) generic to quantum measurement. Despite the challenges, though, with care and the right tools in hand one can fully characterize the measured process' properties.

We introduced two classes of HMMs in Section 3.3.1 —those that are unifilar and those that are not. The following then explains why measurement induces complex statistics. Specifically, the following establishes that (i) nonunifilarity arises in the measured process HMM, (ii) this is generic for projective measurements, and (iii) complex statistics arise in the measured process due to an exponential explosion of the predictive feature set. Along the way, we will use another concept defined in Section 3.3.2: that of generative and predictive presentations—those that can be used to produce process realizations and those that, in addition, can be used to optimally predict realizations.

5.3. Measured QSSP Complexity

As stated above, a given stochastic process can be generated by many different HMMs. Each is called a *presentation* of the given process. The essential requirement is that a presentation describes all and only a process' realizations and their probabilities. HMM presentations are either *unifilar* or *nonunifilar*. Unifilarity controls how useful a presentation is to quantitatively analyzing a process.

Looking ahead, we must distinguish between processes with finite predictive presentations and those without. We now introduce a concept that will allow us to do just that.

Definition 20 (Irreducible Nonunifilarity). *A stochastic process is* irreducibly nonunifilar *if there exists no finite unifilar HMM presentation that predicts it. This does not rule out that there can exist a finite nonunifilar HMM that generates the stochastic process.*

5.3.1. Measurement-induced Nonunifilarity. The tools are now in place to address the origins of measurement-induced complexity in observed quantum-state stochastic processes. First, we identify the emergence of nonunifilarity. Second, we argue this happens frequently and, in fact,

is generic in measured QSSPs. Third, we explore the consequence—explosive complexity. Finally, we identify the underlying mechanism driving this in quantum state indistinguishability.

PROPOSITION 2. *Quantum measurement of a QSSP generated by a cCQS can lead to an classical process represented by a nonunifilar measured cCQS.*

PROOF. *Consider cCQS hidden state σ with outgoing transitions to two distinct hidden states σ_I and σ_{II}. The first transition emits quantum state ρ' and the second, ρ''. Now, performing a measurement $\{E_0, E_1\}$ on the emitted quantum states, both measured transitions have nonzero probability of emitting the same symbol. Recall from Equation (5.6):*

$$T^x_{\sigma\sigma'} = \sum_{\rho \in \mathcal{A}_Q} \mathbb{T}^\rho_{\sigma\sigma'} \Pr(x|\rho) \ .$$

Note that above and in what follows we suppress explicit mention of the measurement protocol in the conditional probabilities, simplifying $\Pr(x|\rho, \mathcal{E})$ to $\Pr(x|\rho)$ for ease of notation. Now, consider a ρ that gives a nonzero probability of obtaining measurement outcome x. Note that this is the case for $\sigma' = \sigma_I$ and $\sigma' = \sigma_{II}$. Then, as long as $\mathbb{T}^\rho_{\sigma\sigma'}$ is nonzero for this ρ, both $T^x_{\sigma\sigma_I}$ and $T^x_{\sigma\sigma_{II}}$ will be nonzero. This makes the observed transition out of state σ nonunifilar and, thus, makes the measured cCQS nonunifilar.

The implications become more apparent later, when discussing how HMM nonunifilarity almost always implies that the process it generates is irreducibly nonunifilar. In this way, measurement can—and as discussed later, almost always will—induce irreducible nonunifilarity of the measured process.

5.3.2. Nonunifilarity is Generic. We say a property is *measurement generic* over a set of measurements—for instance, qubit projective measurements—if it holds true for almost all measurements but a measure zero subset. Similarly, a property is *source generic* if it holds true for the QSSPs generated by almost all cCQSs.

PROPOSITION 3. *A measured cCQS HMM is generically nonunifilar. This is true both source generically and measurement generically over the set of projective measurements.*

PROOF. *Consider the constraints that give rise to unifilar transitions. Recall the entries of the labeled-transition matrices for a measured cCQS, as defined in Equation (5.6):*

$$T^{x}_{\sigma\sigma'} = \sum_{\rho \in \mathcal{A}_Q} \mathbb{T}^{\rho}_{\sigma\sigma'} \Pr(x|\rho) \ .$$

Each entry is composed of a sum of terms. For the measured cCQS to maintain unifilarity there should be at most one nonzero term per row of each labeled transition matrix. That is, for each $x \in \mathcal{A}$ and $\sigma \in \boldsymbol{\mathcal{S}}$ pair, the term $T^{x}_{\sigma\sigma'}$ is nonzero for at most one value of σ'.

For unifilarity to hold, the following conditions on the underlying cCQS and measurement must be satisfied:

(1) A hidden state with an outgoing transition to only one other hidden state maintains unifilarity.

(2) All cCQS hidden states can have at most two outgoing transitions (to distinct states). Denote the quantum states emitted on the outgoing edges ρ_a and ρ_b and the two destination states σ_a and σ_b, respectively. Thus, for each cCQS hidden state σ there are at most two nonzero transition elements: $\mathbb{T}^{\rho_a}_{\sigma\sigma_a}$ and $\mathbb{T}^{\rho_b}_{\sigma\sigma_b}$, say. When determining the measured cCQS's labeled transitions $T^{x}_{\sigma\sigma'}$, for each σ and x, there will be at most two contributions. The following condition ensures that these two contributions do not result in two or more nonzero values for each σ.

(3) If the state has two outgoing transitions then, to maintain unifilarity, it must satisfy:

- *The two emitted quantum states ρ_a and ρ_b on the outgoing transitions must be orthogonal to each other.*
- *The measurement basis must be aligned with ρ_a and ρ_b. That is, each measurement operator must project into the quantum states that the process emits—ρ_a and ρ_b.*

For a given pair of x and σ, these ensure that the only potentially nonzero terms are $T^{\rho_a}_{\sigma\sigma_a}$ and $T^{\rho_b}_{\sigma\sigma_b}$. If the projective measurement with outcome x is aligned with either ρ_a or ρ_b, however, then only one of the terms of the form $\mathbb{T}^{\rho}_{\sigma\sigma'} \Pr(x|\rho)$ can be nonzero for a given ρ. And so, there will be at most one nonzero term in the σ row of transition matrix T^{x}. This guarantees that the measured cCQS remains unifilar.

Conditions 1, 2, and 3 are highly restrictive in the space of cCQSs. That is, almost none of the possible labeled transition matrices $\mathbb{T}^{\rho}$ satisfy them. In turn, this means that measured cCQSs are source-generically nonunifilar.

For a given cCQS, Condition 3 is highly restrictive in the set of projective measurements and is only satisfied for one measurement choice out of a continuous set of possible measurement choices. Therefore, the measured cCQS is measurement-generically nonunifilar.

It is also important here to note that Conditions 2 and 3 refer to the case in which the cCQS emits qubits. For a cCQS that emits qudits, these conditions can be generalized to allow for d distinct transitions to other states. The generalization is straightforward: To maintain unifilarity the output quantum states associated with those d transitions must be mutually orthogonal and the measurement chosen must be able to distinguish perfectly between those d states.

As Sec. 5.3.4 develops in more detail, Prop. 3 says that measured processes are typically highly complex, in the sense that they generically have an uncountable infinity of predictive features (causal states), divergent statistical complexity, and a positive entropy rate.

5.3.3. Variations. Several observations are in order on generic nonunifilarity for qubit processes and how generic nonunifilarity trades-off against the Hilbert space dimension of the QSSP's quantum states.

Structurally, a binary alphabet highly restricts the possibilities for a particular HMM's topology to support unifilarity. This could lead to a rushed conclusion that restricting to projective measurements plays a determinant role in nonunifilarity of measured quantum processes. In fact, however, allowing for POVMs does not change this aspect. Suppose a particular cCQS hidden state has two outgoing edges with nonorthogonal quantum states ρ_a and ρ_b. In any POVM with two or more measurement operators at least one has a nonzero probability of being an outcome when applied to both ρ_a and ρ_b. This means that in the measured cCQS there are at least two distinct outgoing transitions with the same symbol. And, this again yields nonunifilar dynamics. Section 6.2.1 explores this for an example POVM measurement protocol.

In general, when the quantum states emitted by the cCQS are restricted to qubits, the relatively low dimensionality of the Hilbert space means that we generically recover nonunifilar machines.

This is due to the fact that the only case in which a measurement with two or more operators (not necessarily projectors) can perfectly distinguish between two quantum states is when they are orthogonal. (Distinguishing quantum states here means that none of the operators have nonzero probability to be the measured outcome on both of the quantum states.) And, even in this case, distinguishability holds only for a particular measurement basis that aligns with the two orthogonal states to measure.

The second observation concerns a potentially-useful generalization of how to reduce nonunifilarities in higher dimension. The preceding establishes that irreducible nonunifilarity dominates in measured QSSPs, and this is true generically. However, if the quantum states emitted by the cCQS are qudits, there is more "room" to reduce the nonunifilar transitions in the measured cCQSs when the Hilbert space dimension is larger than $d = 2$. For instance, when the number of outgoing transitions from one hidden state to distinct hidden states is at most d, one can partition the set of quantum states emitted in those transitions into mutually orthogonal subsets. In that case one can devise a measurement that captures each orthogonal subset as a distinct measurement outcome. This, in a sense, constrains the nonunifilarities to be only in the outgoing transitions that have output quantum states with nonzero overlap. Effectively, the measured cCQS loses information about which specific state was output within each orthogonal subset by turning those into nonunifilar transitions, but maintains the information about which output states where mutually orthogonal. This somewhat reduces the complexity of the measured cCQS.

As a simple example, consider a cCQS that outputs qutrits. A particular hidden state has three outgoing edges to distinct states, each outputting qutrits in states $|0\rangle$, $|+\rangle$, and $|2\rangle$. One can devise a measurement that outputs 0 if the qutrit is in the subspace spanned by $|0\rangle$ and $|1\rangle$, and outputs 1 if the qutrit is orthogonal to that subspace. In this case the measured cCQS has nonunifilarity only in the $|0\rangle$ and $|+\rangle$ transitions. This reduction of nonunfilarity is only viable as long as there is an orthogonal subset within the set of possible output states $\mathcal{A}_Q$. So, it is still rare, but the larger the Hilbert space of output states is, the more opportunities there are for reducing nonunifilarity.

Naturally, these measurements should also be physically motivated by the information the experimenter is trying to extract from the underlying quantum dynamic. Thus, when working with quantum information from a classical reality, there is a tradeoff between the complexity of the

observed dynamics and how coarsely or finely one probes the quantum state through measurement. Compared to general qudits, the space of qubits offers much less room for coarser measurements.

The practical upshot of these arguments is that analyzing a measured cCQS requires working with nonunifilar presentations of the observed classical stochastic process.

5.3.4. Explosive Complexity. Consider a process that is generated by a finite nonunifilar presentation. One can construct a unifilar presentation for it. The reasons for doing so are outlined in Sections 3.5 and 3.6: essentially, a unifilar, or predictive presentation is required in order to characterize the stochastic process. The states of this unifilar presentation are Blackwell's *mixed states* [**68**]. These are identified using the *Mixed State Algorithm* introduced in References [**78**, **79**], explained in detail in Reference [**17**], and reviewed in Section 3.5. Said simply, by tracking the "states of knowledge" about an HMM's internal states as revealed indirectly by emitted symbols, one builds a unifilar hidden Markov chain whose states are the mixed states and whose transitions are the mixed-state to mixed-state transitions. The result is known as the process' *mixed state presentation* (MSP). The MSP then provides an insightful and calculationally efficient way to determine many, if not all, of a process' statistical and informational properties.

We restate here the general idea. Given a process' N-state HMM presentation M, one constructs M's set $\boldsymbol{\mathcal{R}}$ of mixed states as the conditional probability distributions $\eta(x_{-\ell:0}) = \Pr(\mathcal{S}_0 = \sigma | X_{-\ell:0} = x_{-\ell:0})$ over the HMM's hidden states $\sigma \in \boldsymbol{\mathcal{S}}$ given all possible sequences $x_{-\ell:0} \in \mathcal{A}^\ell$. Given M and an observed symbol sequence $x_{-\ell:0}$, there is a unique mixed state $\eta(x_{-\ell:0})$ that represents the best guess as to M's current internal state. Moreover, the set of the process' allowed sequences of all lengths $\ell \in \mathbb{N}$ induces a invariant measure μ on the state distribution $(N-1)$-dimensional simplex. We simply denote this as the *mixed state distribution* $\mu(\boldsymbol{\mathcal{R}})$. An HMM's mixed state set $\boldsymbol{\mathcal{R}}$ together with the transition dynamic $\mathcal{W}$ between mixed states induced by observed sequences form the HMM's MSP: $\mathrm{MSP}(M) = \{\boldsymbol{\mathcal{R}}, \mathcal{W}\}$.

Most importantly, by construction an HMM's MSP is a unifilar presentation of the stochastic process generated by the HMM. Additionally, the set of mixed states $\boldsymbol{\mathcal{R}}$ corresponds to the process' set of causal states. The consequence is that the MSP, up to minimizing state redundancies, is the unique optimally predictive model of the stochastic process—its ϵ-machine [**18**].

Conjecture. *A stochastic process generated by a nonunifilar presentation generically is an irreducibly nonunifilar stochastic process. That is, it requires an infinite number of predictive features (causal states) for optimal prediction.*

Blackwell indirectly introduced this conjecture in his seminal 1957 work on classical stochastic processes [**68**]. As stated in Section 3.6, in that work he developed several of the first information-theoretic results for what he called *functions of Markov chains*, what we call HMMs. Moreover, for very specific cases Blackwell showed that the set of (predictive) features that a process stores from observed sequences can be finite or countable. In all other instances, the predictive features set is uncountably infinite. These predictive features are equivalent to the process' MSP mixed states $\boldsymbol{\mathcal{R}}$. The primary lesson is that the predictive complexity of irreducibly nonunifilar processes explodes, despite them being generated by a finite mechanism—a finite-state HMM.

Long experience and extensive explorations of HMM space support Blackwell's claims and this conjecture, which has also been recorded elsewhere [**15**,**17**,**18**,**63**,**69**]. Reference [**19**] goes into great detail about the mechanisms by which these stochastic processes generate and process information. It reviews the arguments and evidence that the conjecture holds quite broadly. Finally, for measured cCQSs we have not encountered a single violation. That said, establishing the conjecture for the general or the quantum settings remain open problems.

5.3.5. Quantum State Indistinguishability. Section 5.3.2 detailed the structural reasons that make measured cCQSs generically nonunifilar. Behind these lies a simple physical property that is responsible for irreducible nonunifilarity and, thus, explosion in complexity of measured quantum processes. When applying a measurement to a QSSP that emits qubits in two or more distinct quantum states, a single measurement will generally have a nonzero probability of not being able to distinguish which quantum state it measured. This indistinguishability between quantum states therefore acts as a source of noise. And, this makes direct reading of the QSSP's underlying structure markedly more memory intensive. This, in turn, radically increases the predictive complexity of the measured process with respect to the QSSP.

One can quantify how distinguishable or indistinguishable two quantum states are using the trace distance [**77**] for instance. If a particular hidden state in a cCQS has outgoing transitions to two

distinct hidden states that emit two different quantum states, a measurement makes the distinction ambiguous (noisy) unless the trace distance between the two quantum states is unity. In that case, the quantum states have orthogonal supports. Moreover, there is the further requirement for not inducing nonunifilarity that the measurement distinguish between the two states. If these criteria are met, then nonunifilarity in the measured process is not created and there is no explosion in predictive complexity. However, these criteria are very restrictive and so explosive complexity is to be expected in measured cCQSs.

5.4. Measured Process Characterization

Simply establishing explosive complexity is insufficient. One needs yardsticks for analysis and comparison. This section makes explicit the use of metrics introduced in Section 3.6 for quantifying randomness and structure in the classical stochastic processes resulting from measured QSSPs. As stated above, the mathematics for these metrics depend critically on the stochastic process presentation, whether it is unifilar or nonunifilar. The latter is particularly relevant, as the above showed that the measured processes are overwhelmingly irreducible nonunifilar. We review how to compute entropy rate and statistical complexity from unifilar HMMs. The bulk of the effort and interest, though, arise in adapting these to nonunifilar presentations, which follows shortly.

5.4.1. Unifilar Generators. When an HMM is unifilar, there is a one-to-one or one-to-finite correspondence between a sequence of observed symbols and the sequence of hidden states that generated it. This allows direct, closed-form calculation of process intrinsic randomness and predictive memory from the HMM's internal Markov chain.

5.4.1.1. *Information Creation.* Process randomness—the rate at which the process generates information—is quantified through the Shannon *entropy rate* h_μ, introduced in Section 3.6.1. It is defined directly for a process, but can be computed in a useful way from a process' presentation. If the measured cCQS is unifilar, the entropy rate is computed as in Equation (3.9).

5.4.1.2. *Information Storage.* The information storage, or memory structure of a process is quantified by its *statistical complexity* C_μ or its *statistical complexity dimension* d_μ, as introduced in Section 3.6.2. For a process with a finite unifilar presentation, C_μ converges to a finite value and

is therefore the correct quantifier. It can be directly computed from a unifilar presentation of the process as in Equation (3.13). C_μ is the minimal memory required to optimally predict the future, or how much information is stored in the *causal states* of a process' ϵ-machine.

5.4.2. Nonunifilar Generators. If the only description available for a measured QSSP is a nonunifilar HMM presentation, quantifying the process' stochasticity and structure becomes markedly more complicated due to the explosive complexity demonstrated above, as described in more detail in Sections 3.3.1, 3.3.2 and 3.6. In the case of intrinsic randomness, Equation (3.9) overestimates the entropy rate. In the case of structure, the Shannon entropy of the nonunifilar HMM's hidden states only quantifies the memory used by that particular (likely nonunique) presentation. More to the point, it does not provide information on how much memory is minimally required to *optimally* predict the process.

There is yet another complication at this stage. While constructing the MSP from a process' nonunifilar HMM produces a unifilar HMM, it is rarely finite-state. As we discussed, the typical case is an infinite-state HHMM, generically with uncountably infinite states [**17**, **63**, **69**]. Generally then, the ϵ-machine, the minimized MSP, has an uncountable set of causal states. As a consequence, the statistical complexity of Equation (3.13) diverges and the expression for entropy rate in Equation (3.9) becomes inadequate.

5.4.2.1. *Information Creation.* If the measured cCQS is nonunifilar, then its MSP must be computed or approximated, and the entropy rate is then computed using Equation (3.11). The latter procedure is significantly more complicated and computationally intensive, as a consequence of the much higher complexity of the stochastic process.

That said, the entropy rate can always be computed and will generically be nonzero. As stated above, the entropy rate quantifies the amount of information per observed symbol that the process creates.

5.4.2.2. *Information Storage.* As discussed above, the statistical complexity C_μ of these infinite-causal-state processes diverges. To track the rate at which this divergence of memory resources required for prediction grows, we use the *statistical complexity dimension* d_μ of the process. This

tracks the rate at which memory requirements for optimal prediction grow with increasing precision of the prediction.

Section 3.6.2 goes into more detail about how this quantity is defined and the complications for its computation. So far, reliable techniques to compute d_μ exist only for stochastic processes that have a nonunifilar presentation with $N = 2$ hidden states. For $N = 3$ states the calculation can be done for certain specific stochastic processes but not in general. These are solvable but algorithmically complicated problems that are left for future work.

All of the quantities above—with the exception of d_μ—can be computed using the QSSP package described in Appendix A. Specifically: the entropy rate, the statistical complexity and the mixed state presentation of a process. For mixed state presentations, handling 10^6 mixed states takes only about 10 minutes in a single CPU.

To give a firmer, even visual, grounding to the preceding results and metrics, the next section explores three examples representative of distinct classes of measured cCQSs and how the above metrics characterize them.

5.5. Classifying Measured Quantum Processes

The metrics for randomness and structure of a measured quantum process depend on the cardinality of the mixed state set $\boldsymbol{\mathcal{R}}$ generated by the measured cCQS. There are three distinct classes: processes for which the number of mixed states is finite, countably infinite, and uncountably infinite. The following examples illustrate processes in these classes.

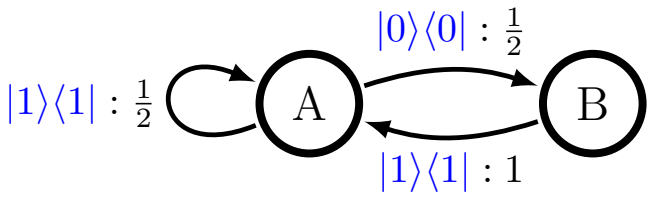


FIGURE 5.1. Unifilar presentation for the *Observation Basis Golden Mean (OB-Golden Mean) process*: A simple cCQS.

5.5.1. Finite-State. The first quantum process is generated by the unifilar cCQS shown in Figure 5.1. It consists of all random sequences without consecutive $|0\rangle\langle 0|$s. Measuring in the

observation basis $E_0 = |0\rangle\langle 0|$ and $E_1 = |1\rangle\langle 1|$ yields a unifilar HMM that generates the *Golden Mean Process* consisting of all random sequences without consecutive 0s. Figure 5.2 shows its minimal presentation—its ϵ-machine: a unifilar HMM with two states. Being unifilar one readily calculates that it has an entropy rate of $h_\mu = 2/3$ bits/symbol from Equation (3.9) and a statistical complexity of $C_\mu = 0.918$ bits from Equation (3.13).

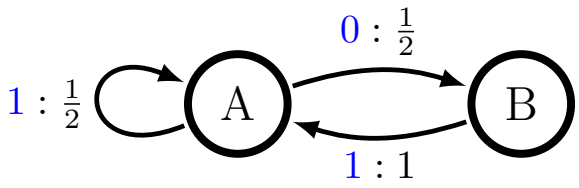


FIGURE 5.2. Measured cCQS of the stochastic process resulting from measuring the quantum process generated in Figure 5.1 in the observation basis.

Although unnecessary in this case, computing the MSP of this presentation—or any other finite unifilar HMM, for that matter—results in an HMM with a finite number of states. In the present case both the measured process' entropy rate and the statistical complexity are finite. They are readily computed via Eqs. (3.9) and (3.13), respectively.

Section 5.3.2 showed that quantum processes in this class are relatively rare in the space of measured cCQSs. They occur only under very constrained circumstances. This observation will become clearer as we consider more complex classes.

5.5.2. Countably-Infinite-State. The next quantum process is generated by the cCQS in Figure 5.3. This is a seemingly slight variation on the previous example. Now, the quantum alphabet $\mathcal{A}_Q$ consists of nonorthogonal states. Instead of emitting quantum states in the observation basis, this cCQS emits qubits in state $|0\rangle\langle 0|$ and others in state $|+\rangle\langle +|$. In this, we define $|+\rangle$ and $|-\rangle$ in the conventional way: $|\pm\rangle = (1/\sqrt{2})(|0\rangle \pm |1\rangle)$. When the process generated by this cCQS is measured in the basis $E_0 = |+\rangle\langle +|$ and $E_1 = |-\rangle\langle -|$, the measured cCQS has the HMM presentation shown in Figure 5.4.

Notice, though, that Figure 5.4's measured cCQS is nonunifilar. Specifically, knowledge of being in state A and emitting a 0 does not determine the next HMM state. The next state could be either A or B. Thus, to compute the entropy rate for this process one must construct its MSP. The latter is

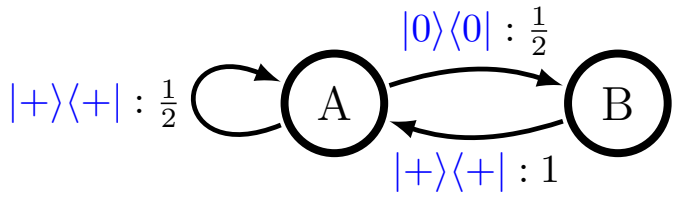


FIGURE 5.3. Structurally, this cCQS is similar to that in Fig. 5.1. However, not all emitted quantum states are orthogonal. This guarantees that the measured process is more complex, as Fig. 5.4 shows.

shown in Figure 5.5. It has a countable infinity of causal states. Helpfully, as annotated there, the state transition probabilities can be parametrized analytically.

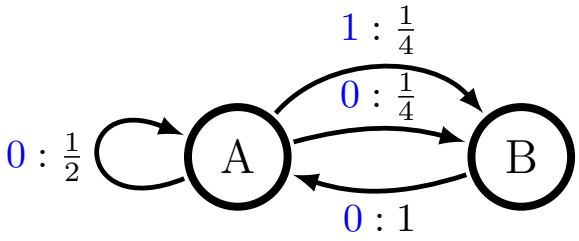


FIGURE 5.4. Measured cCQS for the process generated by measuring the quantum process generated by the cCQS in Fig. 5.3. This HMM is nonunifilar: if in state A and emitting a 0, the next hidden state may be A or B.

Using Figure 5.4 one can follow the logic for constructing the MSP. Independent of any knowledge of the HMM state, seeing symbol 1 the observer concludes with absolute certainty that the measured cCQS is in state B. This is what we referred to previously as a state of knowledge (or a mixed state) represented by hidden state I in Figure 5.5. In point of fact, the mixed state associated with state I is $\eta(1) = (0, 1)$. After that, observing symbol 0 or a sequence of 0s means that the measured cCQS has a certain probability of being in each cCQS state A or B. Each additional observation of a 0 then updates the present state of knowledge to one of the mixed states II $= \eta(10)$, III $= \eta(100)$, IV $= \eta(1000)$, ... depending on how many 0s are observed before seeing a 1, when the MSP resets to state I $= \eta(100 \ldots 01)$.

The measured process' entropy rate can be computed from the HMM in Figure 5.4 using the methods for nonunifilar HMMs described in Section 5.4.2. Note, though, that for processes whose MSP has a

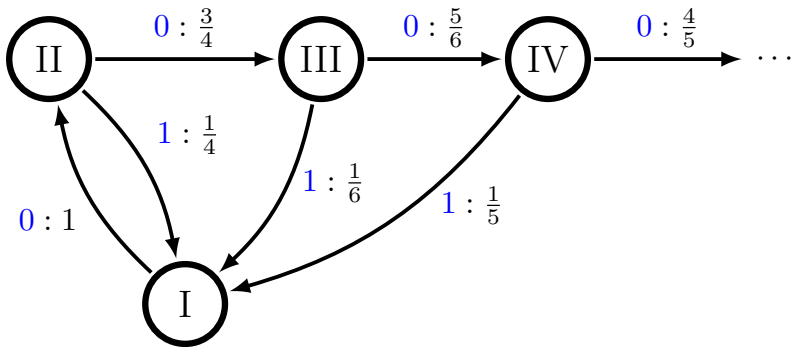


FIGURE 5.5. Mixed state presentation of the process generated by Fig. 5.4's measured cCQS.

countable infinity of states, as here, a more rudimentary, though convergent and accurate, approach is available.

When observing the stochastic process, the probability of observing consecutive 0s diminishes with the length of the observed sequence. One then approximates the process' HMM by truncating the MSP at a finite number N of mixed states and then exploring the limiting behavior of both h_μ and C_μ from those unifilar machines as $N \to \infty$. For the example in question, this analysis is illustrated in Figure 5.6. One finds that $h_\mu = 0.599$ bits/symbol and $C_\mu = 3.69$ bits. Note that, although infinite state, the process statistical complexity is finite. This is due to the fact that the asymptotic state distribution π decays exponentially fast for mixed states reached via increasingly more 0s.

5.5.3. Uncountably-Infinite-State. The preceding two processes are relatively simple, in that they all exhibit a finite or countable set of mixed states. In the typical case, as argued in Sec. 5.3.1, the measured cCQS has an HMM presentation that is nonunifilar and an MSP with an uncountable infinity of states. Section 5.3.2 established that this is the typical case for processes generated by cCQSs of two or more states.

To illustrate, consider the cCQS of Figure 5.7, chosen to have three states principally to aid visualizing the MSP's complexity. The cCQS is then measured in the observation basis, which yields the measured cCQS of Figure 5.8.

Note that, as in the example of the countably-infinite state process, the measured cCQS has only a single source of nonunifilarity: the successor state is ambiguous when observing symbol 0 with the HMM in state A. More generally, however, none of the symbols 0 or 1 allow the observer

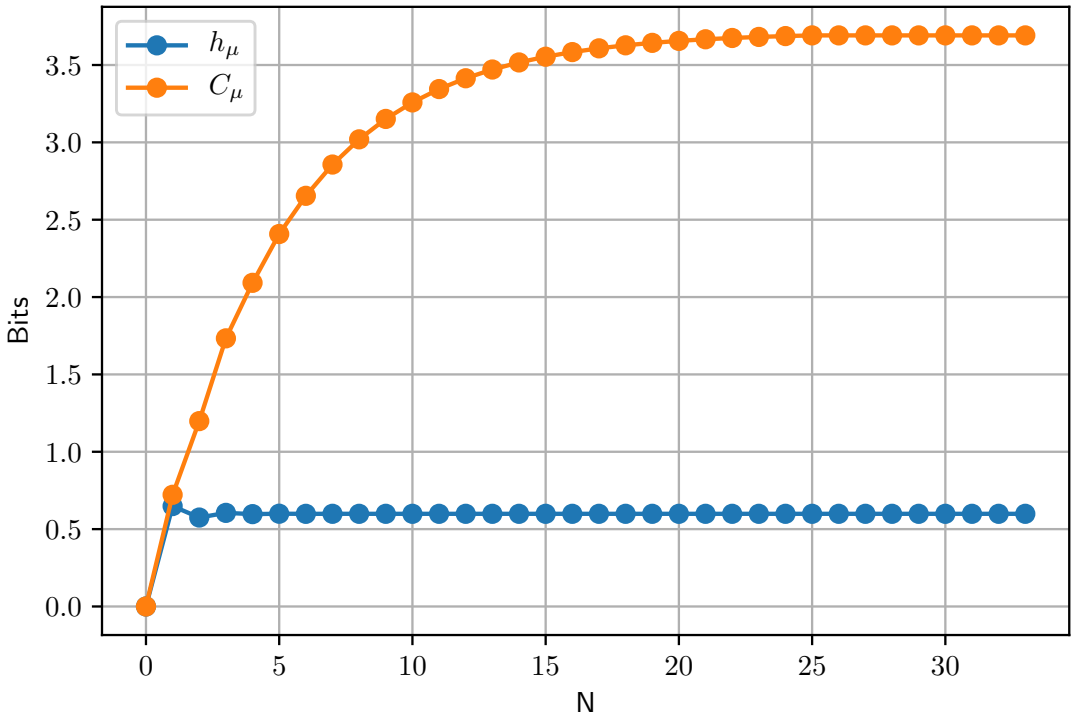


FIGURE 5.6. Entropy rate h_μ (blue) and statistical complexity C_μ (orange) of N-state HMM approximations of the MSP shown in Fig. 5.5. Notice that h_μ converges rapidly, while C_μ has a stronger dependence on the number of states, but stabilizes around $N = 25$. The values obtained are $h_\mu = 0.599$ bits/symbol and $C_\mu = 3.69$ bits.

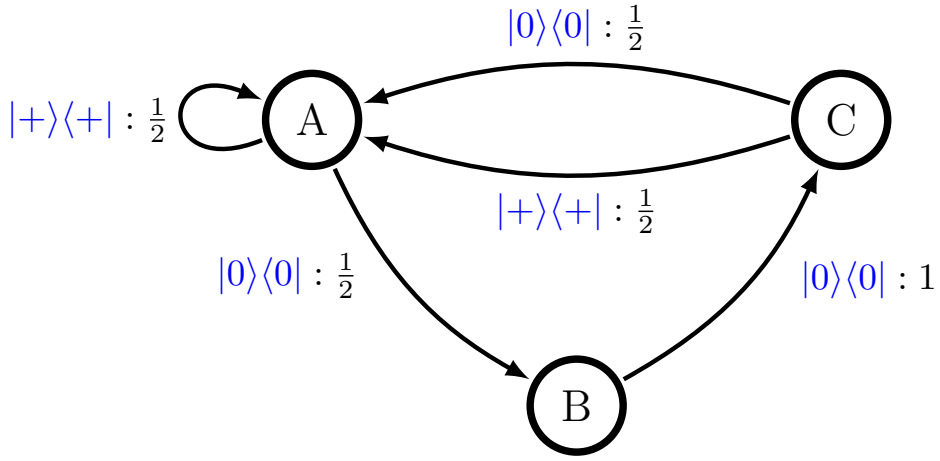


FIGURE 5.7. Nonorthogonal Nemo Quantum Process: Three-state cCQS that emits qubits in states $|0\rangle\langle 0|$ and $|+\rangle\langle +|$.

to "synchronize" to the process. That is, observation of a particular symbol does not give an observer certainty in the measured cCQS's state. As argued above mathematically and as is now constructively clear in Figure 5.9, this effectively translates into the fact that the MSP of the measured quantum process has an uncountable infinity of mixed states. The MSP—these states together with their transition probabilities—are a markedly less tractable presentation than in the previous two quantum processes.

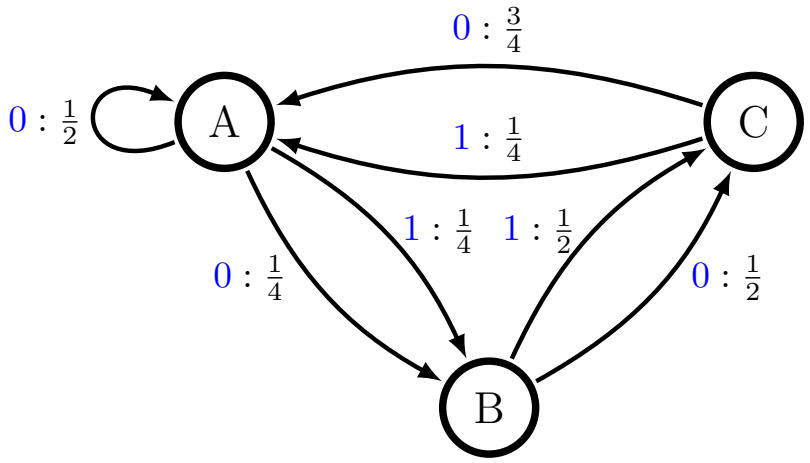


FIGURE 5.8. Measured cCQS presentation of the stochastic process produced when measuring the quantum process generated by Fig. 5.7's cCQS measured in the observation basis.

The MSP with all of its states and state transitions cannot be explicitly displayed as with the previous HMMs. Nonetheless, Figure 5.9 gives a sense of the MSP's structure and complexity. It presents a plot of 2×10^6 MSP states in the mixed-state simplex $\boldsymbol{\mathcal{R}}$. In fact, it shows $\mu(\eta)$ and its variation in probability density via a histogram with a coarse-graining of 1000×1000 bins.

The measured process' entropy rate is computed using Equation (3.11) and has a value of $h_\mu = 0.8896$ bits/symbol. The statistical complexity dimension d_μ of Equation (3.14) is computed as described in Refs. [**18**, **19**]: $d_\mu = 1.38$.

5.5.4. Remarks. As seen from the three examples above, the cardinality of the mixed state set of the distinct measured stochastic processes can vary from a finite state set to a countable infinity of states and on to an uncountable infinity of states. This cardinality affects the way in

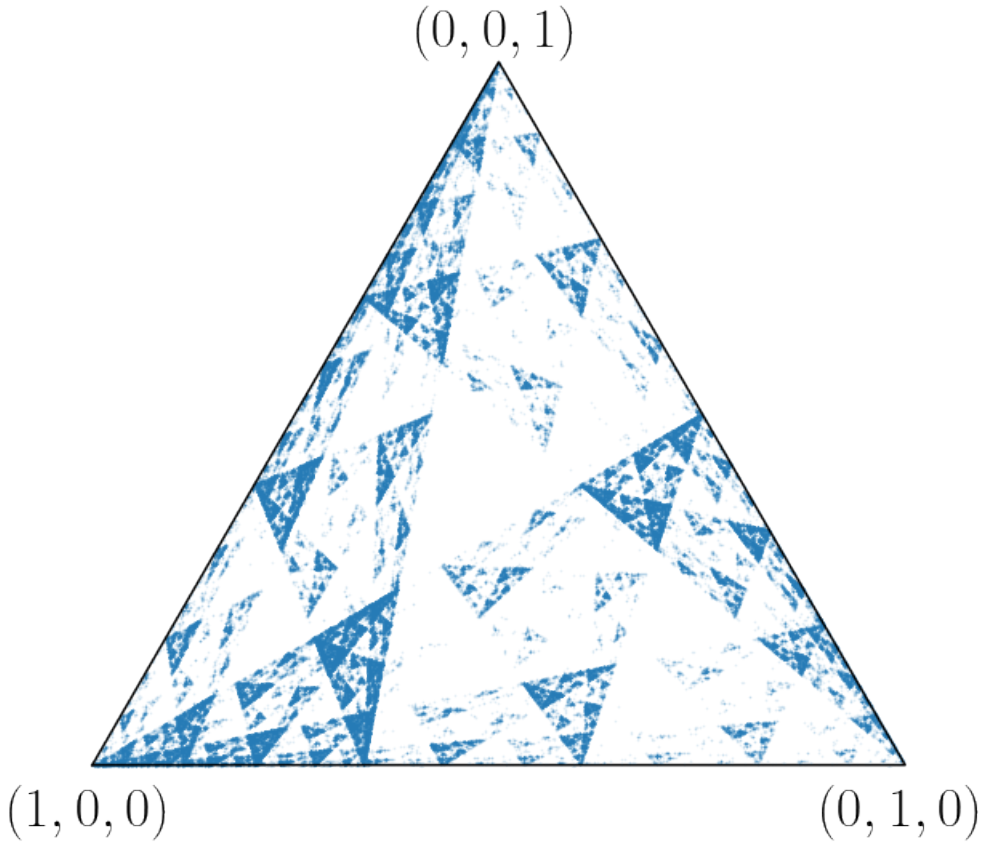


FIGURE 5.9. MSP's asymptotic invariant measure $\mu(\eta)$ in the mixed-state simplex $\mu(\eta)$. Each mixed state is a point of the form (p_A, p_B, p_C) with p_σ the probability of being in state σ of the measured cCQS in Fig. 5.8.

which the metrics of randomness and structure for the process are computed, but also the values they can take.

For processes with finite sets, the statistical complexity and the entropy rate will generally be finite positive values. This implies that these processes have a certain degree of stochasticity, but that they can be optimally predicted with finite memory resources.

Notably, excepting very special cases, this is also true for processes whose mixed state set has a countably-infinite number of states, as in the second example. These processes have positive entropy rate, signaling that they have an intrinsic degree of randomness. And, while they do require an infinite number of causal states to optimally predict, these states are structured such that one can simulate an optimal predictor of arbitrary precision with a finite amount of memory.

The third case is significantly more complicated than the previous two. A mixed-state presentation with an uncountable infinity of states implies that the statistical complexity of the process diverges. This means that it takes infinite memory to optimally predict these processes. That said, there is an asymptotic invariant measure over the mixed states in $\boldsymbol{\mathcal{R}}$. And, by being able to compute these measures, one can then estimate the process' entropy rate h_μ and also the growth rate d_μ of the memory required for optimal prediction.

This third case, of processes with MSPs that have an uncountably infinite number of states, turns out to be the norm for MSPs of measured cCQSs, as argued above and as we elaborate shortly below. The implications of this are that, in general, the classical processes that we recover from measuring QSSPs generated by cCQSs are highly complex and require infinite memory for optimal prediction. That is, measuring a QSSP greatly obscures the underlying quantum stochastic process. Fortunately, we have metrics to characterize these processes and to develop a quantitative understanding of how measurement affects the measured quantum processes.

5.5.5. Genericity of Complexity. The tools are in place now to quantitatively analyze the measured QSSPs that are represented by measured cCQSs. Here, we use the tools to draw broader conclusions about what one should expect and how measurement choice changes the randomness and complexity of measured quantum processes.

The main lesson from the preceding is that one expects explosive complexity and this is reflected in the information-theoretic metrics of the measured process.

Proposition 4. *A measured quantum process, with a measured cCQS presentation, generically is highly complex in two specific ways: it has nonzero entropy rate and statistical complexity dimension. That is, it requires uncountably infinite states to optimally predict.*

Proof. *This follows as a corollary of Sec. 5.3.1's structural propositions—specifically Props. 2 and 3 and Sec. 5.3.4's conjecture—though translated into the information metrics of Sec. 5.4.*

As discussed above and extensively in Refs. [**17**–**19**, **69**], nonunifilar HMMs lead to causal state sets of uncountably infinite cardinality and divergent statistical complexity. As the preceding demonstrated, measured quantum processes have presentations that fall into this class.

CHAPTER 6

Measurement Dependence

Having laid out the progression from quantum sources to quantum state processes and their presentations to measured processes and their metrics, we are now ready to illustrate uses and benefits. The following does these via three applications: measurement choice, alternate measurement protocols, and optimal measurements.

6.1. Measurement Variation and Choice

Equations (5.4) and (5.6) directly show that choice of measurement basis changes the observed process. This, in turn, means that a process' entropy rate and its MSP's statistical complexity dimension also depend on measurement choice. Fortunately, the changes are well behaved.

Conjecture. *Measured process complexity depends piecewise smoothly on both the underlying QSSP and choice of measurement.*

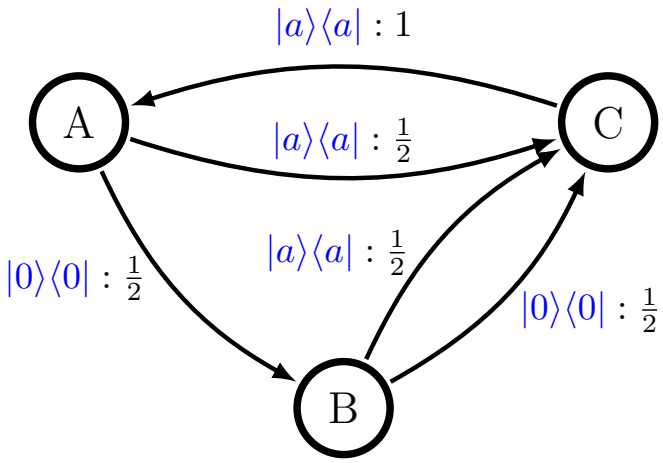


Figure 6.1. cCQS that generates a quantum process with qubits in quantum states $|0\rangle$ and $|a\rangle = \cos \pi/5 \, |0\rangle + \sin \pi/5 \, |1\rangle$.

REMARK. *Given the extensive development up to this point, the following refrains from presenting formal proofs. These will appear elsewhere. Nonetheless, it is worthwhile to illustrate how the results can be used to outline a construction that supports observed behavior and is backed by formal proofs in parallel problem settings.*

Note that:

(1) The measured process' entropy rate and statistical complexity dimension depend smoothly on its MSP's invariant measure, as can be seen from Eqs. (3.10) and (3.14).

(2) Equations (5.4) and (5.6) state that the parameters (transition probabilities) of the measured cCQS HMM depend smoothly on the underlying QSSP and measurement operator parameters.

Therefore, if the MSP's invariant measure depends smoothly on the parameters of the measured cCQS HMM, then the entropy rate and statistical complexity dimension of the measured process depend smoothly on the underlying QSSP and measurement parameters.

Smoothness dependence of the MSP's invariant measure with respect to HMM parameters is not only consistent with observation, which is illustrated shortly, but has been established for many classes of iterated function system *(IFS). For more detail, Reference [**17**] outlines how any HMM can be cast as an IFS—a stochastic dynamical system with a unique attractor (equivalent to an HMM's MSP) that has an invariant measure. Both the attractor and the invariant measure vary smoothly as a function of IFS parameters under contractivity conditions [**80**–**82**]. These conditions are generally satisfied by HMMs, thus indicating that both the MSP and invariant measure of an HMM depend smoothly on the HMM parameters. The caveat of piecewise smoothness as opposed to smoothness stems from the fact that the MSP can have abrupt jumps in cardinality for a finite set of parameters, potentially causing finite discontinuities in the statistical complexity dimension, as will be illustrated in the examples of the following section.*

Beyond smooth dependence, we ask more specifically, How do the mixed-state invariant measure and the associated complexity measures change as a function of the measurement angles θ and ϕ? To answer these questions, we explore two specific examples. For each we choose a quantum process generated by a particular cCQS. We then obtain the measured cCQSs resulting from measuring the

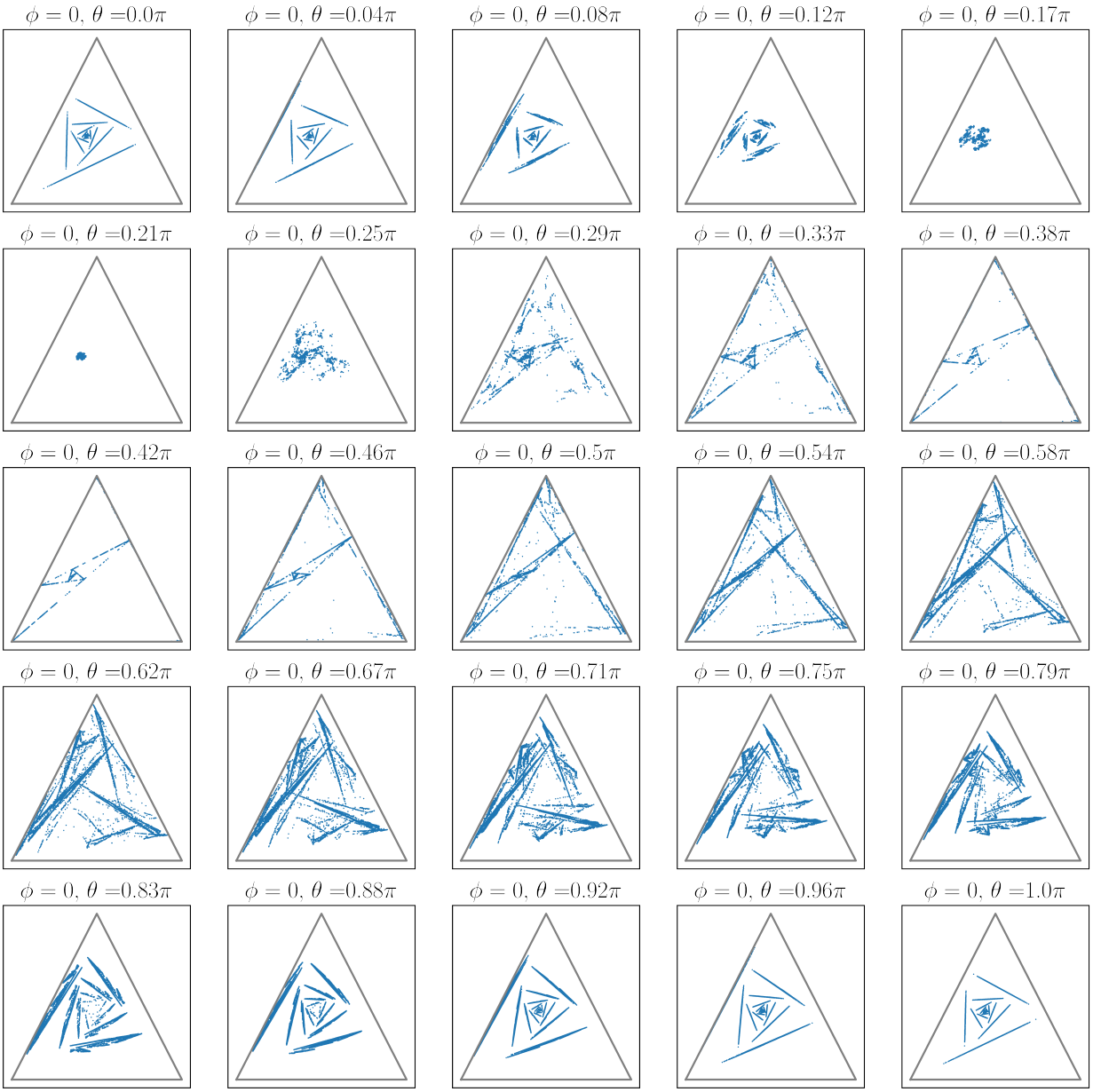


FIGURE 6.2. Mixed-state presentation of the process resulting from measuring the quantum process generated by the cCQS in Fig. 6.1 in the measurement bases parametrized by (ϕ, θ), as indicated in each subfigure. For each value of the parameters 25000 mixed states are plotted.

quantum process with bases in which one of the angles is held fixed and the other sweeps across its range of possible values.

The two example processes below were chosen since together they illustrate the general properties of measurement dependence of QSSPs. The first example is the three-hidden-state cCQS depicted in Figure 6.1, which is then measured in many different qubit bases, holding $\phi = 0$ and varying θ. The second example is the two-hidden-state quantum process generated by the cCQS in Figure 5.3, which is then measured following the same procedure.

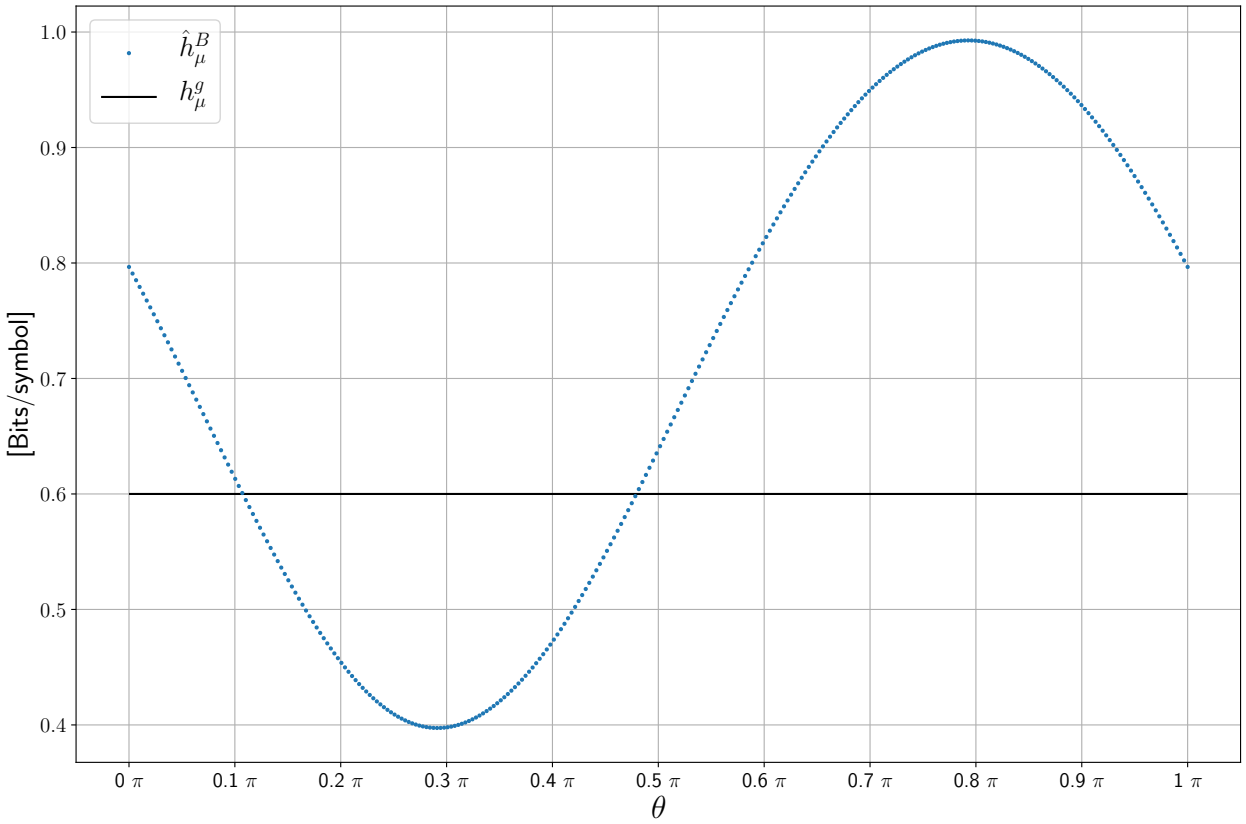


FIGURE 6.3. Entropy rate of the measured cCQSs resulting from measuring the quantum process generated by the cCQS of Fig. 6.1 as a function of measurement angle θ, as in Eq. (5.4) at 300 θ values with the value $\phi = 0$ fixed. Entropy rate h_μ^g (black line) of the cCQS that generates the measured process.

6.1.1. Random Insertion Process. First, we track changes in the invariant measure on the mixed-state simplex. Figure 6.2 shows these for the measured process generated by the cCQS of Figure 6.1 in 10 different measurement bases (ϕ, θ), as noted there. The structure of the invariant sets $\boldsymbol{\mathcal{R}}$ varies substantially with measurement basis. For most (but one, discussed below) measurement

bases the set has an uncountable infinity of states, yet these states have distinct structures that vary smoothly with choice of measurement basis.

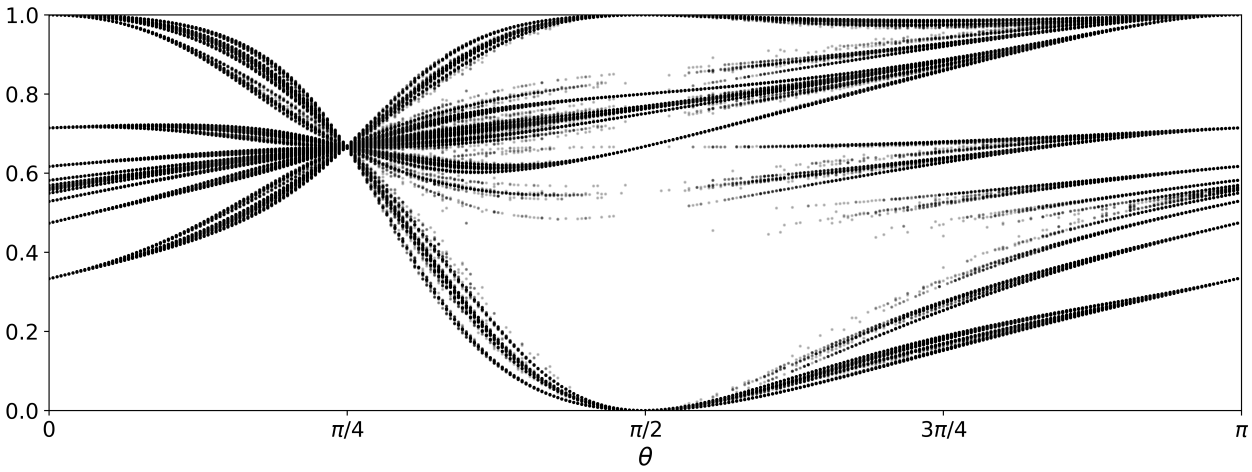


FIGURE 6.4. Mixed-state presentations of the processes resulting from measuring the quantum process generated by the cCQS in Fig. 5.3 in the bases parametrized by $(\phi = 0, \theta)$. Each vertical line represents the 1D simplex $\boldsymbol{\mathcal{R}}$ and the points in it are the mixed states corresponding to the measured cCQS at that particular value of θ.

Second, we determine the entropy rate as a function of measurement basis in Figure 6.3. The process is measured in 300 different bases, holding the value of $\phi = 0$ and varying $\theta \in [0, \pi]$. By comparing to the entropy rate h_μ^g of the cCQS that generates the underlying QSSP, Figure 6.3 clearly demonstrates that measurement both increases and decreases the randomness (h_μ).

While this example serves to graphically illustrate the high complexity of predicting the classical stochastic processes measured from the QSSP, there are limitations to estimating the MSP's statistical complexity dimension. For reasons explained in detail in Reference [**18**], estimating the statistical complexity dimension for measured cCQSs with MSPs in two and higher dimension simplices is computationally intensive and there is as yet no efficient algorithm. To illustrate the behavior of the statistical complexity dimension in these stochastic processes, though, we turn to an example of a QSSP generated by a two-state cCQS with an MSP in the 1-simplex.

6.1.2. Golden Mean Process. This example analyzes the QSSP generated by the cCQS in Figure 5.3. It is measured in different qubit bases holding $\phi = 0$ fixed and varying θ uniformly from 0 to π. Each measurement yields a measured cCQS that is nonunifilar, the MSP is then computed.

Figure 6.4 displays its invariant measures. Each vertical unit interval corresponds to a 1-simplex that shows the MSP at that particular value of θ. From the figure we observe that the majority of the MSPs have a complex fractal-like structure. However, what the figure makes evident is that this structure varies smoothly with respect to the measurement parameter θ, consistent with Sec. 6.1's Conjecture.

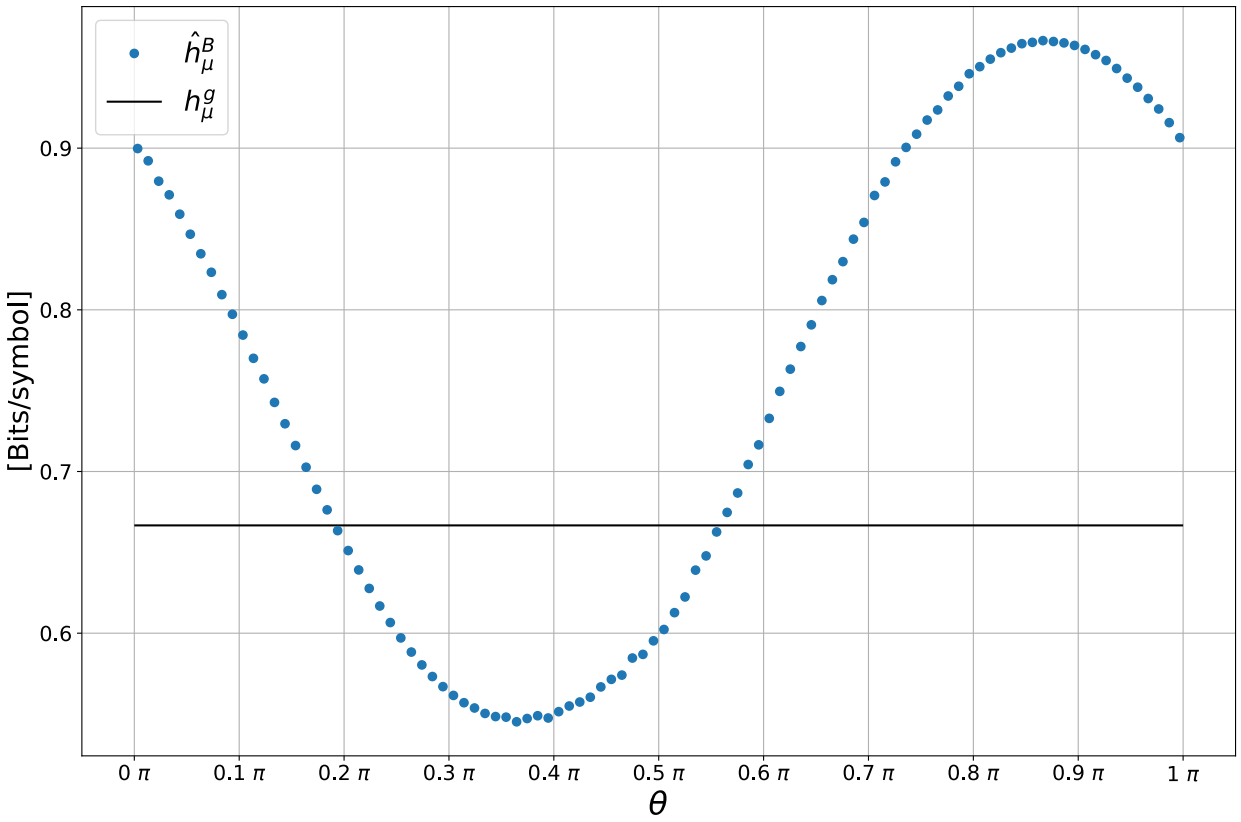


FIGURE 6.5. Entropy rate of the measured cCQSs resulting from measuring the quantum process generated by the cCQS of Fig. 5.3 as a function of measurement angle θ, as in Eq. (5.4) with the value $\phi = 0$ fixed, at 100 θ values. Entropy rate h_μ^g (black line) of the cCQS that generates the underlying QSSP.

Figures 6.5 and 6.6 track how both entropy rate h_μ and statistical complexity dimension d_μ vary with respect to measurement basis.

As with the previous example, we see that h_μ of the measured process both increases and decreases with respect to h_μ^g—that of the original QSSP—depending on measurement basis.

The statistical complexity diverges for most, in contrast with the finite statistical complexity of the underlying QSSP. That said, the statistical complexity dimension d_μ smoothly varies. To a

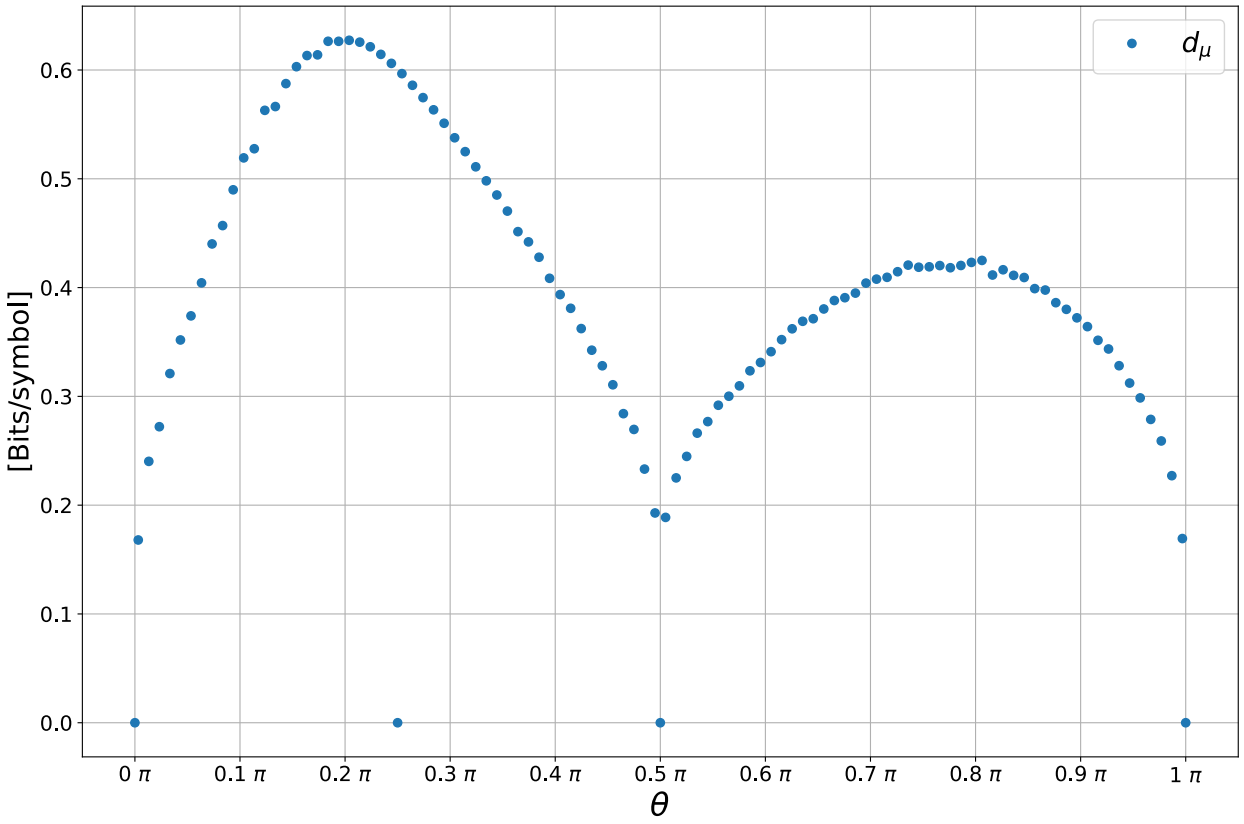


FIGURE 6.6. Statistical complexity dimension d_μ of the measured cCQSs resulting from measuring the quantum process generated by the cCQS of Fig. 5.3 as a function of measurement angle θ, as in Eq. (5.4) with the value $\phi = 0$ fixed, at 100 θ values.

certain extent this reflects what is see from the MSPs in Figure 6.4. Figure 6.6 also reveals four values of θ for which $d_\mu = 0$. These are are $\theta \in \{0, \pi/4, \pi/2, \pi\}$. When θ takes the values 0 or π the measurement is in the observation basis and one of the measurement operators aligns with the quantum state $|0\rangle$. This simplifies the process and thus the measured cCQS has a countably infinite number of mixed states. This also happens at the value of $\theta = \pi/2$, in which one of the measurement operators aligns with the quantum state $|+\rangle$ that is output by the cCQS. Notice that in these three cases, $d_\mu = 0$ is reached smoothly.

The exception to smoothness is the discontinuous jump to $d_\mu = 0$ when the process is measured at $\theta = \pi/4$. This special case is discussed more, shortly. In general terms, though, for that particular basis the measurement does not distinguish between the two distinct emitted qubit states $|0\rangle$ and $|+\rangle$. And so, all of the structural information about the underlying quantum process is lost, except

for the probability of obtaining one measurement outcome or the other. The process becomes memoryless and so has a single-state presentation.

6.1.3. General Features. With the previous two examples in hand, and after an exhaustive exploration of example processes in this fashion, we review several common characteristics. Of particular interest are the smooth behaviors of h_μ and d_μ with well defined maxima and minima. It is also apparent that the MSP invariant sets exhibit marked structural variations. However, in agreement with Sec. 6.1's Conjecture, they appear to vary smoothly with respect to measurement change.

A feature that immediately warrants attention in Figure 6.2 is the drop in structural complexity of the MSP at $\theta = \pi/5$. With that particular measurement basis, the statistical complexity dimension vanishes, indicating that the measured cCQS is finite. On closer inspection, the HMM corresponding to the measured cCQS is not only finite, but has a single causal state. This indicates that the measured process consists of independent identically distributed (i.i.d.) random variables. At each time step, the observed symbols are 0 with probability $p_0 = \cos^2 \pi/5$ and 1 with probability $p_1 = \sin^2 \pi/5$. This seemingly special case is not a fluke.

PROPOSITION 5 (Memoryless measurements). *For any cCQS with quantum alphabet $\mathcal{A}_Q$ consisting of two distinct quantum states ρ_a and ρ_b, there exists a set of measurement bases for which the resulting measured process is memoryless and $C_\mu = 0$.*

PROOF. *We establish this by construction. Without loss of generality and for ease of notation we align both quantum states with the zx-plane, such that one of the quantum states ρ_a is at the top of the Bloch sphere. We further denote the angle between the two states by α. We then write $\rho_a = |0\rangle\langle 0|$, and $\rho_b = |b\rangle\langle b|$ such that $|b\rangle = \cos\alpha/2\,|0\rangle + \sin\alpha/2\,|1\rangle$.*

Consider the projective measurement bases for which one measurement operator projects onto a state $|\psi_0\rangle$, such that $|\langle\psi_0|0\rangle| = |\langle\psi_0|b\rangle|$. That is, $|\psi_0\rangle$ lies in the Bloch sphere circumference that bisects the angle between $|0\rangle$ and $|b\rangle$. Then, the set of measurements that project onto $|\psi_0\rangle$ and $|\psi_0^\perp\rangle \equiv |\psi_1\rangle$ are such that the probability distributions over measurement outcomes are the same whether the measured qubit state was ρ_a or ρ_b. That is, for this particular set of measurements, we have $\Pr(i|\rho_j) = p_i$ for $i \in \{0, 1\}$ and for all ρ_j with p_i a constant and $p_0 + p_1 = 1$.

Together with Equation (5.6), this observation says that the measured cCQS transition matrices are, for $i \in \{0, 1\}$:

$$\begin{aligned} T^i &= \mathbb{T}^{\rho_a} Pr(i|\rho_a) + \mathbb{T}^{\rho_b} Pr(i|\rho_b) \\ &= p_i(\mathbb{T}^{\rho_a} + \mathbb{T}^{\rho_b}) \\ &= p_i \mathbb{T} \ . \end{aligned} \tag{6.1}$$

That is, both labeled transition matrices are proportional to each other and to $\mathbb{T}$. Note also that $T = T^0 + T^1 = \mathbb{T}$, so for simplicity we refer to the internal Markov chain transition matrix as T.

Both labeled transition matrices being proportional to T implies that the MSP yields a biased coin process with biases p_0 and p_1, respectively. There is a single recurrent mixed state, namely π. This follows by definition, since π is an eigenvector of T. And so, evolving the mixed state $\eta_t = \pi$ gives:

$$\begin{aligned} \eta_{t+1} &= \frac{\pi \cdot T^i}{p_i} \\ &= \pi \cdot T \\ &= \pi \ . \end{aligned}$$

Physically, memoryless measurements project states onto a basis whose components are symmetric with respect to the pure states in $\mathcal{A}_Q$. Consequently, the measurement cannot distinguish between the pure states and so the act of measurement effectively leads to a complete loss of information about the cCQS's internal structure.

The fact that these memoryless measurements maximize the loss of information about the cCQS's internal structure, naturally leads to the question of whether there exists a set of measurements that maximally preserves information about the cCQS's internal structure. These would be measurements that optimally distinguish between the quantum states. In the case of POVMs these measurements are well studied for the case of distinguishing between 2 or 3 states. And, as explored in Sec. 6.2.1 they in fact yield special measured processes.

Shortly, we return to explore the issue of optimal and extremizing measurement bases.

Another point to make is that in all of the examples above, only the θ parameter of the measurement bases was varied, and the phase ϕ was fixed to zero. This choice was for simplicity and visualization purposes only. The variation of ϕ does not change the analysis or the conclusions in any way. To clearly illustrate this, Figure 6.7 shows the invariant measure of the MSPs obtained by measuring the cCQS in Figure 6.1 while holding parameter $\theta = \pi/2$ fixed, and varying the values of ϕ from 0 to π. The values of ϕ from π to 2π are skipped because they are redundant in this case. Essentially, Figure 6.7 serves the same purpose as Figure 6.2, but it illustrates the smooth variation of the MSP as we vary parameter ϕ as opposed to parameter θ. As expected, we observe some repeated patterns as we vary ϕ, with a symmetry around $\phi = \pi/2$. This is a consecuence of the geometric symmetry of the measurement bases around this value with respect to the output qubits of the underlying cCQS, $|0\rangle$ and $|a\rangle$, which lie along the $\phi = 0$ meridian. Also as expected $\phi = \pi/2$ and $\theta = \pi/2$ is a *memoryless measurement* for this process, since the measurement basis is agnostic to the output quantum states in the sense that it has the same probability of measuring a 0 or a 1 regardless of which output quantum state ($|0\rangle$ or $|a\rangle$) is being measured.

An additional notable point here is that when the measurement parameters both define a measurement basis that is not close to being aligned with any of the output quantum states and is poor at distinguishing between them, then that will in general result in processes that have h_μ larger than the underlying cCQS and, as is generally the case, divergent memory.

Appendix B graphically demonstrates this with two animations that sweep the angles θ and ϕ while monitoring entropy rate and mixed states. One animation shows how the mixed state presentation and h_μ vary a function of parameter θ. The other animation shows $h_\mu(\theta)$ plots as in Figure 6.3 while sweeping ϕ from 0 to 2π.

In general, as seen from Figure 6.3, different choices of measurement increase or decrease the randomness of the measured quantum process. Furthermore, even if the general case is that quantum measurement dramatically increases the structural complexity of the measured stochastic process with respect to the underlying quantum process, the existence of memoryless measurements shows that particular choices of measurement, in fact, can mask the quantum process' structural complexity.

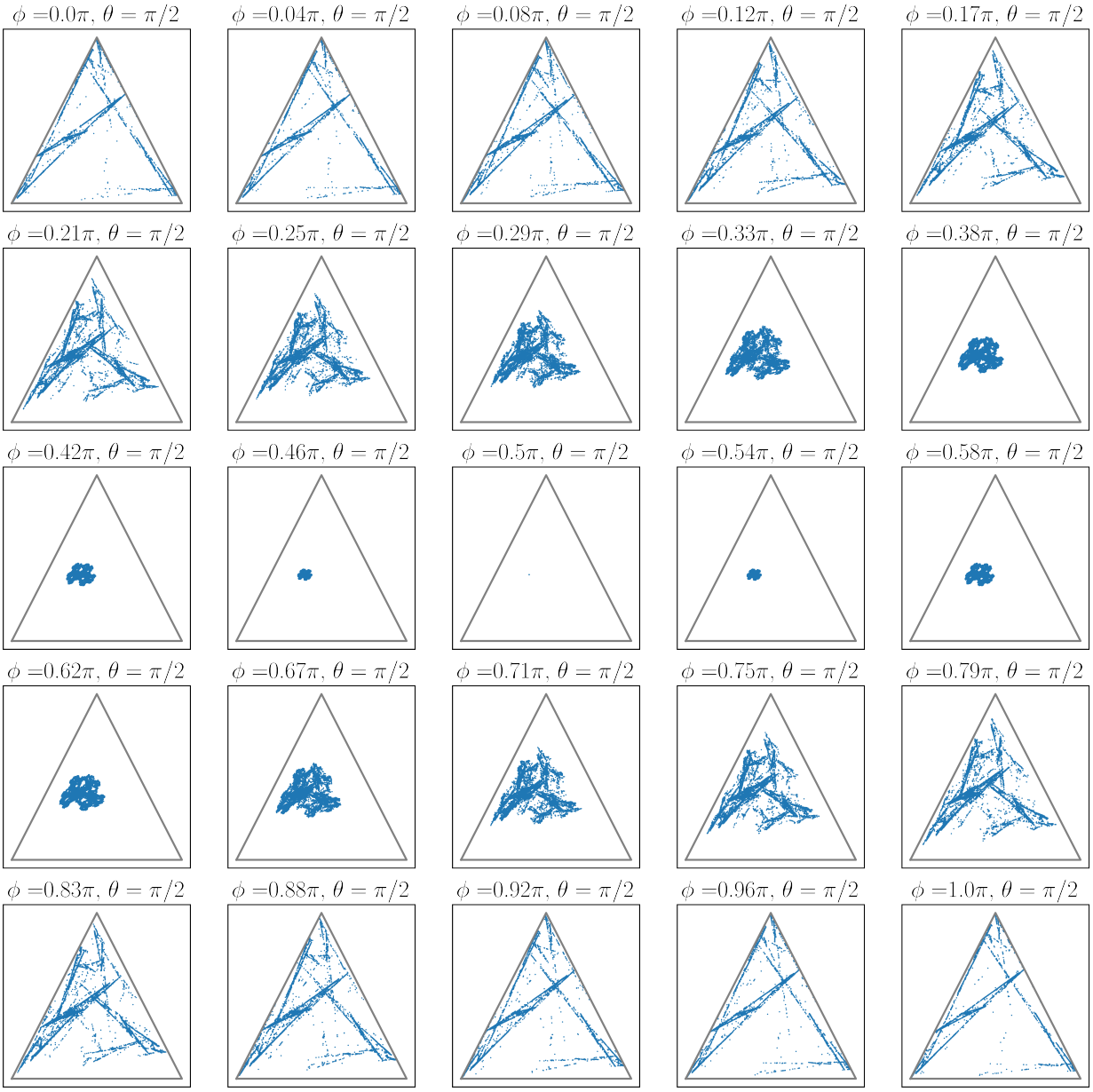


FIGURE 6.7. Mixed-state presentation of the process resulting from measuring the quantum process generated by the cCQS in Fig. 6.1 in the measurement bases parametrized by (ϕ, θ), as indicated in each subfigure. For each value of the parameters 25000 mixed states are plotted.

6.2. Alternate Measurement Protocols

While simplicity dictated that the preceding concentrate on protocols in which the same projective measurement is applied at every time step, there are many alternative protocols to explore. As an

example, the following considers processes that result from applying more general measurements to each emitted qubit.

6.2.1. Positive Operator-Valued Measurements. The development to this point investigated the consequences for the observed classical stochastic process of employing only projective measurements. However, the natural generalization is to more flexible positive operator-valued measurements.

Definition 21 (Positive Operator-Valued Measurement on a QSSP). *A* positive operator-valued measurement *(POVM) $\mathcal{I}$, consists of a finite set of positive semi-definite operators $\{E_x\}$, on the Hilbert space $\mathcal{H}^d$ of dimension d. The operators satisfy the condition $\sum_x E_x = \mathbb{I}_d$. When measurement $\mathcal{I}$ acts on a quantum system in state ρ_t, emitted by a QSSP, the outcome is x_t, corresponding to operator E_x, with probability:*

$$\Pr(x_t|\rho_t) = \mathrm{Tr}(\rho E_{x_t}) \ .$$

Applying a POVM to every quantum state emitted by the QSSP yields a classical stochastic process over the values of $x \in \mathcal{A}_M$—the alphabet of the measured process.

When the measurement $\mathcal{I}$ consists of a POVM, the number of operators $\{E_j\}$ and possible outcomes can be any positive integer. This increase in possible measurement outcomes results in a larger alphabet for the classical measured quantum process. At first glance, this suggests finding a wider range of unifilar measured HMMs, but it is not. This is a direct result of indistinguishability.

Generally, when performing a POVM on any two qubit states (even distinguishable ones), at least one of the measurement outcomes has a nonzero probability of being observed on both of the quantum states. This is due to the fact that POVMs generally have nonorthogonal measurement operators.

When measuring with a POVM, consider a cCQS hidden state with two outgoing transitions on distinct quantum states. Applying the POVM on those transitions means that at least one of the symbols in the classical alphabet $\mathcal{A}_M$ is present in two outgoing transitions for the same hidden state in the measured cCQS. This makes the dynamic of the measured cCQS nonunifilar. Thus, in

general when using POVMs, measured processes are also highly complex, akin to those obtained when applying projective measurements.

That said, the measured processes produced using POVMs reveal a new collection of notable special cases, which remain to be broadly explored. To illustrate just one, the following develops a simple but illuminating example using the unambiguous state discrimination POVM. The resulting flexibility then leads, in the following section, into the challenge of optimizing measurement protocols to achieve various ends.

6.2.2. Unambiguous State Discrimination. Recall that when measuring qubits either in state $|\psi\rangle$ or $|\phi\rangle$, the POVM yielding the highest probability of unambiguously distinguishing between them is given by [83–85]:

$$E_\psi = \frac{1}{1 + |\langle\phi|\psi\rangle|} |\phi^\perp\rangle\langle\phi^\perp| \tag{6.2a}$$

$$E_\phi = \frac{1}{1 + |\langle\phi|\psi\rangle|} |\psi^\perp\rangle\langle\psi^\perp| \tag{6.2b}$$

$$E_? = \mathbb{I} - E_\psi - E_\phi \ . \tag{6.2c}$$

Applying this measurement scheme to the quantum process generated by the cCQS in Figure 4.2 produces the classical process emitted by the HMM of Figure 6.8. There $\{E_\psi, E_\phi, E_?\}$ are relabeled $\{E_0, E_1, E_2\}$, $p_\psi = \mathrm{Tr}(E_0 |\psi\rangle\langle\psi|)$, and $p_\phi = \mathrm{Tr}(E_1 |\phi\rangle\langle\phi|)$.

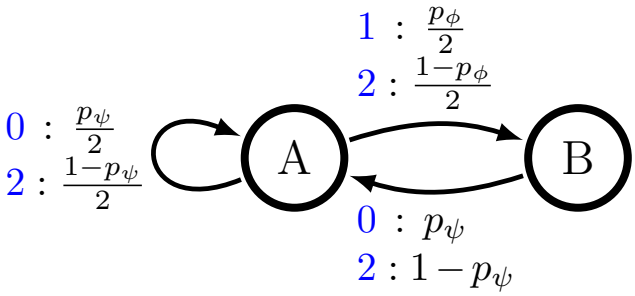


FIGURE 6.8. HMM presentation of the process resulting from measuring the QSSP depicted in Fig. 4.2 with the POVM in Eq. (6.2).

Note that symbol 2, corresponding to an inconclusive measurement, is present in all transitions. Yet observing 0 or 1 is synchronizing since they each determine the next HMM state. This property is preserved from the cCQS that is being measured and need not occur generally. That said, if the cCQS under study outputs only two distinct quantum states, then measuring it with the unambiguous state discrimination POVM in Equation (6.2) results in an HMM presentation that preserves the internal topology. However, each HMM transition is corrupted with a nonzero probability of observing symbol 2, rendering an inconclusive measurement.

For this example, constructing the MSP for the process generated the cCQS in Figure 6.8 produces the presentation depicted in Figure 6.9, where transition probabilities are not shown to reduce clutter.

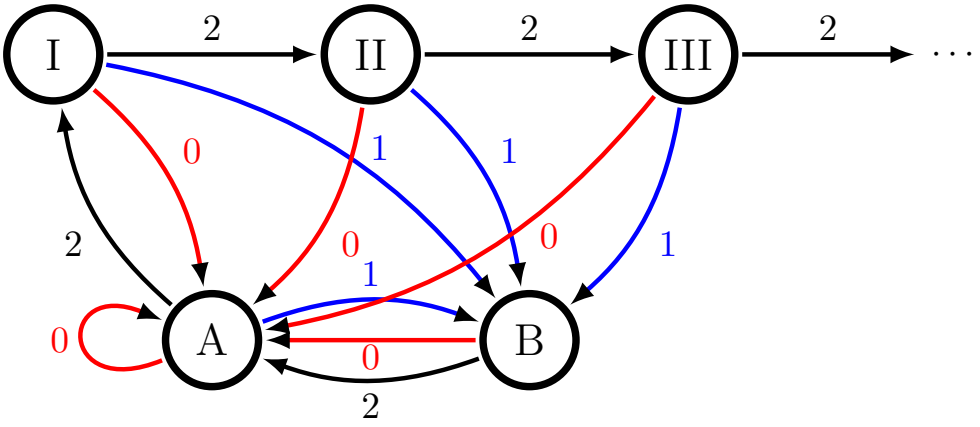


FIGURE 6.9. State transition diagram of the MSP constructed from the HMM in Fig. 6.8. Transition probabilities omitted for clarity. Observing 1s (blue transitions) leads to state B; observing 0s (red transitions) to state A. Both cases are synchronizing.

There is a subset of MSP states with topology similar to the original cCQS, but augmented with the mixed states that capture the observation of 2s. This generally holds when the generator is a unifilar cCQS that emits two distinct nonorthogonal quantum states if the process is measured via the POVM in Eqs. (6.2). In the measured cCQS, there will be a subset of MSP states that mimic the cCQS's internal dynamics, but the latter is augmented by inconclusive measurement outcomes. And, there are chains of mixed states that drive the process away from the original dynamic whenever a sequence of "inconclusive results" or a nonsynchronizing symbol is observed.

To explore this process further set:

$$|\phi\rangle = |0\rangle$$
$$|\psi\rangle = \cos(\alpha/2)\,|0\rangle + \sin(\alpha/2)\,|1\rangle \ ,$$

with $\alpha \in (0, \pi)$. Then, $p_\psi = p_\phi = 1 - \cos(\alpha/2)$. When constructing the MSP of Figure 6.9, all transitions that emit a 2 (black) have an associated probability of $\cos(\alpha/2)$, while the probabilities of the blue and red transitions depend on the specific transition. As the number of 2s observed approaches infinity, the mixed states visited approach the stationary state distribution $\pi = (2/3, 1/3)$ of the nonunifilar measured cCQS.

This MSP, while requiring a countable infinity of states, is so well behaved that it allows for direct calculation of the mixed states and the numerical computation of both its entropy rate and statistical complexity by approximating the MSP with a finite but sufficiently large set of mixed states. Figure 6.10 plots both the entropy rate and statistical complexity when the processes are approximated by a MSP with 500 hidden states.

Figure 6.10 reveals edge cases that match expectations. At one extreme, when $\alpha = 0$, the process reduces to a sequence of qubits in state $|0\rangle$. Thus, both process randomness and structure vanish. At the other extreme, when $\alpha = \pi$, the alphabet emitted by the cCQS is orthogonal $\{|0\rangle, |1\rangle\}$. Hence, the unambiguous state discrimination POVM reduces to the projective measurement aligned with the observation basis. This means that the measured process is true to the original quantum source and both its entropy rate and statistical complexity coincide with h_μ^q and C_μ^q, respectively. The plots also highlight that both entropy rate and statistical complexity coincide with the generator values with different nonorthogonal alphabets and that their maximum values are attained for different alphabets as well. While both randomness and structure of the measured process depend on the cCQS's quantum alphabet, they both have distinct meanings and, thus, the dependencies are not equivalent.

When varying the quantum alphabets and exploring α's whole range, it becomes apparent that the values of both entropy rate and statistical complexity can be lower or higher than those of the original cCQS. This means that for a given cCQS and a given measurement both the randomness and structure of the measured process with respect to the QSSP can be reduced or increased. This

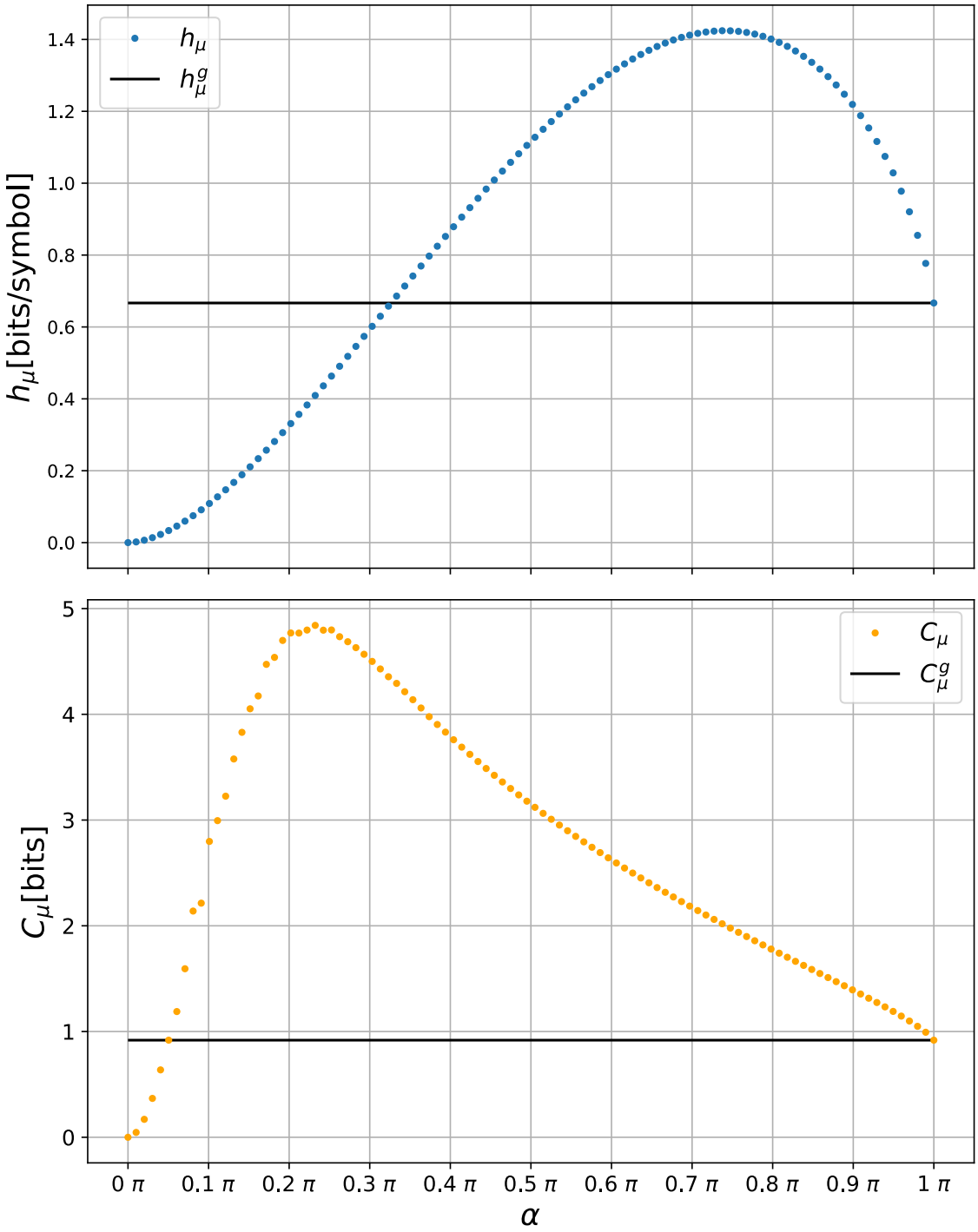


FIGURE 6.10. (Top) Randomness h_μ and (bottom) statistical complexity C_μ as a function of the angle α between the two states emitted by the cCQS in Fig. 4.2, with $|\phi\rangle = |0\rangle$ and $|\psi\rangle = \cos(\alpha/2)\,|0\rangle + \sin(\alpha/2)\,|1\rangle$. The horizontal black lines show the values of h_μ^g and C_μ^g of the cCQS that generates the original quantum state process.

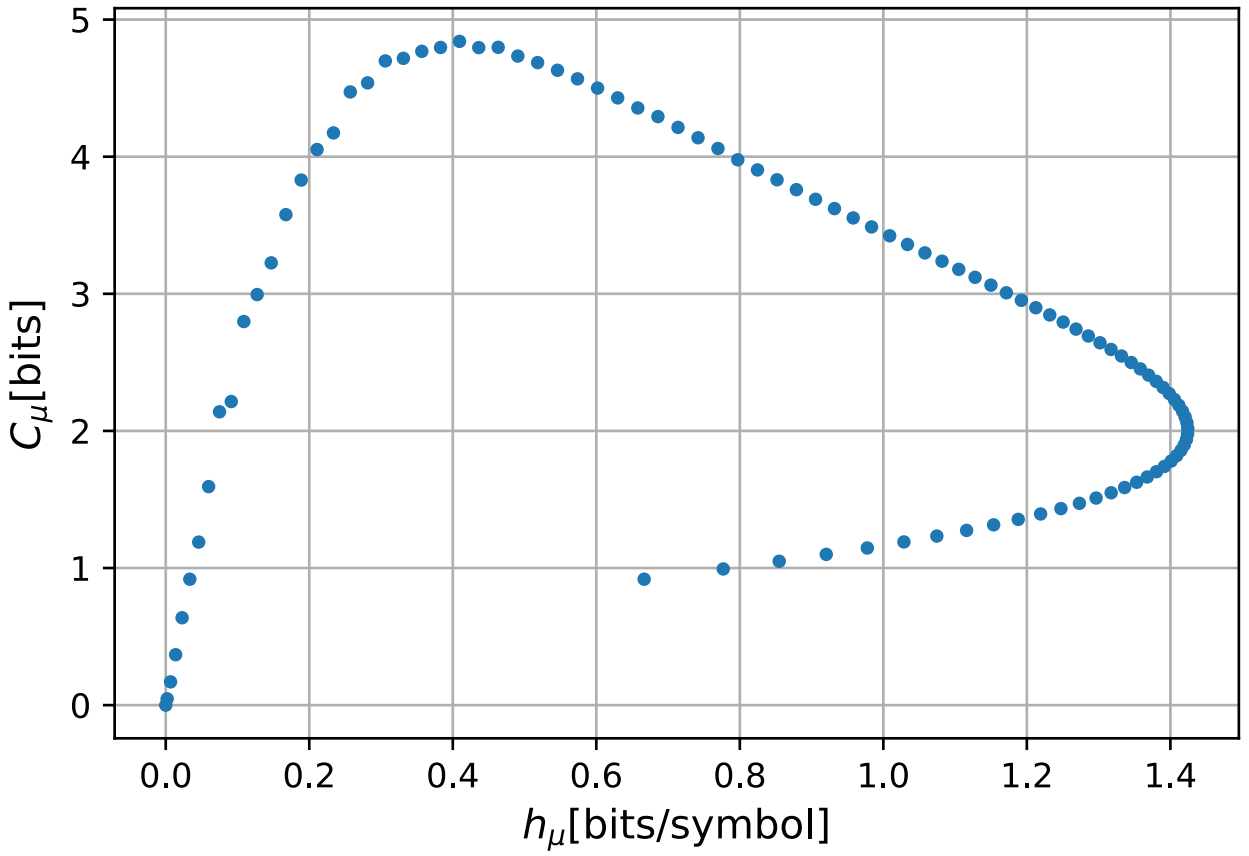


FIGURE 6.11. Complexity-entropy diagram capturing the purely informational character of the measured quantum-state process: Parametric plot of $h_\mu(\alpha)$ and $C_\mu(\alpha)$ over $\alpha \in [0, \pi]$ illustrating how the intrinsic informational properties depend on each other without reference to model parameter α. Cf. complexity-entropy plots in Ref. [86].

makes plain the possibly ambiguous effects of measurement and what the latter can add to or remove from the underlying quantum process.

Figure 6.11's *complexity-entropy diagram* [86] offers a more concise display of the QSSP's achievable information generation and storage—its *intrinsic computation*—when measured in this particular POVM across the range of alphabets. Notice that for small h_μ the system's structure or memory requirements C_μ are low. Then memory increases with increased randomness until a peak is reached at about $h_\mu = 0.4$ bits/symbol. Above this, increased randomness requires fewer memory resources and a given randomness can be achieved at more than one memory value C_μ.

Overall, this example illustrates a situation in which a particular choice of measurement protocol leads to a very tractable measured process. While the statistical complexity diverges for most measured processes, this example shows that tailoring measurement schemes still leads to complex, but more tractable dynamics. That is, the observed dynamics can be leveraged to better understand the dynamic that produces the underlying QSSP.

6.3. Optimal Measurements

The previous sections established the two main characters of measured quantum processes—their unpredictability and temporal correlation. And, they demonstrated how measurement can increase or decrease observed randomness and structure. These metrics naturally broach the challenge of defining and demonstrating the existence of informationally-optimal measurements. The possibility of these optimizations is greatly facilitated by the piecewise smooth dependence of the informational metrics on the QSSP and on measurement operators.

Eschewing details, the following lays out several avenues for future exploration, illustrating various kinds of optimality using the tools now in hand. We consider, in turn, measurements that lead to minimal structural complexity and to various forms of maximal informativeness. The following cases only address projective measurements, though extension to POVMs is in some cases straightforward and of interest in general.

Developing algorithms and calculational methods to find and implement these optimal measurements is left to the future.

Let's briefly recall relevant notation. A given projective measurement protocol is denoted $\mathcal{E}$. Given a QSSP $R_{-\infty:\infty}$ and a measurement protocol $\mathcal{E}$, the corresponding measured process is $X_{-\infty:\infty} = \mathcal{E}(R_{-\infty:\infty})$. To simplify the notation the following introduces $\mathcal{E}(M)$—the MSP of the measured cCQS of $X_{-\infty:\infty}$ for a given cCQS M.

The following treats the alternative informational metrics as operators themselves. So that for HMM M, $h_\mu(M)$ is that HMM's entropy rate, $C_\mu(M)$ is its statistical complexity, and $d_\mu(M)$ its statistical complexity dimension.

6.3.1. Minimal Structural Complexity. There are settings where it is useful to identify and use measurements that lead to the least complex, smallest-memory observed process. Such measurement schemes are specified as follows.

Definition 22 (Minimal Structural Complexity Measurement). *Given a cCQS M, the projective measurement $\mathcal{E}_{\underline{C_\mu}}$ that leads to the measured process with the minimal structural complexity is, when $\mathcal{E}(M)$ is finite state:*

$$\mathcal{E}_{\underline{C_\mu}} = \arg\min_{\{\mathcal{E}\}} C_\mu(\mathcal{E}(M))$$

and when $\mathcal{E}(M)$ uncountably infinite state:

$$\mathcal{E}_{\underline{C_\mu}} = \arg\min_{\{\mathcal{E}\}} d_\mu(\mathcal{E}(M)) \ .$$

While this kind of measurement is the least informative about the underlying quantum dynamics, it has also proven in multiple examples to be the measurement that yields a classical process requiring the least memory resources to simulate and predict. This remains to be proven but is consistent with the fact that the measurement effectively is the most efficient at discarding information about the structure of the underlying process, which need not be stored to represent the resulting measured process.

6.3.2. Maximally Informative Measurements. Perhaps most naturally, one can employ a measurement scheme that maximizes the amount of information per symbol in the measured process. Such measurement schemes are specified as follows.

Definition 23 (Maximally Informative Measurement). *Given a cCQS M, the projective measurement $\mathcal{E}_{\overline{h_\mu}}$ that leads to the measured process with maximally informative measurement outcomes is:*

$$\mathcal{E}_{\overline{h_\mu}} = \arg\max_{\{\mathcal{E}\}} h_\mu(\mathcal{E}(M)) \ .$$

Recalling basic dynamical systems, this is a natural choice of optimal measurement in that it mimics the essence of what a generating partition is, as defined by Kolmogorov and proven by Sinai [87]. In this case, each observation of a measurement outcome X_t results in the maximum possible amount of new information.

6.3.3. Maximally Mutually-Informative Measurements. One is often interested in monitoring how measurement outcomes reveal (or not) the internal generating mechanism. This suggests the following measurement.

Definition 24 (Maximally Mutually-Informative Measurement). *Given a QSSP $R_{-\infty:\infty}$, the measurement protocol $\mathcal{E}_{\overline{R:X}}$ is* the maximally mutually informative measurement *when:*

$$\mathcal{E}_{\overline{R:X}} = \arg\max_{\{\mathcal{E}\}} I(R_{-\infty:\infty} : X_{-\infty:\infty}) \ ,$$

where $I(R_{-\infty:\infty} : X_{-\infty:\infty})$ is the mutual information [88] between the QSSP and the measured process.

This measurement maximizes the information shared between the quantum-state stochastic process and the measured quantum process. That is, observation of the classical process maximally reduces uncertainty of the quantum process. In contrast to the *maximally informative measurement*, the *maximally mutually-informative measurement* does not yield the maximal amount of information learned per observation of the measured process. Rather, it provides the maximal amount of information learned about the QSSP given observation of the measured process.

6.3.4. Dynamically Informative Measurements. Finally, one may be interested in finding the measurement that yields a stochastic process most similar to the underlying QSSP. This requires adhering to a particular definition of distance or similarity between stochastic processes. One option when working with a particular cCQS and its corresponding measured cCQS is to use a measure of distance between HMMs. The measurement that minimizes the distance between two HMMs is the most informative about the internal structure of the QSSP generated by the cCQS. Said simply, it is the measurement yielding a classical stochastic process that is most informative about the QSSP's dynamical structure.

Selecting an appropriate measure of distance between HMMs is not a straightforward problem. Many have been proposed [89–91], each with their own nuances. Determining which distance measure better suits the problem at hand is left for future work.

This and the above notions of optimality are distinct and so are of interest in different operational settings. It is important to emphasize that what differentiates these optimality criteria from other notions of measurement optimality is that they depend on the QSSP's time correlations and not only on the particular quantum state of a single quantum system or its evolution.

CHAPTER 7

Conclusions

To investigate temporal complexity—unpredictability and structure—in quantum dynamics we developed an intentionally simplified setting—one that excluded sequential qubit entanglement. This allowed deploying classical multivariate information theory as a quantitative analysis tool. Importantly, this led directly to isolating the problem of how measurement affects the appearance of quantum processes—processes to which one must apply a quantum measurement to observe. The simple lesson is that measurement can both increase or decrease randomness and structure. In point of fact, and somewhat unanticipated, observing a quantum process through projective measurements results in an observed classical process of explosively high structural complexity. The detailed analysis enriched this by identifying the mechanisms through which this complexity arises.

In general, quantum-state stochastic processes observed through projective measurements result in observed classical processes that require storing an infinite number of predictive features to allow for optimal prediction. The sets of predictive features for most processes are rich in structure and, making use of that, we implemented newly developed tools to quantify their structural complexity and the intrinsic randomness of the measured process. In addition, the development shed light on the influence that the chosen measurement basis has on the complexity of the observed process.

Irreducible nonunifilarity was identified as the driving mechanism of these features. The low dimensionality of the quantum state Hilbert space is the physical cause. Nonalignment between the measurement basis and emitted quantum states of the cCQS is the root physical mechanism that leads to induced indistinguishability of quantum states.

Even allowing for the framework's simplifications, the typical complexity of the measured process complicates not only its study, but also makes the task of learning about the underlying quantum process a difficult one. We made progress in understanding why that is and how to characterize the measured processes. That progress came from adapting new methods from ergodic theory

and random dynamical systems to this setting. The result is a powerful toolset for quantitatively analyzing measured quantum processes.

We showed that the underlying cCQSs have distinct signatures of structure and randomness as a function of the measurement parameters and that dependence is systematic and smooth. One remaining task is to characterize the possible underlying quantum sources that generate a given measured cCQS or, at the very least, their statistical properties.

In this way, the results lay a path to fully characterizing quantum state stochastic processes. However, many steps remain unexplored. We conjecture that success in these will have broad impacts. One of those steps is to find the spectral decompositions of the processes by use of meromorphic functional calculus [**92**, **93**]. While these tools are not engaged here, they will be necessary when studying CQSs generated by purely quantum controllers. Another essential step is to model the internal controller in such a way that it generates entangled QSSPs. There are helpful starting points for this in both finite-memory classical controllers [**70**] and quantum protocols [**94**, **95**] for sequential generation of matrix product states. The latter are of particular interest in the study of many-body entangled states.

We only briefly explored a measurement protocol that used single qubit POVM measurements. This showed that exploring different measurement protocols has the potential to bring novel results and to move closer to more physically realistic settings. For example, Reference [**71**] looks at measurements that allow synchronization to the underlying QSSP in a setting similar to that explored here. As in the study of classical dynamical systems, though, understanding the informational and statistical effects of choosing a particular measurement instrument or protocol can aid in optimizing particular tasks. The study of optimal quantum measurements for QSSPs remains as a challenging open problem.

A major challenge is to extend the current setting to quantally-controlled qubit sources (qCQS), as just noted. And, then, from there to develop a quantum communication channel setting in which qubits are input, quantally processed, and then output. Advances in this will more directly impact information processing and computing performed by quantum dynamical systems. Beyond that, qubit source timing issues should be addressed, moving away from the admittedly simple use of discrete time here to continuous-time processes. Fortunately, the cCQS model can easily

be extended to discrete-event continuous-time hidden semi-Markov models using the methods of Reference [96]. This will immediately give metrics of quantum randomness and structure, paralleling the development here.

Extending the present results along these lines will naturally complement existing quantum descriptions of classical stochastic processes [97–99]. They also flag a starting point from which to understand the statistical and structural properties of quantum-state time series. That step will provide tools necessary not only for furthering our understanding of fundamental quantum dynamics, but also grasping the operational meaning of their informational properties in the context of quantum computation.

APPENDIX A

QSSP Python Package

Written in collaboration with David Gier.

A.1. Summary

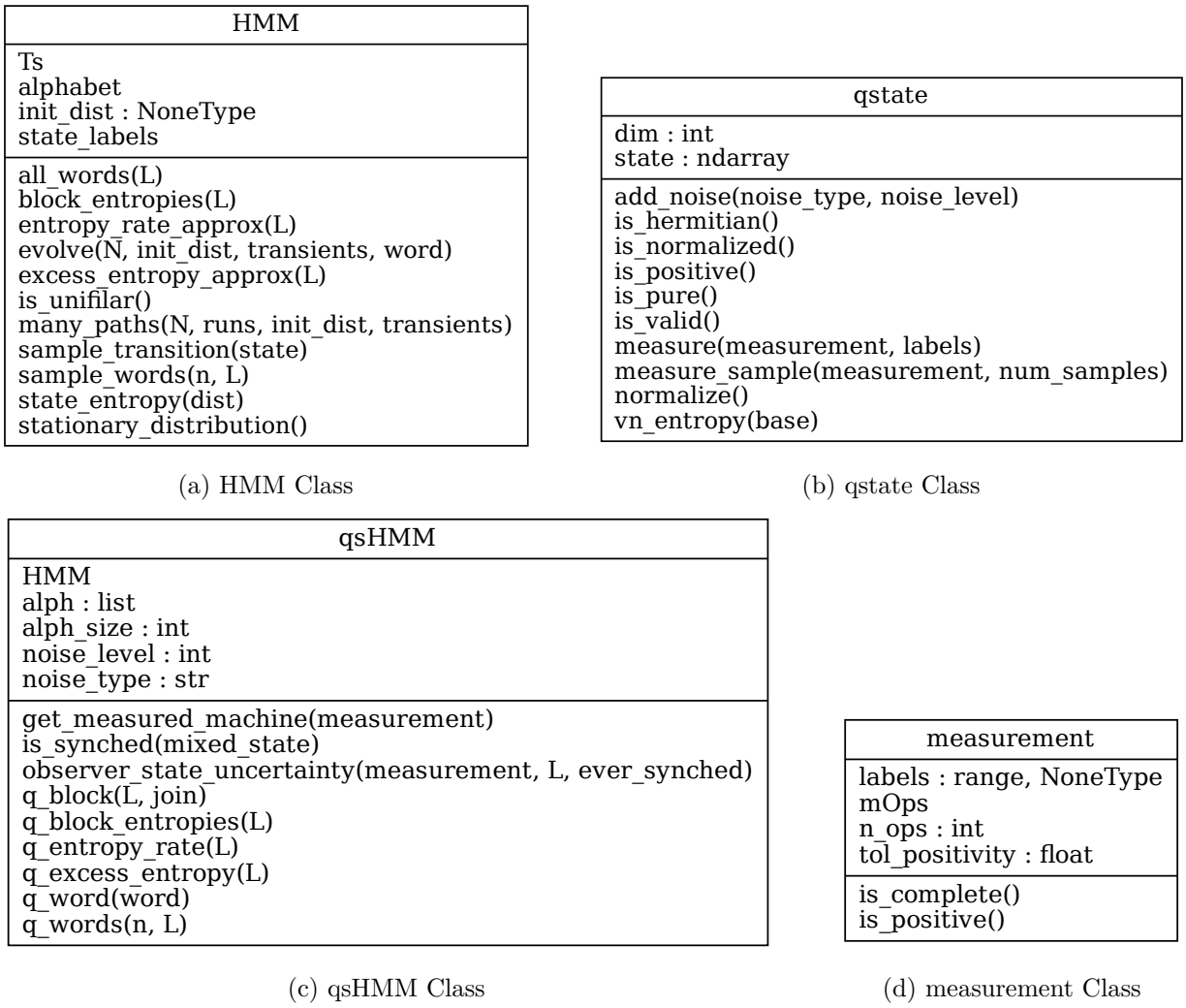
QSSP stands for Quantum-State Stochastic Processes. The Python package QSSP has been created with the purpose of generating, manipulating, and characterizing Quantum-State stochastic processes as in References [**15**, **16**, **71**]. It is open source and available at https://github.com/ari-VL/qssp.

QSSP uses basic Python libraries such as NUMPY, SCIPY and MATPLOTLIB. It is otherwise self contained. The details of the objects and what they do will be provided shortly, here we summarize. The basic functionality of the package allows the user to construct edge-emitting Hidden Markov Models (HMMs), generate realizations of stochastic processes with them, and quanitfy basic information properties of these processes from the HMMs. It is also capable of handling basic quantum states (qudit states) as desity matrices or kets, as well as Positive Operator Valued Measurements (POVMs) to measure these states. The main object of the package is the QSHMM, which models the classical controllers for Quantum-State Stochastic Processes described in the references above.

A.2. The Building Blocks

Figure A.1 describes the main objects of QSSP.

The `HMM` class is defined by inputing a numpy array of labeled transition matrices, and an optional 'initial distribution' which would describe an internal state distribution. The alphabet, or set of output symbols and the set of hidden state labels of the `HMM` are initialized by default by reading the number of labeled transition matrices and their size, respectively. The methods then allow the user to ask whether the `HMM` is unifilar or not. The user can also generate: the set of all allowed



(a) HMM Class

(b) qstate Class

(c) qsHMM Class

(d) measurement Class

FIGURE A.1. Class diagrams for the QSSP package

words (sequences of output symbols) of length L, a sample of a state transition from a given hidden state, n samples of length L of the stochastic process modeled by the `HMM`, and the stationary hidden state distribution of the `HMM`. For `HMM` characterization the user can compute: the state entropy in a given hidden state distribution (by default the stationary state distribution), the block entropies up to length L of the stochastic process, the entropy rate approximation at length L and the excess entropy approximation at length L.

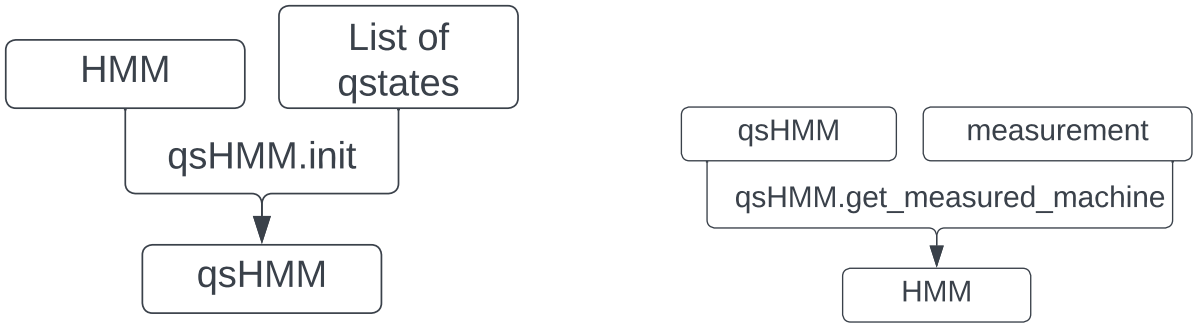
The `qstate` class represents quantum states described in the computational basis (qudits). It is constructed by inputting the amplitudes as a numpy array of amplitudes, either as a density matrix

or as a ket (row vector). The methods allow the user to check whether the state is: normalized—and if not, normalize it—, hermitian, positive, a valid quantum state and a pure quantum state. The user can also compute the von Neumann entropy of the state. Together with use of the `memasurement` class, the user can measure the staten and obtain a distribution over outcomes or sample the measurement and obtain a number of outcomes from the correct distribution.

The `measurement` class is initialized by providing an array of measuement operators, and an optional set labels for the outcomes, which are otherwise taken to be the first nonnegative integers. The initialization also checks the the measurement satisfies the requirements to be a valid POVM.

Last, the `qshmm` class makes use of the previous objects to implement a classical HMM controller as a source of a Quantum-State Stochastic Process(QSSP). The `qshmm` is initialized by inputing a controler `HMM` and a list of `qstates` which make up the alphabet of output quantum states of the QSSP. The user can then generate one or many realizations of length L of the QSSP, as well as compute the density matrix that describes all possible sequences of realizations of length L together with their probabilities. The user can also compute: the quantum block entropy of length L of the QSSP, the quantum entropy rate of the QSSP and the quantum excess entropy of the QSSP. When using an instance of a `measurement` object, the user can also obtain a `HMM` that models the classical measured stochastic process that results from measuring the QSSP generated by the `qshmm`. Said process can then be characterized by using the methods of the `HMM` object. Finally, the user can compute the average hidden state uncertainty of measuring up to L single qudit states with a fixed basis. Figure A.2 makes clear the interactions that `qshmm` has with other classes in the package.

The matrix necessary to represent a QSSP is d^L by d^L, where d is the dimension of an individual quantum state and L is the length of the sequence. Calculating information properties generically requires diagonalizing these matrices, thus these algorithms have execution time which scales exponentially with L. For instance, for the QSSPs in Reference [**71**], estimates these quantities with $L = 8$ (for $d = 3$), $L = 10$ and $L = 12$, all of which are possible with exact diagonalization on a single CPU. One direction for future development is speeding up these calculations using more sophisticated algorithms for matrix manipulation or converting the numerically-intensive calculations to a language with faster execution such as C or C++.



(a) Classical Process to Quantum Process

(b) Quantum Process to Classical Process

FIGURE A.2. Transforming between Classical and Quantum Processes

The package is tested and documented and has a built-in `utils` file with some of the most used 2- and 3-hidden state HMMs as well as some basic qubit states and measurements. It also has basic plotting capabilities to generate the Mixed State Presentation (MSP) of a 3-State nonunifilar HMM. A Jupyter notebook with simple examples of the central features of the package can be found at https://github.com/ari-VL/qssp/blob/main/qssp_example.ipynb.

APPENDIX B

Measurement Dependence Animations

Two animations illustrate the measurement angle dependence of the MSP; see:

https://csc.ucdavis.edu/~cmg/compmech/pubs/qdic.htm.

The first animation shows shows how the mixed state presentation and h_μ vary as a function of measurement parameter θ for the QSSP output by the cCQS in Figure 5.7.

The second animation shows plots like that in Fig. 6.3, while sweeping ϕ from 0 to 2π. In this case, for the process generated by the cCQS in Figure 6.1. In particular, Figure 6.3 coincides with the first frame of the animation.

Bibliography

[1] Cosma Shalizi and James Crutchfield. Pattern discovery in time series, Part I: Theory, algorithm, analysis, and convergence. *J. Mach. Lear. Res.*, 2002.

[2] Kristina Lisa Shalizi, Cosma Rohilla Shalizi, and James P Crutchfield. Pattern discovery in time series, Part II: Implementation, evaluation, and comparison. *J. Mach. Lear. Res.*, 2002.

[3] P. Schindler, J.T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt. Experimental repetitive quantum error correction. *Science*, 332(6033), 2011.

[4] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt. Quantum computations on a topologically encoded qubit. *Science*, 345(6194):302–305, 2014.

[5] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. Devoret, and R. Shoelkopf. Extending the lifetime of a quantum bit with error correction in superconducting qubits. *Nature*, 536:441–445, 2016.

[6] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout. Detecting crosstalk errors in quantum information processors. *Quantum*, 4:321, 2020.

[7] R. Harper, S. T. Flammia, and J. J. Wallman. Efficient learning of quantum noise. *Nature Phys.*, 16(1184-1188), 2020.

[8] A. Rivas, S. F. Huelga, and M. B. Plenio. Quantum non-Markovianity: characterization, quantification and detection. *Rep. Prog. Phys.*, 77(094001), 2014.

[9] I. de Vega and D. Alonso. Dynamics of non-Markovian open quantum systems. *Rev. Mod. Phys.*, 89:015001, Jan 2017.

[10] G. A. L. White, C.D. Hill, F. A. Pollock, L. C. L. Hollenberg, and K. Modi. Demonstration of non-Markovian process characterization and control on a quantum processor. *Nature Commun.*, 11(6301), 2020.

[11] R. Roy, S. Hohng, and T. Ha. A practical guide to single-molecule fret. *Nat. Methods*, 2008.

[12] C.-B. Li, H. Yang, and T. Komatsuzaki. Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proceedings of the National Academy of Sciences USA*, 105:536–541, 2008.

[13] M. Pirchi, R. Tsukanov, R. Khamis, T. Tomov, Y. Berger, D. Khara, H. Volkov, G. Haran, and E. Nir. Photon-by-photon hidden markov model analysis for microsecond single-molecule fret kinetics. *J. Phys. Chem. B*, 120:13065–13075, 2016.

[14] Tatsuhiro Furuta, Keisuke Hamada, Masaru Oda, and Kazuma Nakamura. Hidden Markov model analysis for fluorescent time series of quantum dots. *Phys. Rev. B*, 106:104305, Sep 2022.

[15] A. E. Venegas-Li, A. M. Jurgens, and J. P. Crutchfield. Measurement-induced randomness and structure in controlled qubit processes. *Phys. Rev. E*, 102:040102, 2020.

[16] A. Venegas-Li and J. P. Crutchfield. Optimality and complexity in measured quantum-state stochastic processes. *J. Stat. Phys.*, 190(106), 2023.

[17] A. Jurgens and J. P. Crutchfield. Shannon entropy rate of hidden Markov processes. *J. Stat. Phys.*, 183(32):1–18, 2020.

[18] A. Jurgens and J. P. Crutchfield. Divergent predictive states: The statistical complexity dimension of stationary, ergodic hidden Markov processes. *Chaos*, 31(8):0050460, 2021.

[19] A. Jurgens and J. P. Crutchfield. Ambiguity rate of hidden Markov processes. *Phys. Rev. E*, 104:064107, 2021.

[20] J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Lett.*, 63:105–108, 1989.

[21] C. R. Shalizi and J. P. Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *J. Stat. Phys.*, 104:817–879, 2001.

[22] J. P. Crutchfield. Between order and chaos. *Nature Physics*, 8(January):17–24, 2012.

[23] J. Parrondo, J. Horowitz, and T. Sagawa. Thermodynamics of information. *Nature Physics*, 11:131–139, 2015.

[24] T. Conte, E. DeBenedictis, N. Ganesh, T. Hylton, J. Paul Strachan, R. S. Williams, A. Alemi, L. Altenberg, G. E. Crooks, J. P. Crutchfield, L. del Rio, J. Deutsch, M. R. DeWeese, K. Douglas, M. Esposito, M. P. Frank, R. Fry, P. Harsha, M. D. Hill, C. Kello, J. Krichmar, S. Kumar, S.-C. Liu, S. Lloyd, M. Marsili, I. Nemenman, A. Nugent, N. Packard, D. Randall, P. Sadowski, N. Santhanam, R. Shaw, A. Z. Stieg, E. Stopnitzky, C. Teuscher, C. Watkins, D. Wolpert, J. J. Yang, and Y. Yufik. Thermodynamic computing. *CoRR*, abs/1911.01968, 2019.

[25] Melvin Lax. Quantum noise. xi. multitime correspondence between quantum and classical stochastic processes. *Phys. Rev.*, 172:350–361, Aug 1968.

[26] E. B. Davies. Quantum stochastic processes. *Commun. Math. Phys.*, 15(4):277–304, December 1969.

[27] E. B. Davies. Quantum stochastic processes II. *Commun. Math. Phys.*, 19(2):83–105, June 1970.

[28] E. B. Davies. Quantum stochastic processes III. *Commun. Math. Phys.*, 22(1):51–70, March 1971.

[29] Göran Lindblad. Non-Markovian quantum stochastic processes and their entropy. *Commun. Math. Phys.*, 65(3):281–294, October 1979.

[30] K. Burke, K. A. Mitchell, B. Wyker, S. Ye, and F. B. Dunning. Demonstration of turnstiles as a chaotic ionization mechanism in Rydberg atoms. *Phys. Rev. Lett.*, 107:113002, 2011.

[31] Francesco Ciccarello, Salvatore Lorenzo, Vittorio Giovannetti, and G. Massimo Palma. Quantum collision models: Open system dynamics from repeated interactions. *Physics Reports*, 954:1–70, 2022.

[32] C. Gardiner and P. Zoller. *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. Springer, 2004.

[33] A. A. Clerk, M. H. Devoret, S. M. Girvin, Florian Marquardt, and R. J. Schoelkopf. Introduction to quantum noise, measurement, and amplification. *Rev. Mod. Phys.*, 82:1155–1208, Apr 2010.

[34] H.-P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, United Kingdom, 2007.

[35] A. Rivas and S. F. Huelga. *Open Quantum Systems: An Introduction*. Springer, Heidelberg, Germany, 2012.

[36] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac. Assessing non-Markovian quantum dynamics. *Phys. Rev. Lett.*, 101:150402, Oct 2008.

[37] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini. Non-Markovian dynamics in open quantum systems. *Rev. Mod. Phys.*, 88:021002, Apr 2016.

[38] L. Li, M. J. W. Hall, and H. Wiseman. Concepts of quantum non-Markovianity: A hierarchy. *Phys. Rep.*, 759:1–51, 2018.

[39] C.-F. Li, G.-C. Guo, and J. Piilo. Non-Markovian quantum dynamics: What is it good for? *Europhysics Letters*, 128(3):30001, 2020.

[40] C.-F. Li, G.-C. Guo, and J. Piilo. Non-Markovian quantum dynamics: What does it mean? *Europhysics Letters*, 127(5):50001, 2019.

[41] A. A. Markov. Primer statisticheskogo issledovaniya nad tekstom "Evgeniya Onegina", illyustriruyuschij svyaz' ispytanij v cep'. *Izv. Akad. Nauk, SPb*, 93:153–162, 1913.

[42] A. A. Markov. *Ischislenie veroyatnostej*. Spb, 1900; 2-e izd., spb, 1908 edition, 1913. Translated into German, Wahrscheinlichkeits-Rechnung, Teubner, Leipzig-Berlin, 1912; 3-e izd., SPb, 1913; 4-e izd., Moskva, 1924.

[43] Simon Milz, Fattah Sakuldee, Felix A. Pollock, and Kavan Modi. Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories. *Quantum*, 4:255, April 2020.

[44] Simon Milz and Kavan Modi. Quantum stochastic processes and quantum non-Markovian phenomena. *PRX Quantum*, 2:030201, Jul 2021.

[45] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi. Non-Markovian quantum processes: Complete framework and efficient characterization. *Phys. Rev. A*, 97:012127, Jan 2018.

[46] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi. Operational Markov condition for quantum processes. *Phys. Rev. Lett.*, 120:040405, Jan 2018.

[47] P. Taranto, S. Milz, F. A. Pollock, and K. Modi. Structure of quantum stochastic processes with finite Markov order. *Phys. Rev. A*, 99:042108, Apr 2019.

[48] P. Taranto, F. A. Pollock, S. Milz, M. Tomamichel, and K. Modi. Quantum Markov order. *Phys. Rev. Lett.*, 122:140401, Apr 2019.

[49] A. Efros and D. Nesbitt. Origin and control of blinking in quantum dots. *Nature Nanotech*, 11:661–671, 2016.

[50] Roberto N. Muñoz, Laszlo Frazer, Gangcheng Yuan, Paul Mulvaney, Felix A. Pollock, and Kavan Modi. Memory in quantum dot blinking. *Phys. Rev. E*, 106:014127, Jul 2022.

[51] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov. Single-photon sources and detectors. *Review of Scientific Instruments*, 82(071101), 2011.

[52] N. Somaschi, V. Giesz, L. De Santis, J. C. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Antón, J. Demory, C. Gómez, I. Sagnes, N. D. Lanzillotti-Kimura, A. Lemaítre, A. Auffeves, A. G. White, L. Lanco, and P. Senellart. Near-optimal single-photon sources in the solid state. *Nature Photonics*, 10(5):340–345, May 2016.

[53] Pascale Senellart, Glenn Solomon, and Andrew White. High-performance semiconductor quantum-dot single-photon sources. *Nature Nanotechnology*, 12(11):1026–1039, November 2017.

[54] Evan Meyer-Scott, Christine Silberhorn, and Alan Migdall. Single-photon sources: Approaching the ideal through multiplexing. *Review of Scientific Instruments*, 91(4):041101, 2020.

[55] Sarah Thomas and Pascale Senellart. The race for the ideal single-photon source is on. *Nature Nanotechnology*, 16(4):367–368, April 2021.

[56] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: Levels of entropy convergence. *CHAOS*, 13(1):25–54, 2003.

[57] P. M. Ara, R. G. James, and J. P. Crutchfield. The elusive present: Hidden past and future dependence and why we build models. *Phys. Rev. E*, 93(2):022143, 2016.

[58] G. P. Basharin, A. N. Langville, and V. A. Naumov. The life and work of A. A. Markov. *Linear Algebra and its Applications*, 386:3–26, 2004.

[59] A. A. Markov. An example of statistical investigation of the text "Eugene Onegin" concerning the connection of samples in chains. *Science in Context*, 19:591–600, 2006.

[60] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Many roads to synchrony: Natural time scales and their algorithms. *Phys. Rev. E*, 89:042135, 2014.

[61] L. R. Rabiner and B. H. Juang. An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1):4–16, 1986.

[62] L. R. Rabiner. A tutorial on hidden Markov models and selected applications. *IEEE Proc.*, 77:257, 1989.

[63] J. P. Crutchfield. The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75:11–54, 1994.

[64] S. Marzen and J. P. Crutchfield. Informational and causal architecture of discrete-time renewal processes. *Entropy*, 17(7):4891–4917, 2015. SFI Working Paper 14-08-032; arxiv.org:arXiv:1408.6876 [cond-mat.stat-mech].

[65] S. Marzen and J. P. Crutchfield. Statistical signatures of structural organization: The case of long memory in renewal processes. *Phys. Lett. A*, 380(17):1517–1525, 2016.

[66] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, second edition, 2006.

[67] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948.

[68] D. Blackwell. The entropy of functions of finite-state Markov chains. volume 28, pages 13–20, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957. Held at Liblice near Prague from November 28 to 30, 1956.

[69] S. E. Marzen and J. P. Crutchfield. Nearly maximally predictive features and their dimensions. *Phys. Rev. E*, 95(5):051301(R), 2017.

[70] G. M. Crosswhite and D. Bacon. Finite automata for caching in matrix product algorithms. *Phys. Rev. A*, 78(012356), 2008.

[71] D. Gier and J. P. Crutchfield. Intrinsic and measured information in separable quantum processes. *arXiv:2303.00162*, 2023.

[72] J. Bechhoefer. Hidden Markov models for stochastic thermodynamics. *New. J. Phys.*, 17:075003, 2015.

[73] P. Riechers and J. P. Crutchfield. Spectral simplicity of apparent complexity, Part I: The nondiagonalizable metadynamics of prediction. *Chaos*, 28:033115, 2018.

[74] P. Riechers and J. P. Crutchfield. Spectral simplicity of apparent complexity, Part II: Exact complexities and complexity spectra. *Chaos*, 28:033116, 2018.

[75] C. Moore and J. P. Crutchfield. Quantum automata and quantum grammars. *Theoret. Comp. Sci.*, 237:1-2:275–306, 2000.

[76] J. von Neumann and R.T. Beyer. *Mathematical Foundations of Quantum Mechanics*. Goldstine Printed Materials. Princeton University Press, 1955.

[77] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, United Kingdom, tenth anniversary edition, 2011.

[78] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney. Time's barbed arrow: Irreversibility, crypticity, and stored information. *Phys. Rev. Lett.*, 103(9):094101, 2009.

[79] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *J. Stat. Phys.*, 136(6):1005–1034, 2009.

[80] P. M. Centore and E. R. Vrscay. Continuity of attractors and invariant measures for iterated function systems. *Canadian Math. Bull.*, 37(3):315–329, 1994.

[81] F. Mendivil. A generalization of ifs with probabilities to infinitely many maps. *Rocky Mountain J. Math.*, 28(3):1043–1051, 1998.

[82] B. Kloeckner. Optimal transportation and stationary measures for Iterated Function Systems. *Mathematical Proceedings*, 2021. https://hal.archives-ouvertes.fr/hal-02276750.

[83] I. D. Ivanovic. How to differentiate between non-orthogonal states. *Phys. Let. A*, 123(6):257–259, 1987.

[84] D. Dieks. Overlap and distinguishability of quantum states. *Phys. Let. A*, 126(5):303–306, 1988.

[85] A. Peres. How to differentiate between non-orthogonal states. *Phys. Let. A*, 128(1):19, 1988.

[86] D. P. Feldman, C. S. McTague, and J. P. Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *CHAOS*, 18(4):043106, 2008.

[87] Ja. G. Sinai. On the notion of entropy of a dynamical system. *Dokl. Akad. Nauk. SSSR*, 124:768, 1959.

[88] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 1991.

[89] M. Falkhausen, H. Reininger, and D. Wolf. Calculation of distance measures between hidden Markov models. In *EUROSPEECH*, 1995.

[90] S. M. E. Sahraeian and B.-J. Yoon. A novel low-complexity HMM similarity measure. *IEEE Signal Proc. Let.*, 18(2):87–90, 2011.

[91] J. Zeng, J. Duana, and C. Wu. A new distance measure for hidden Markov models. *Expert Systems with Applications*, 37(2):1550–1555, 2010.

[92] James P. Crutchfield, Christopher J. Ellison, and Paul M. Riechers. Exact complexity: The spectral decomposition of intrinsic computation. *Physics Letters A*, 380(9):998–1002, 2016.

[93] P. M. Riechers and J.P. Crutchfield. Beyond the spectral theorem: Decomposing arbitrary functions of non-diagonalizable operators. *A.I.P Advances*, 8:065305, 2018.

[94] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf. Sequential Generation of Entangled Multiqubit States. *Phys. Rev. Lett.*, 95(11):110503, 2005.

[95] C. Schön, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano. Sequential generation of matrix-product states in cavity QED. *Phys. Rev. A*, 75(3):032311, 2007.

[96] S. Marzen and J. P. Crutchfield. Structure and randomness of continuous-time discrete-event processes. *J. Stat. Physics*, 169(2):303–315, 2017.

[97] M. Gu, K. Wiesner, E. Rieper, and V. Vedral. Quantum mechanics can reduce the complexity of classical models. *Nature Comm.*, 3(762):1–5, 2012.

[98] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam's quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel. *Scientific Reports*, 6:20495, 2016.

[99] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Minimized state-complexity of quantum-encoded cryptic processes. *Phys. Rev. A*, 93(5):052317, 2016.