Spring 2019

Dynamical Entropy of Quantum Random Walks

Duncan Wright

Dynamical Entropy of Quantum Random Walks

by

Duncan Wright

Bachelor of Arts
University of Northern Iowa 2012

Master of Arts
University of Northern Iowa 2014

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Mathematics

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

George Androulakis, Major Professor

Stephen Dilworth, Committee Member

Éva Czabarka, Committee Member

Anton Schep, Committee Member

Yuriy Pershin, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

Dedication

To my beautiful wife, April. None of this would have been possible without your loving support, encouragement and patience. I look forward to continuing our journey together.

Acknowledgments

First, I would like to thank George Androulakis. Thanks for accepting me as a student. You allowed me to follow my own intuitions while developing this thesis, providing just the right amount of criticism when I was getting too far off-track. Most importantly, thank you for reading all of my work thoroughly and preventing me from making unsubstantiated claims of grandeur.

Thank you to all of the professors who have taught me during my tenure here. To Éva Czabarka, László Székely and Josh Cooper, thank you for teaching me the ways of graph theory and discrete mathematics; I hope that we can collaborate on future work. Stephen Dilworth, thank you for encouraging my questions on martingales and working with me to produce results on those questions. Thank you to all the members of the Analysis Seminar: Stephen Dilworth, George Androulakis, Anton Schep, Maria Girardi, Ralph Howard, Daniel Dix and Changhui Tan. You listened to me babble about my work consistently, even when I couldn't figure out how to give a fifty minute talk. Your comments and patience were extremely helpful. A special thanks to Éva Czabarka. Without your wisdom and encouragement, I would have floundered when it came to the job search.

Thank you to all of the staff who have picked up the slack left behind by me and the others in LeConte. To DeeAnn Moss, thank you for your casual conversations and always being available to help whenever I got overwhelmed. To Julia Clark-Spohn and Lauren Jahrsdoerfer, thank you for answering all of my silly questions and providing me support even when it was not your responsibility. You three have made my experience at U of SC much easier and more enjoyable.

Thank you to my parents, Cecil and Penny, and my brothers whose names are too many to list here. You have always pushed me to be my best and I am eternally grateful for your support. Thank you for excusing all the missed family functions and holidays while I have pursued my Ph.D.

Lastly, I would like to thank my friends and fellow graduate students: Hays Whitlatch, Harsh Mehta, Robert Vandermolen, Chris Edgar and so many others. Thanks for keeping me sane throughout this entire process and never letting me forget my mistakes (e.g. that trash can thing). Hays and Rob, thanks for reading through all of my proposals and statements, and providing me with much needed feedback. I am especially grateful to (Dr.) Hays Whitlatch. You kept me on schedule with everything this last year. Without your help I would not be looking forward to a wonderful Postdoc position and I would probably not even be graduating this Spring.

This work was supported in part by a SPARC Graduate Fellowship from the Office of the Vice President for Research at the University of South Carolina.

Abstract

In this manuscript, we study discrete-time dynamics of systems that arise in physics and information theory, and the measure of disorder in these systems known as dynamical entropy. The study of dynamics in classical systems is done from two distinct viewpoints: random walks and dynamical systems. Random walks are probabilistic in nature and are described by stochastic processes. On the other hand, dynamical systems are described algebraically and deterministic in nature. The measure of disorder from either viewpoint is known as dynamical entropy.

Entropy is an essential notion in physics and information theory. Motivated by the study of disorder for the positions and velocities of gas molecules, the notion of entropy was first introduced mathematically by Boltzmann near the end of the 19th Century and gives rise to the second law of thermodynamics. Almost eighty years after Boltzmann, Shannon became the father of the new field of information theory when he produced his groundbreaking works where he used entropy as a measure of information transfer between two sources. In the last two years of the 1950's, Kolmogorov and Sinai extended the notions of Boltzmann to a dynamical entropy. The Kolmogorov-Sinai dynamical entropy gives a measure for the disorder of a system of particles (e.g. gas molecules) averaged over time, quantifying the uncertainty in the dynamics of a system.

The advent of quantum mechanics and its pervasiveness in nature has required the development of non-commutative generalizations of dynamics and dynamical entropy to the quantum regime. Many of each have been proposed. In particular, we recall the definitions of quantum random walks, dynamical systems and Markov chains.

We motivate each generalization by relating to its classical counterpart. Quantum dynamical entropy (QDE) is a generalization of the Kolmogorov-Sinai dynamical entropy to quantum mechanics. There have been numerous definitions for QDE beginning with that of Connes, Narnhofer and Thirring in 1987. We focus on the semi-classical approach given by Słomczyński and Życzkowski in 1994 and the quantum Markov chain approach which started with Accardi, Ohya and Watanabe in 1997.

Linearity of a dynamical entropy means that the dynamical entropy of the n-fold composition of a dynamical map with itself is equal to n times the dynamical entropy of the map for every positive integer n. We show that the quantum dynamical entropy introduced by Słomczyński and Życzkowski is nonlinear in the time interval between successive measurements of a quantum dynamical system. This is in contrast to Kolmogorov-Sinai dynamical entropy for classical dynamical systems, which is linear in time. We also compute the exact values of quantum dynamical entropy for the Hadamard walk with varying Lüders-von Neumann instruments and partitions.

In 1948, Shannon proved the Source Coding Theorem which gives upper and lower bounds on the minimal expected codeword length in terms of the entropy of a random variable. This theorem can be leveraged to give the minimal expected average codeword length for a string of symbols in terms of entropy rate, which can be interpreted as the dynamical entropy of a stochastic process. In 1994, Schumacher defined indeterminate-length quantum codes and proved a Quantum Source Coding Theorem. We introduce the notion of stochastic ensembles of pure states and give a novel representation in terms of quantum Markov chains. Moreover, this representation allows us to extend the Quantum Source Coding Theorem, giving the minimal expected average codeword length of an indeterminate-length quantum code in terms of quantum dynamical entropy.

This manuscript includes joint work with Dr. George Androulakis (Mathematics, University of South Carolina).

Table of Contents

Chapter 1

Introduction

In this manuscript, we study discrete-time dynamics of systems that arise in physics and information theory, and the measure of disorder in these systems known as dynamical entropy.

1.1 Dynamics and Entropy in Classical Systems

The study of dynamics in classical systems is done from two distinct viewpoints: random walks and dynamical systems. The origins of the classical random walk traces back to the study of gambling problems, such as the gambler's ruin problem which dates back to 1656. Random walks are probabilistic in nature and are described by stochastic processes. Over the years random walks have seen much success in many fields ranging from finance ([60]), computer optics ([27]) and computer science ([49]) to biology ([26]) and neurology ([25]). On the other hand, a dynamical system is a probability space together with a measure-preserving automorphism, thereby giving deterministic dynamics. Dynamical systems are often used in physics to model an ensemble of particles whose state varies over time. Motivated by his study of the three-body problem in celestial mechanics, many people regard Henri Poincaré as the founder of dynamical systems (see e.g. [47]).

In both random walks and dynamical systems, entropy is used as a measure of uncertainty or chaos. In physics and dynamical systems theory, entropy is used to describe chaos and gives rise to the second law of thermodynamics. In information theory, entropy is used to quantify many properties of information (e.g. the resources needed to store information or data compression; see [33, 54]). The notion of entropy was first introduced mathematically by Boltzmann near the end of the 19th century in [15] as a tool to measure disorder for the positions and velocities of gas molecules, although the name "entropy" dates back to Clausius in 1865 according to [65]. Almost eighty years after Boltzmann, Shannon became the father of the new field of information theory when he produced his groundbreaking works where he used entropy as a measure of information transfer between two sources in [54, 55].

The entropy described by both Boltzmann and Shannon is static. That is, it is used to quantify the amount of disorder in a system at a fixed point in time. To quantify the amount of disorder in the dynamics of a system, the limit of the time-averaged entropy of joint probabilities is used. Due to the Césaro Mean Theorem, this limit can also be thought of as disorder of a system at the present time conditioned on the past. For stochastic processes, this dynamical entropy is known as entropy rate. In dynamical systems, this dynamical entropy was introduced by Kolmogorov and Sinai in 1958 and is now referred to as Kolmogorov-Sinai (KS) entropy ([32, 57]). Motivated by the problem of characterizing isometric dynamical systems, the KS entropy is a metric invariant for dynamical systems. The connection between entropy rate and KS entropy can be seen through the symbolic dynamics space which is discussed in Sections 2.3 and 5.3.

Another important property of KS entropy is its linearity in time. Linearity of the KS entropy means that, for a dynamical system (Ω, Σ, μ, f) and a positive integer n, the KS entropy of the n-fold composition of f is equal to n times the KS entropy of f and is given by the equation

$$h^{KS}(f^n) = nh^{KS}(f), \quad \text{for all } n \in \mathbb{N}.$$

In contrast, the entropy rate of a stochastic process is not linear in time due to its probabilistic nature. This difference, amongst others, is discussed in Section 5.4.

1.2 Dynamics in Quantum Systems

The advent of quantum mechanics and its pervasiveness in nature has required the development of non-commutative generalizations of dynamics to the quantum regime. Many descriptions of discrete-time, non-commutative dynamics have been proposed and studied over the years. In particular, we recall the definitions of quantum random walks, dynamical systems and Markov chains in Chapter 4.

Quantum random walks (QRWs) are a non-commutative generalization of dynamics from the random walk point of view, the properties of which are significantly different from their classical courterparts (see e.g. [8]). QRWs were first introduced by Aharonov et al. in [5] and independently by Meyer in [43] to describe closed system dynamics. Closed systems in quantum mechanics are described by unitary operators or, more precisely, unitary transformations; QRWs for closed systems are no different and are referred to as unitary QRWs (UQRWs). In 2012, a new variant of the QRW was introduced by Attal et al. in [12] to describe open system dynamics; i.e. a system coupled with an environment. QRWs for open systems are referred to as open QRWS (OQRWs). The evolution of an OQRW is described by a completely positive trace preserving map or quantum channel. The study of QRWs has enjoyed much interest in recent years (see e.g. [7, 29, 30, 34, 48, 63]), and applications have been found in many areas including quantum computing [35, 37], the study brain networks [41], and biology [50, 51]. In their seminal paper on OQRWs ([12]), the authors show that unitary and open QRWs differ only by a single step in their realization procedure. Namely, the step that requires decoherence and hence simulation of interaction with an environment. In Subsection 4.1.3, we develop this idea further.

Quantum dynamical systems (QDSs) are a non-commutative generalization of dynamical systems motivated by the algebraic representation of automorphisms on a probability space (Ω, Σ, μ) as $L^\infty(\Omega)$. $L^\infty(\Omega)$ can then be embedded into a commutative subalgebra of a matrix algebra M, the whole of which is non-commutative (see e.g. [11]). Matrix algebras are C^*-algebra and the general description of QDSs is done using C^*-algebra quantum mechanics.

Quantum Markov chains (QMCs) were first introduced by Accardi in [1]. A stochastic process is generally defined by its joint probability mass functions (pmfs). A Markov process is a special case of a stochastic process in which the joint pmf for one step into the future is only influenced by its current distribution. Therefore a Markov process can be uniquely defined by its initial distribution, μ, and a family of conditional expectation operators, $(P_n)_{n\in\mathbb{N}}$. From this viewpoint, we can denote the Markov process as the pair $\{\mu, P_n\}$. The QMC approach to non-commutative dynamics is based on this viewpoint. In QMCs, the initial distribution is represented by a density operator, ρ, in a matrix algebra and the conditional expectations are generalized to a family of transition expectations, $(\mathcal{E}_n)_{n\in\mathbb{N}}$, giving a QMC $\{\rho, \mathcal{E}_n\}$. The QMC approach is then a straightforward generalization of dynamics from the random walk perspective. On the other hand, by careful selection of transition expectations, the QMC approach can also be used to describe the course-grained measurements of a QDS. This viewpoint allows us to view QMCs as a generalization of dynamical systems as well. We will discuss this point in Sections 4.3 and 7.

1.3 Quantum Dynamical Entropy

The main focus of this work is the description of dynamical entropy in quantum systems. There have been many successful attempts to generalize KS entropy to a quantum dynamical entropy (QDE) in [18, 6, 59, 44, 3, 38], beginning with Connes, Narnhofer and Thirring in 1987. The Connes-Narnhofer-Thirring (CNT) [18], Alicki-Fannes (AF) [6], Accardi-Ohya-Watanabe (AOW) [3] and Kossakowski-Ohya-Watanabe (KOW) [38] entropies have had the most attention in the literature as they have been computed exactly for several examples of quantum dynamical systems. Moreover, the relationship between the different definitions is not fully understood, although some work in this direction has been done (e.g. [4, 61, 46, 14]).

We are particular interested in the Słomczyński-Życzkowski (SZ) quantum dynamical entropy [59] and the QMC approach to dynamical entropy introduced in [3] in terms of the AOW entropy. SZ entropy uses a semi-classical approach and was developed using the general notions of measurements, instruments, phase space and state space developed by Edwards in [23], and Davies and Lewis in [21, 20]. In contrast to both the CNT and AF entropies, SZ dynamical entropy can obtain nonzero values for quantum systems with finite-dimensional Hilbert spaces. Quantum algorithms are a natural example where this property is desirable.

In the past four decades there has been a lot of interest in trying to develop a quantum computer which requires a new field of computing, known as quantum computing (see [45]). In 1996, Lov Grover created a quantum database searching algorithm, now referred to as Grover's algorithm, which was shown to be quadratically faster than the classical analogue ([28]). Certain quantum algorithms have even obtained exponential speed-up over their classical counterparts, such as Shor's algorithm for factorizing integers ([56]). To demonstrate the applicability of SZ dynamical entropy for quantum algorithms, we will apply it to UQRWs which have been shown to be universal for quantum computation in [40] and give exact computations of the SZ dynamical entropy for the Hadamard walk measured with varying Lüders-von Neumann instruments. In doing so, we also give explicit calculations verifying the nonlinearity of SZ entropy. This is in contrast to KS entropy, which is linear in time, and gives further evidence to the probabilistic nature of measurements in a quantum system.

The QMC approach to dynamical entropy has been studied by many. It has been modified from its original context in terms of AOW entropy to include an approach for the study of the AF entropy developed by Tuyls in [61]. AF entropy was originally introduced in the context of QDSs and the approach of Tuyls lays the groundwork for the QMC representation for QDSs. A generalization of both QMC approaches was given in [38], where the authors introduced the KOW entropy. Furthermore, in Section 7.2, we give a QMC approach for SZ dynamical entropy. This representation allows us to show that the AOW entropy is also nonlinear in time.

In Section 7.3, we recall the notion of indeterminate-length quantum codes which were first considered by Schumacher in 1994 ([52]). In subsequent work with Westmoreland, Schumacher is able to extend the Source Coding Theorem to a quantum analog, establishing upper and lower bounds on the minimal expected quantum data compression in terms of von Neumann entropy. In classical data compression, the Source Coding Theorem can be leveraged to give the minimal average expected data compression rate in terms of the entropy rate of the corresponding stochastic process. We introduce the notion of a stochastic ensemble of pure states for indeterminate-length quantum codes and provide a QMC whose dynamical entropy is equal to the minimal average expected quantum data compression rate in Theorem 7.3.9.

Throughout this manuscript we will be interested in discrete-time processes. Whenever suitable definitions for continuous time have been made, we will try to mention it and give references to related material.

Chapter 2

Dynamics in a Classical Setting

In this chapter we will recall two different formalisms of dynamics in a classical (or commutative setting). We recall random walks in probability theory in Section 2.1 and try to highlight the graph-theoretic intuitions of these random processes. In Section 2.2, we recall dynamical systems which are dynamics from an algebraic point of view. We then introduce symbolic dynamics of a random walk in Section 2.3 which is a dynamical system connecting the two different pictures of dynamics. We finish off the chapter by delving more deeply into the algebraic structure of dynamical systems in Section 2.4, giving us a first peek at quantum dynamical systems. All of the systems of this chapter will have entropy introduced on them in Chapter 5 and will be extended to non-commutative (or quantum) settings in Chapter 4.

2.1 Random Walks in Probability Theory

Let (Ω, Σ, μ) be a probability space and let $(E, \mathcal{E})$ be a measurable space. Whenever a set Ω which is finite or countably infinite is equipped with the power set σ-algebra $\mathcal{P}(\Omega)$, we will refer to the measurable space $(\Omega, \mathcal{P}(\Omega))$ (or simply the set Ω) as a **discrete space**. An $\boldsymbol{(\Omega, E)}$ **random variable** X is a measurable map $X : \Omega \to E$. For any $S \in \mathcal{E}$, the probability that X takes values in S is given by $\mu(X \in S) := \mu(X^{-1}(S))$. We say that the random variable $\boldsymbol{X}$ **is discrete** whenever its range E is a discrete space. In that case, X is determined by its **probability mass function** (pmf) $p_X : E \to [0, 1]$ given by $p_X(x) = \mu(X = x)$ for each $x \in E$. We will simply write $p(x)$ as opposed to $p_X(x)$ when there is no confusion about the random variable.

We will be mainly interested in discrete random variables in this manuscript.

Given a finite collection, $(X_k)_{k=1}^n$, of (Ω, E) discrete random variables the **joint pmf of $(X_1, \dots, X_n)$** is given by

$$p_{X_1,\dots,X_n}(x_1, \dots, x_n) = \mu(X_1 = x_1, \dots, X_n = x_n) \quad \text{for all } x_1, \dots, x_n \in E. \tag{2.1}$$

Furthermore, $(X_1, \dots, X_n)$ is a discrete (Ω, E^n) random variable.

The **conditional pmf of X_n given $(X_1, \dots, X_{n-1})$** is given by

$$\begin{aligned} p_{X_n|(X_1,\dots,X_{n-1})}(x_n|x_1, \dots, x_{n-1}) &:= \mu(X_n = x_n|X_1 = x_1, \dots, X_{n-1} = x_{n-1}) \\ &= \frac{p_{X_1,\dots,X_n}(x_1, \dots, x_n)}{p_{X_1,\dots,X_{n-1}}(x_1, \dots, x_{n-1})}, \end{aligned} \tag{2.2}$$

for all $x_1, \dots, x_n \in E$. Whenever there is no confusion about the random variables in question we simply write $p(x_1, \dots, x_n)$ for the joint pmf and $p(x_n|x_1, \dots, x_{n-1})$ for the conditional pmf.

If (Ω, Σ, μ) is a probability space and $(E, \mathcal{E})$ is a measurable space, then a **(discrete time) (Ω, E) stochastic process** is an indexed sequence of (Ω, E) random variables. Throughout this manuscript we will only consider discrete time stochastic processes and the sequences will all be indexed by $\mathbb{N}$, where the index is meant to represent time. Whenever $(E, \mathcal{E})$ is a discrete space we will refer to any stochastic process with range in E as a **discrete stochastic process**. In this case the stochastic process, $\mathbf{X} = (X_n)_{n=0}^\infty$, is determined by its joint pmf, denoted by $p_\mathbf{X}$, and given by $p_\mathbf{X}(x_1, \dots, x_n) = p_{x_1,\dots,X_n}(x_1, \dots, x_n)$ for each $n \in \mathbb{N}$ and $x_1, \dots, x_n \in E$.

Given any discrete stochastic process there is a natural way of viewing the associated joint and conditional probabilities as being associated to a *random walker* moving around on a graph. We call this perspective the classical random walk perspective and describe it below. Fix a discrete (Ω, E) stochastic process $\mathbf{X}$. Then, from the classical random walk perspective, we interpret $p_{x_1}(x)$ as the probability that a random walker inhabits the site x initially at time 1 and $p_{X_n}(x)$ as the probability that a random walker inhabits the site x at time n, for any $x \in E$ and $n \in \mathbb{N}$.

Furthermore, for any $x_1, \ldots, x_n \in E$ and $n \in \mathbb{N}$, we interpret $p_(x_1, \ldots, x_n)$ as the probability that a random walker takes the path $x_1 \to x_1 \to \cdots \to x_n$ at times $0, 1, \ldots, n$. Moreover, if there is a specific graph the random walker is moving on, the conditional probabilities reflect this by giving probability 0 whenever there is no edge for the walker to traverse.

A discrete stochastic process $(X_n)_{n=1}^{\infty}$ is called **stationary** whenever its joint pmf is invariant with respect to shifts of the time index; i.e.

$$\mu(X_1 = x_1, \ldots, X_n = x_n) = \mu(X_l = x_1, \ldots, X_{n+l} = x_n),$$

for all $n, l \in \mathbb{N}$ and $x_1, \ldots, x_n \in E$. In the literature, a stationary stochastic process is sometimes referred to as being time invariant (see e.g. [19, Page 61]).

A simple example of a discrete stochastic process is one in which each random variable depends only on the one proceeding it in the sequence; i.e.

$$\mu(X_{n+1} = x_{n+1} | x_1 = x_1, \ldots, X_n = x_n) = \mu(X_{n+1} = x_{n+1} | X_n = x_n), \tag{2.3}$$

for all $n \in \mathbb{N}$ and $x_1, \ldots, x_{n+1} \in E$. A discrete stochastic process which satisfies Equation (2.3) is referred to as a **Markov process.** In particular, we are interested in those discrete Markov processes, $\mathbf{X} = (X_n)_{n=1}^{\infty}$, whose conditional pmfs do not vary with time; i.e.

$$\mu(X_2 = x | X_1 = y) = \mu(X_{n+1} = x | X_n = y) \text{ for all } x, y \in E, \text{ and } n \in \mathbb{N}. \tag{2.4}$$

In this case, we will set $p_{x,y} := \mu(X_2 = x | X_1 = y)$ and define the $|E| \times |E|$ matrix P to have (x, y)-entry given by $p_{x,y}$, for all $x, y \in E$. Then P is a **transition (column-stochastic) matrix**; i.e. for all $x, y \in E$, $0 \leq p_{x,y} \leq 1$ and, for all $y \in E$, $\sum_{x \in E} p_{x,y} = 1$. From the classical random walk perspective, the (x, y)-entry of P, $p_{x,y}$, is interpreted as the conditional probability that a random walker will move in one step from site y to site x. Moreover, $p_{x,y}$ will be equal to 0 whenever (y, x) is not a (possibly directed) edge on the corresponding graph.

Given a discrete measurable space $(E, \mathcal{P}(E))$, we will represent a probability measure μ on $(E, \mathcal{P}(E))$ as the column vector $\mu = \{\mu_e\}_{e\in E}$, where $\mu_e := \mu(\{e\})$ for each $e \in E$, which we will refer as a **probability vector**. Then, given a transition matrix P on E, we define $P\mu$ by matrix multiplication. We say that $\boldsymbol{\mu}$ **is** $\boldsymbol{P}$**-invariant** whenever $P\mu = \mu$. In particular, whenever $\mathbf{X}$ is a discrete (Ω, E) Markov process governed by the transition matrix P, we take the initial measure μ to be p_{x_1}. In this case, notice that $\mathbf{X}$ is stationary if and only if p_{x_1} is P-invariant.

2.2 Dynamical Systems

Let (Ω, Σ) be a measurable space. Define $\mathcal{P}_{ar}(\Omega)$ to be the collection of all finite or countably infinite measurable partitions of Ω. Define a partial ordering on $\mathcal{P}_{ar}(\Omega)$ such that, for any $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$, $\mathcal{D} \leq \mathcal{C}$ whenever, for every $D \in \mathcal{D}$ there exists $\mathcal{C}_D \subseteq \mathcal{C}$ such that $D = \cup\mathcal{C}_D$. If $\mathcal{D} \leq \mathcal{C}$ we say that $\boldsymbol{\mathcal{C}}$ **is finer than** $\boldsymbol{\mathcal{D}}$ or that $\boldsymbol{\mathcal{D}}$ **is coarser than** $\boldsymbol{\mathcal{C}}$**.**

Whenever Ω is a discrete space we will refer to the partition of Ω into singletons $\{\{\omega\}\}_{\omega\in\Omega}$ as the **atomic partition**. In this case it is clear that the atomic partition $\mathcal{A}$ of Ω is countable and measurable; i.e. $\mathcal{A} \in \mathcal{P}_{ar}(\Omega)$. Furthermore, in this case, $\mathcal{A}$ is the finest partition in $\mathcal{P}_{ar}(\Omega)$; i.e. $\mathcal{C} \leq \mathcal{A}$ for any $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$.

For any $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$, the join (or least upper bound) of $\mathcal{C}$ and $\mathcal{D}$ is given by the partition $\mathcal{C} \vee \mathcal{D}$ which contains all sets of the form $C \cap D$ for all $C \in \mathcal{C}$ and $D \in \mathcal{D}$. Given a finite collection of partitions $\{\mathcal{C}_k\}_{k=1}^n \subseteq \mathcal{P}_{ar}(\Omega)$, the join $\vee_{k=1}^n \mathcal{C}_k$ can be defined recursively from the join of two partitions so that $\vee_{k=1}^n \mathcal{C}_k$ is the partition containing exactly the sets of the form $\cap_{k=1}^n C_k$, where $C_k \in \mathcal{C}_k$ for all $1 \leq k \leq n$.

Fix a probability space (Ω, Σ, μ). Given any partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, consider the discrete space $(\mathcal{C}, \mathcal{P}(\mathcal{C}))$. We then define the **probability measure associated to** $\boldsymbol{\mathcal{C}}$, denoted by $\mu_{\mathcal{C}}$, to be given by

$$\mu_{\mathcal{C}}(C) = \mu(C) \quad \text{for each } C \in \mathcal{C}. \tag{2.5}$$

The measure $\mu_{\mathcal{C}}$ can be thought of as a course-graining of the measure μ. (This idea of course-graining will be discussed further in the paragraph following Example 3.2.3.) Indeed, one can easily define $\mu_{\mathcal{C}}$ as a restriction of μ on (Ω, Σ) analogously to above and think of the sets $C \in \mathcal{C}$ as the possible outcomes of a particular measurement. Measurements will be discussed in further detail in Chapter 3.

We also can now draw our first connection to random walks from probability theory introduced in Section 2.1. To this end, for each $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ we define a discrete $(\Omega, \mathcal{C})$ random variable $X^{\mathcal{C}}$ by setting $X^{\mathcal{C}}(\omega) = C$ whenever $\omega \in C$, for any $C \in \mathcal{C}$. It is easy to see that the pmf of $X^{\mathcal{C}}$ is equal to $\mu_{\mathcal{C}}$; i.e.

$$p_{X^{\mathcal{C}}}(C) = \mu_{\mathcal{C}} \quad \text{for all } C \in \mathcal{C}. \tag{2.6}$$

Furthermore, given any finite collection of partitions $(\mathcal{C}_k)_{k=1}^n \in \mathcal{P}_{ar}(\Omega)$, it is easy to see that

$$p_{X_{\mathcal{C}_1},\ldots,X_{\mathcal{C}_n}}(C_1, \ldots, C_n) = \mu_{\vee_{k=1}^n \mathcal{C}_k}(C_1 \cap \cdots \cap C_n), \tag{2.7}$$

where $C_k \in \mathcal{C}_k$ for all $k \in \{1, \ldots, n\}$ and the joint pmf $p_{X_{\mathcal{C}_1},\ldots,X_{\mathcal{C}_n}}$ is given by Equation (2.1).

Next, we recall that, for any two sets $C, D \in \Sigma$, the **conditional probability of C given D** is given by $\mu(C|D) := \mu(C \cap D)/\mu(D)$. Given a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ and a set $D \in \Sigma$, we define the **conditional probability measure of $\mathcal{C}$ given D** to be given by $\mu_{\mathcal{C}|D}(C) = \mu(C|D)$ for all $C \in \mathcal{C}$. Further, given any two partitions $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$, we define the **conditional probability measure of $\mathcal{C}$ given $\mathcal{D}$** to be given by $\mu_{\mathcal{C}|\mathcal{D}}(C|D) = \mu_{\mathcal{C}|D}(C)$ for all $C \in \mathcal{C}$ and $D \in \mathcal{D}$. Again, it is clear that

$$p_{X^{\mathcal{C}}|X^{\mathcal{D}}}(C|D) = \mu_{\mathcal{C}|\mathcal{D}}(C|D) \quad \text{for all } C \in \mathcal{C} \text{ and } D \in \mathcal{D}, \tag{2.8}$$

where the conditional pmf $p_{X^{\mathcal{C}}|X^{\mathcal{D}}}$ is given by Equation (2.2).

Next we wish to introduce some dynamics on the probability space (Ω, Σ, μ). This role will be played by a measurable map $f : \Omega \to \Omega$. We call the quadruple

(Ω, Σ, μ, f) a **dynamical system** (DS). Furthermore, whenever $\mu(A) = \mu(f^{-1}(A))$ for all $A \in \Sigma$, we say that **μ is f-invariant** and call the DS (Ω, Σ, μ, f) **stationary.** It is worth noting that stationary DSs are often referred to as *measure-preserving* in the literature.

Remark 2.2.1. *In the literature it is common to only refer to* (Ω, Σ, μ, f) *as a DS whenever* μ *is* f*-invariant. We do not adopt that convention here.*

Remark 2.2.2. *In Section 2.1, the dynamics are described by a stochastic process. Indeed, in the case of time-homogeneous Markov processes governed by a transition matrix* P*, it is easy to see that* P *describes explicitly the dynamics of the system (which are probabilistic in nature). For a DS* (Ω, Σ, μ, f)*, the dynamics are played by a deterministic map making the evolution fundamentally different than that of stochastic processes. However, we can still associate probabilities to the dynamics via the probability measures described in Equations* (2.6)*,* (2.7) *and* (2.8)*.*

Fix a DS (Ω, Σ, μ, f) and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$. For each $k \in \mathbb{N}_0$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, define the k^{th}-preimage of $\mathcal{C}$ under f by $f^{-k}(\mathcal{C}) := \{f^{-k}(C)\}_{C \in \mathcal{C}}$, where f^0 denotes the identity map. Note that, for each $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, $f^{-1}(\mathcal{C}) \in \mathcal{P}_{ar}(\Omega)$ and hence $f^{-k}(\mathcal{C}) \in \mathcal{P}_{ar}(\Omega)$ for every $k \in \mathbb{N}_0$. The probabilistic description of the dynamical system is then given by the family of probability measures $(\mu_{\bigvee_{k=0}^{n-1} f^{-k}(\mathcal{C})})_{n=1}^{\infty}$, given in Equation (2.5). We will refer to the family $(\mu_{\bigvee_{k=0}^{n-1} f^{-k}(\mathcal{C})})_{n=1}^{\infty}$ as the **joint probabilities of (Ω, Σ, μ, f) with respect to $\mathcal{C}$**. These joint probabilities can also be thought of as the joint pmfs of the appropriate stochastic process as in Equation (2.7). For any $n \in \mathbb{N}$ and $C_1, \dots, C_n \in \mathcal{C}$, the probability $\mu_{\bigvee_{k=0}^{n-1} f^{-k}(\mathcal{C})}(C_1 \cap \dots \cap f^{-(n-1)}(C_n))$ can be thought of as the probability of a system whose dynamics are determined by f being measured at times $1, 2, \dots, n$ with outcomes $C_1, C_2, \dots, C_n$. We will discuss measurements further in Chapter 3. Due to this interpretation of measurements at subsequent times, we would like to introduce the following notation for the joint probabilities of (Ω, Σ, μ, f):

$$\mu_n^{(f,\mathcal{C})}(C_1, \ldots, C_n) := \mu_{\vee_{k=0}^{n-1} f^{-k}(\mathcal{C})}(C_1 \cap \cdots \cap f^{-(n-1)}(C_n)). \tag{2.9}$$

Remark 2.2.3. *Given a DS (Ω, Σ, μ, f) and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, it is clear that $\vee_{k=0}^{n-1} f^{-k}(\mathcal{C})$ consists exactly of sets of the form $f^{-(n-1)}(C_n) \cap \cdots \cap f^{-1}(C_2) \cap C_1$, for all $C_1, \ldots, C_n \in \mathcal{C}$.*

2.3 Symbolic Dynamics: Random Walks and Dynamical Systems from an Algebraic Point of View

In this section we present the symbolic dynamics representation of both probabilistic random walks (see Section 2.1). The symbolic dynamics representation can be done analogously for dynamical systems and many of the results here hold, although we will not dwell on this point.

The symbolic dynamics (also known as the code space or path space) representation for a stochastic process is a DS which encodes the possible paths that a random walker can take as the points in the space. We will see that the joint pmfs for a discrete stochastic process are encapsulated as the joint probabilities associated to the symbolic dynamics.

Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ a (not necessarily discrete) measurable space and $\mathbf{X} = (X_n)_{n=1}^{\infty}$ an (Ω, E) stochastic process. Consider the measurable space $(E^*, \mathcal{E}^*)$, where $E^* := E^{\mathbb{N}}$ and $\mathcal{E}^* := \sigma(\cup_{n=1}^{\infty} \mathcal{E}^n)$ be the σ-algebra generated by $\cup_{n=1}^{\infty} \mathcal{E}^n$. For all $n \in \mathbb{N}$, collection of integer times $1 \leq t_1 < t_2 < \cdots < t_n$ and $A_1, \ldots, A_n \in \mathcal{E}$, we define the **cylinder set**

$$C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right) := \{x = (x_i)_{i \in \mathbb{N}} \in E^* : x_{t_k} \in A_k \text{ for } k \in \{1, \ldots, n\}\}.$$

For $\mathcal{C} \in \mathcal{P}_{ar}(E)$, we say $C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right)$ is a **$\mathcal{C}$-cylinder set** if $A_1, \ldots, A_n \in \mathcal{C}$. Also, we define the partition, $\hat{\mathcal{C}} \in \mathcal{P}_{ar}(E^*)$, by

$$\hat{\mathcal{C}} := \{C\left(\begin{smallmatrix} A \\ 1 \end{smallmatrix}\right)\}_{A \in \mathcal{C}}$$

and the set

$$\hat{\mathcal{P}_{ar}}(E) := \{\hat{\mathcal{C}} : \mathcal{C} \in \mathcal{P}_{ar}(E)\} \subset \mathcal{P}_{ar}(E^*).$$

Notice that the collection of all cylinder sets in E^* is a π-system which generates the σ-algebra $\mathcal{E}^*$. Therefore, any measure on $(E^*, \mathcal{E}^*)$ is uniquely defined by its values on the cylinder sets. We define the process-dependent measure, $\mu^{\mathbf{X}}$, on the cylinder sets by

$$\mu^{\mathbf{X}}(C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right)) = \mu(\cap_{k=1}^n (X_{t_k} \in A_k)), \tag{2.10}$$

for all $n \in \mathbb{N}$, $1 \leq t_1 < \cdots < t_n$, and $A_1, \ldots, A_n \in \mathcal{E}$. We interpret $\mu^{\mathbf{X}}(C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right))$ as the probability that a random walker, governed by the stochastic process $\mathbf{X}$, is in the set A_1 at time t_1 and takes the path $A_1 \to A_2 \to \cdots \to A_n$ at times $t_1, \ldots, t_n$. Notice that this interpretation is similar to that of the joint pmf for discrete stochastic processes, with the only difference being that the walker's path is through sets here, as opposed to singletons before.

Remark 2.3.1. *Fix an (Ω, E) stochastic process $\mathbf{X}$, $A_1, \ldots, A_n \in \mathcal{E}$, integer times $1 \leq t_1 < \cdots < t_n$ and $t \in \mathbb{N} \backslash \{t_1, \ldots, t_n\}$ such that $t_i < t < t_{i+1}$ for some $1 \leq i < n$. Then, since $\mu(\cap_{k=1}^n (X_{t_k} \in A_k)) = \mu(\cap_{k=1}^n (X_{t_k} \in A_k) \cap (X_t \in E))$, we have that*

$$\mu^{\mathbf{X}}(C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right)) = \mu^{\mathbf{X}}(C\left(\begin{smallmatrix} A_1 & \cdots & A_i & E & A_{i+1} & \cdots & A_n \\ t_1 & \cdots & t_i & t & t_{i+1} & \cdots & t_n \end{smallmatrix}\right)).$$

Similarly, if $1 \leq t < t_1$ or $t > t_n$. Therefore the measure $\mu^{\mathbf{X}}$ is well defined.

We define the shift map

$$s : E^* \to E^* \text{ by } s(x) = y \text{ where } y_i = x_{i+1}, \tag{2.11}$$

for each $i \in \mathbb{N}$, and refer to the quadruple, $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$, as the **symbolic dynamics of X.** Notice that $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$ is a DS.

Of particular interest is the joint probabilities of $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$ with respect to the partitions in $\widehat{\mathcal{P}_{ar}}(E)$. For each $\mathcal{C} \in \mathcal{P}_{ar}(E)$, define the $(\Omega, \mathcal{C})$ stochastic process

$\mathbf{X}_{\mathcal{C}} = (X_n^{\mathcal{C}})_{n=0}^{\infty}$ where, for each $n \in \mathbb{N}$, $X_n^{\mathcal{C}}$ is equal to $X^{s^{-n}(\mathcal{C})}$ defined just before Equation (2.6). More explicitly, we can see that, for each $n \in \mathbb{N}$, $X_n^{\mathcal{C}} = i_{\mathcal{C}} \circ X_n$, where $i_{\mathcal{C}} : E \to \mathcal{C}$ is the natural map that assigns to each $e \in E$ the unique $C \in \mathcal{C}$ such that $e \in C$. Moreover, whenever E is a discrete space and $\mathcal{A}$ is the atomic partition of E, the values of $\mathbf{X}_{\mathcal{A}}$ are singletons and it is clear that $\mathbf{X}$ can be identified with $\mathbf{X}_{\mathcal{A}}$.

The following proposition shows that the joint probabilities of $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$ with respect to $\hat{\mathcal{C}}$ are equal to the joint pmfs of $\mathbf{X}_{\mathcal{C}}$.

Proposition 2.3.2. *Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ a measurable space, $\mathbf{X}$ an (Ω, E) stochastic process and $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$ the symbolic dynamics of $\mathbf{X}$. Then for each $\mathcal{C} \in \mathcal{P}_{ar}(E)$,*

$$\mu_n^{(s,\hat{\mathcal{C}})}(C_1, \ldots, C_n) = p_{\mathbf{X}_{\mathcal{C}}}(C_1, \ldots, C_n), \quad \textit{for each } n \in \mathbb{N} \textit{ and } C_1, \ldots, C_n \in \mathcal{C}.$$

Moreover, whenever E is a discrete space,

$$\mu_n^{(s,\hat{\mathcal{C}})}(x_1, \ldots, x_n) = p_{\mathbf{X}}(x_1, \ldots, x_n), \quad \textit{for each } n \in \mathbb{N} \textit{ and } x_1, \ldots, x_n \in E,$$

where $\mathcal{A}$ is the atomic partition of E and the x_i's are the singleton sets $\{x_i\}$, for each i.

Proof. This is an immediate consequence of Equation (2.7) using the notation introduced in Equation (2.9). The moreover statement is just a special case due to the facts that, whenever E is a discrete space, the atomic partition $\mathcal{A}$ is in $\mathcal{P}_{ar}(E)$ and $\mathbf{X}$ can be identified with $\mathbf{X}_{\mathcal{A}}$. □

A good resource for symbolic dynamics of Markov processes is [31].

Remark 2.3.3. *Let (Ω, Σ, μ, f) be a dynamical system and $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ a partition. We have already seen in Equation (2.6) that we can associate random variables to the course-grained measurements associated to a particular partition. Moreover, by extending Equation (2.7) to account for any finite joins of the form $\vee_{k=0}^{n-1} f^{-k}(\mathcal{C})$, it is* *easy to see that the joint probabilities of* (Ω, Σ, μ, f) *with respect to* $\mathcal{C}$ *are exactly equal to the joint pmfs of the associated stochastic process (defined analogously to above). In this case, we denote the measure on the symbolic dynamics DS* $(\Omega^*, \Sigma^*, s, \hat{\mathcal{C}})$ *with respect to a partition* $\mathcal{C}$ *by* $\hat{\mu}$.

2.4 Dynamical Systems from the Algebraic Point of View

Let (Ω, Σ, μ, f) be a DS and consider the collection of all uniformly bounded functions on Ω, $L^\infty(\Omega)$. Let $T_f : L^\infty(\Omega) \to L^\infty(\Omega)$ be the **Koopman operator** given by

$$T_f(g) = f \circ g \quad \text{for all } g \in L^\infty(\Omega). \tag{2.12}$$

It is well known that T_f satisfies the following properties (see e.g. [11]):

(i) T_f is a $*$-automorphism

(ii) $\|T_f\| = 1$

In further establish connections with random walks on a measurable space (Ω, Σ), we say that a mapping $\nu : \Omega \times \Sigma \to [0, 1]$ to be a **Markov kernel** if

(i) $x \mapsto \nu(x, A)$ is measurable for every $A \in \Sigma$, and

(ii) $A \mapsto \nu(x, A)$ is a probability measure for all $x \in \Omega$.

Whenever (Ω, Σ) is discrete, ν is determined by the transition matrix P with entries

$$P(j, i) = \nu(i, \{j\}), \quad \text{for all } i, j \in \Omega.$$

A Markov kernel acts on $L^\infty(\Omega)$ in the following way

$$\nu \circ f(x) := \int_\Omega f(y)\nu(x, dy). \tag{2.13}$$

Furthermore, any linear operator T that acts on $L^\infty(\Omega)$ of the form

$$Tf(x) = \int_\Omega f(y)\nu(x, dy), \tag{2.14}$$

for some Markov kernel ν, is called a **Markov operator**. In a dual manner, we can consider ν acting on the probability measures on (Ω, Σ) (see also Example 3.2.3). Given any probability measure $\mathbb{P}$ on (Ω, Σ), the measure $\mathbb{P} \circ \nu$ given by

$$\mathbb{P} \circ \nu(A) := \int_\Omega \nu(x, A)\mathbb{P}(dx) \tag{2.15}$$

is also a probability measure.

Remark 2.4.1. *In a similar manner to above, we can also associate to any probability measure, $\mathbb{P}$, on (Ω, Σ) a linear functional $\mathbb{P} : L^\infty(\Omega) \to \mathbb{R}$ given by $\mathbb{P}(g) = \int_\Omega g d\mathbb{P}$ for all $g \in L^\infty(\Omega)$. This is a slight abuse of notation, but the context should make it clear whenever $\mathbb{P}$ is being used as a probability measure or a linear functional. (See Example 3.2.3.)*

A key idea which is useful when considering a type of symbolic dynamics for DSs in the quantum regime (see Sections 4.3), specifically with open system dynamics, is what happens when we consider a DS acting on a product space. To this end, let (Ω_s, Σ_s) and (Ω_e, Σ_e) be measurable spaces, where the subscripts s and e are to remind us that Ω_s is the system of interest and Ω_e is the environment, and let f be a map from $\Omega_s \times \Omega_e$ to itself with associated Koopman operator T_f. To each $g \in L^\infty(\Omega_s)$ set $g \otimes \mathbb{1} \in L^\infty(\Omega_s \times \Omega_e)$ be given by

$$g \otimes \mathbb{1}(x, y) = g(x), \quad \text{for all } (x, y) \in \Omega_s \times \Omega_e.$$

Given a probability measure μ on (Ω_e, Σ_e), we define the mapping $T_s : L^\infty(\Omega_s) \to L^\infty(\Omega_s)$ to be given by

$$T_s(g)(x) := \int_{\Omega_e} T_f(g \otimes \mathbb{1})(x, y)d\mu(y) = \int_{\Omega_e} (g \otimes \mathbb{1})f(x, y)d\mu(y). \tag{2.16}$$

We can think of the map T_s as describing a single component of the dynamics of T_f (or f); i.e. when we only have access to Ω_s. It turns out that T_s is indeed a Markov operator (see [11, Theorem 2.2]) and hence we can treat the dynamics on Ω_s exactly as we would without an interaction with an environment. On the other hand, it is possible for T_s to be probabilistic in nature even when T_f is deterministic. This idea emphasizes the loss of information to an environment. It is also worth noting it is not true in general that, for any $n \in \mathbb{N}$, $(T_f^n)_s$, given by Equation (2.16) for the n^{th} iterate of T_f, is not necessarily equal to T_s^n (see [11, Section 2.3]). When extending DSs to the quantum regime, we will revisit the idea here of having access to only one component of a coupled system.

Let us turn back now to the DS (Ω, Σ, μ, f) and let T_f be the associated Koopman operator. To each partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ we can associate a **partition of unity** in $L^\infty(\Omega)$ given by the collection of characteristic functions $\gamma = \{\mathbb{1}_C\}_{C \in \mathcal{C}}$. Partitions of unity will be defined in a greater generality in Section 3.2 where they will be used as the most basic class of measurements for quantum systems. To the DS (Ω, Σ, μ, f), we associate the triple $(L^\infty(\Omega), \mu, T_f)$. This triple is the simplest example of a quantum dynamical system which will be defined in Section 4.2. In that section and the ones succeeding it we will give many different generalizations with similar probabilistic interpretations for the system $(L^\infty(\Omega), \mu, T_f)$, while also utilizing our intuition from the coupled classical system above.

Chapter 3

Measurements

Here we recall the formalism of measurements, developed by Edwards [23] and Davies and Lewis [21, 20]. We follow mainly the notations of Davies and Lewis. We define phase space, state space, observables and instruments. This formalism is general enough that it holds valid for classical mechanics, Hilbert space quantum mechanics, and C^*-algebra quantum mechanics.

3.1 State Space and Phase Space

A **state space** is defined as a pair (X, K), where

(i) X is a real Banach space with norm $\|\cdot\|$,

(ii) K is a closed cone in X,

(iii) if $u, v \in K$, then $\|u\| + \|v\| = \|u + v\|$, and

(iv) if $u \in X$ and $\epsilon > 0$, then there exists $u_1, u_2 \in K$ such that $u = u_1 - u_2$ and $\|u_1\| + \|u_2\| < \|u\| + \epsilon$.

It can be shown that, for any state space (X, K), there exists a unique positive linear functional $\tau : X \to \mathbb{R}$ such that $\tau(u) \leq \|u\|$, for $u \in X$, with equality when $u \in K$. We say that $u \in K$ is a **state** if $\tau(u) = 1$. It should be remarked that all examples of state spaces presented here will satisfy a strengthening of (iv); namely,

(iv′) if $u \in X$, then there exists $u_1, u_2 \in K$ such that $u = u_1 - u_2$ and $\|u_1\| + \|u_2\| = \|u\|$.

The following examples of state spaces appear most frequently in the literature.

(A) *Classical Mechanics* - Let Ω be a locally compact Hausdorff space and $\mathcal{B}$ be the Borel σ-algebra of Ω. Take X to be the real Banach space of all countably additive, regular, real-valued Borel measures on X and take the norm on X to be the total variation norm and the closed cone, K, to be the set of non-negative measures in X. It is clear that (X, K) satisfies conditions (i)-(iii) and (iv$'$) of a state space by taking u_1 and u_2 to be the positive and negative parts, respectively, of $u \in X$ given by the Hahn decomposition. Furthermore, the linear functional τ is given by $\tau(\nu) = \int_\Omega d\nu = \nu(\Omega)$ for any $\nu \in X$.

(B) *Hilbert Space Quantum Mechanics* - Let H be a Hilbert space. Take $X = S_1^{sa}(H)$ to be the real Banach space of self-adjoint, trace class operators on H equipped with the trace class norm and the closed cone, $K = S_1^+(H)$, to be collection of all the positive, trace class operators on H. It is clear that the state space (X, K) satisfies conditions (i)-(iii) and (iv$'$) of a state space by taking u_1 and u_2 to be the positive and negative parts, respectively, of $u \in X$ given by the functional calculus. Furthermore, the linear functional τ is given by the trace, tr.

(C) *C^*-Algebra Quantum Mechanics* - Let $\mathcal{A}$ be a C^*-algebra. Let X be the real Banach space of all bounded, self-adjoint linear functionals on $\mathcal{A}$ with the dual space norm. Set the positive cone $K = \{\omega \in X : \omega(a^*a) \geq 0 \text{ for all } a \in \mathcal{A}\}$. It is clear that the state space (X, K) satisfies conditions (i)-(iv) of a state space. Also, the linear functional τ is given by $\tau(\cdot) = \|\cdot\|$.

It is worth noting that both the classical and Hilbert space mechanics are a special case of the C^*-algebra quantum mechanics if $\mathcal{A}$ is taken to be the algebra of all continuous functions on Ω vanishing at infinity (also see Section 2.4), or if $\mathcal{A}$ is taken to be the algebra of all compact operators on H, respectively.

A **phase space** is defined as an arbitrary measurable space (Ω, Σ), where Ω

represents the collection of possible outcomes of a measurement and is sometimes called the value space in the literature.

Remark 3.1.1. *Throughout this manuscript we try to avoid complicating things with the GNS construction when considering C^*-algebra mechanics whenever possible. Instead we will often assume that the C^*-algebra $\mathcal{A}$ is a subset of $B(H)$ for some Hilbert space H, where $B(H)$ is the bounded operators on H. Whenever this is the case, it will be convenient to consider only the states and general state space elements from the pre-dual of $\mathcal{A}$, $\mathcal{A}_*$, which is known to be a subset of $S_1(H)$ by the trace duality.*

3.2 Observables and Instruments

We say that $x : \Sigma \to X^*$ is an **observable** if, for every $E \in \Sigma$, $0 \leq x(E) \leq \tau$ and $x(\Omega) = \tau$, where the partial ordering on X^* is defined by $\phi \leq \psi$ whenever $\phi(u) \leq \psi(u)$ for all $u \in K$. Given a state $u \in K$, an observable x, and $E \in \Sigma$, we interpret $x(E)u$ as the probability that a system in state u takes values in E when observed with the observable x.

An **operation** is a positive, bounded linear operator $T : X \to X$, such that $0 \leq \tau(Tu) \leq \tau(u)$ for every $u \in K$. We denote by $\mathcal{O} := \mathcal{O}(X)$ the set of all operations on X. Finally, we define an **instrument** as a map $\mathcal{T} : \Sigma \to \mathcal{O}$ such that $\tau(\mathcal{T}(\Omega)u) = \tau(u)$, for all $u \in K$, and $\mathcal{T}(\cup_n E_n) = \sum_n \mathcal{T}(E_n)$, for any disjoint sequence of sets $\{E_n\} \subseteq \Sigma$, where convergence of the sum is in the strong operator topology.

Notice that for any instrument $\mathcal{T}$, one can define a unique observable $x_{\mathcal{T}}$ by setting $x_{\mathcal{T}}(E)u = \tau(\mathcal{T}(E)u)$ for $u \in X$ and $E \in \Sigma$. However, it is possible that two distinct instruments, $\mathcal{T} \neq \mathcal{S}$, give rise to the same observable, $x_{\mathcal{T}} = x_{\mathcal{S}}$. From the above correspondence, given an initial state $u \in K$ and $E \in \Sigma$, we can interpret $\mathcal{T}(E)u/x_{\mathcal{T}}(E)u \in K$ as the state of the system immediately after measuring the system in state u with the instrument $\mathcal{T}$ and obtaining values in the set E.

Next, we give the definitions of partitions and operational partitions of unity that are generally used in C^*-algebra quantum mechanics before introducing their typical use and formulation as instruments for each of the Examples (A) and (B) above. This is possible since each of Examples (A) and (B) are a special case of C^*-algebra quantum mechanics given in Example (C).

Example 3.2.1 (C^*-Algebra Quantum Mechanics)**.** *Let $\mathcal{A}$ be a C^*-algebra with identity $\mathbb{1}$ and let (X, K) and τ be as in Example (C) above. A (finite)* ***operational partition of unity*** *is a collection $\gamma = (\gamma_k)_{k=1}^n \subseteq \mathcal{A}$, for some $n \in \mathbb{N}$, such that*

$$\sum_{k=1}^{n} \gamma_k^* \gamma_k = \mathbb{1}.$$

Given any operational partition of unity γ of size n and any subset E of $\{1, \ldots, n\}$ it is clear that the linear operator, T_E^γ acting on the bounded, self-adjoint linear functionals on $\mathcal{A}$, given by

$$T_E^\gamma(\cdot) := \sum_{k \in E} \gamma_k^* \cdot \gamma_k \tag{3.1}$$

is indeed an operation. We define an instrument $\mathcal{T}_\gamma : \mathcal{P}(\{1, \ldots, n\}) \to \mathcal{O}$ by setting $\mathcal{T}_\gamma(E) = T_E^\gamma$ for each $E \in \mathcal{P}(\{1, \ldots, \})$. We will say $\mathcal{T}_\gamma$ is ***governed by $\boldsymbol{\gamma}$*** *and will simply write $\mathcal{T}$ whenever the operational partition of unity is clear. Of particular interest are the operational partitions of unity $\gamma = (\gamma_k)_{k=1}^n$ such that $\gamma_k^* = \gamma_k$ for each k and $\gamma_k \gamma_i = \delta_{k,i} \gamma_k$ for each k, i. Operational partitions of unity of this form will be simply referred to as* ***partitions of unity****.*

Remark 3.2.2. *One has to take particular care when defining the operations T_E^γ in Equation* (3.1) *due to the different pictures (Schrödinger or Heisenberg) that are often employed when considering quantum mechanics. This point will be discussed further Example 3.2.4 below.*

We are now ready revisit our examples of state spaces and introduce typical instruments that will be applied to them along with their corresponding observables. First, we consider Example (A) and define the sharp measurement instruments.

Example 3.2.3 (Classical Mechanics)**.** *Let Ω, $\mathcal{B}$ and (X, K) be as in Example (A) above and let the measurable space $(\Omega, \mathcal{B})$ be the phase space. We define the (classical)* ***sharp measurement instrument*** *$\mathcal{T}$ by*

$$\mathcal{T}(E)\nu(A) = \nu(A \cap E) \quad \textit{for } \nu \in X \textit{ and } A, E \in \mathcal{B}. \tag{3.2}$$

The corresponding observable is given by

$$x_{\mathcal{T}}(E)\nu = \tau(\mathcal{T}(E)\nu) = \nu(E) \quad \textit{for } E \in \mathcal{B} \textit{ and } \nu \in X.$$

Let $(\Omega, \mathcal{B})$ and (X, K) be as in Example 3.2.3 and let $\mathcal{A} = c_0(\Omega)$ be the algebra of all continuous functions on Ω vanishing at infinity. Consider a partition of unity $\gamma = (\gamma_k)$. By definition it is clear γ can be identified with partition $\mathcal{C} = (C_k) \in \mathcal{P}_{ar}(\Omega)$, exactly as in Section 2.4, through the identification $\gamma_k = \mathbb{1}_{C_k}$ for each k. Since $c_0(\Omega)$ is commutative, T_E^γ in Equation (3.1) simplifies to $T_E^\gamma(\cdot) = \sum_{k \in E} \mathbb{1}_{C_k} \cdot$ for each $E \in \mathcal{P}(\{1, \ldots, k\})$. The instrument $\mathcal{T}_\gamma$ can then be thought of as a *course measurement instrument*, as opposed to the sharp measurement instrument considered in Example 3.2.3. The key difference between course and sharp measurement is that for the course measurement only a predetermined collection of measurement outcomes is possible (corresponding the the partition $\mathcal{C}$), whereas for sharp measurements any outcome is possible.

The next example illustrates measurements in the Hilbert space formalism of quantum mechanics with discrete phase space determined by an operational partition of unity. This formalism will be used in our analysis of quantum random walks in Section 4.1 and Słomczyński-Życzkowski dynamical entropy in Section 6.

Example 3.2.4 (Hilbert Space Quantum Mechanics)**.** *Let H, (X, K) and τ be as in Example (B) and let $\gamma = (\gamma_k)_{k=1}^n$ be an operational partition of unity. Let $(\Omega, \mathcal{P}(\Omega))$ be a discrete phase space of size n. Here, we are in the Schrödinger picture of quantum* *mechanics and we define the instrument $\mathcal{T} = \mathcal{T}_\gamma$ by*

$$\mathcal{T}(E)\rho = \sum_{i \in E} B_i \rho B_i^* \quad \textit{for each } \rho \in X \textit{ and } E \in \mathcal{P}(\Omega),$$

where the sums are taken with respect to the strong operator topology if Ω is countably infinite. In this case, $\mathcal{T}$ is really the dual (or pre-dual depending on context) of the instrument defined via the operations in Equation (3.1) *and will generally be applied to density matrices on H. Moreover, when restricted to actions on K, $\mathcal{T}$ represents a positive operator valued measure. The corresponding observable is given by*

$$x_{\mathcal{T}}(E)\rho = \sum_{i \in E} \operatorname{tr}(B_i \rho B_i^*) \quad \textit{for each } \rho \in X \textit{ and } E \in \mathcal{P}(\Omega).$$

Whenever γ is a partition of unity in Example 3.2.4; i.e. the family of γ_k's are pairwise orthogonal projections, we will often denote them by $(P_k)_{k=1}^n$ and note that $\sum_{k=1}^n P_k^* P_k = \sum_{k=1}^n P_k = \mathbb{1}$. In this case, the corresponding instrument $\mathcal{T}$, governed by γ, is called a **Lüders-von Neumann instrument** and is given by

$$\mathcal{T}(E)\rho = \sum_{k \in E} P_k \rho P_k \quad \text{for } \rho \in X \text{ and } E \in \mathcal{P}(\Omega),$$

where the sums are taken with respect to the strong operator topology if Ω is countably infinite. It is worth noting that $\mathcal{T}$ is defined by the "collapse of wave function formula." The corresponding observable is defined analogously.

Whenever the partition of unity $(P_k)_{k=1}^n$ consists of orthogonal, rank-1 projections, the Lüders-von Neumann instrument $\mathcal{T}$ is called a **coherent states instrument** (see [59, Section IV].) In this paper whenever we refer to a coherent states instrument we will always mean a Lüders-von Neumann instrument given by a family of orthogonal, rank-1 projections as opposed to the more general definition given in [59, Example (M)].

Chapter 4

Non-Commutative Random Walks

In this chapter we recall three different descriptions of discrete-time dynamics in the quantum (or non-commutative) setting. We motivate each description by relating to its classical counterpart and discuss some connections between the different descriptions.

4.1 Quantum Random Walks

Quantum random walks (QRWs) are a generalization of classical random walks to quantum mechanics, the properties of which are significantly different from their classical courterparts (see e.g. [8]). QRWs come in two flavors: unitary QRWs (UQRWs), introduced by Aharonov et al. in [5] and independently by Meyer in [43], used to describe closed system dynamics, and open QRWs (OQRWs), introduced by Attal et al. in [12], for open system dynamics. As the name suggests, UQRWs evolve via unitary transformation, whereas OQRWs evolve via completely positive trace preserving maps (quantum channels). The study of QRWs has enjoyed much interest in recent years [7, 29, 30, 34, 48, 63], and applications have been found in many areas including quantum computing [35, 37], the study brain networks [41], and biology [50, 51].

Both unitary and open QRWs are defined, using Hilbert Space Quantum Mechanics (see Example 3.2.4), on the tensored Hilbert space $H = H_C \otimes H_P$, where the coined Hilbert space, H_C, is meant to represent the internal degrees of freedom and the position Hilbert space, H_P, is meant to represent the position of a random walker. The position Hilbert Space H_P is a separable Hilbert space with some orthonormal basis labeled by a finite or countably infinite vertex set V. For this reason, it is convenient to keep in mind the probabilistic/graph-theoretic random walk model, as a classical comparison, in this section.

4.1.1 Unitary quantum random walks

The unitary quantum random walk (UQRW) is one of the many adaptations of the classical random walk to the quantum domain and, in particular, is an adaptation of classical random walks for closed quantum systems; i.e. systems that do not interact with an environment. Similar to the probabilistic and graph-theoretic random walk perspective (Section 2.1), we define the UQRW on a finite or countably infinite vertex set V.

To consider a collection of vertices in the quantum domain, Hilbert space quantum mechanics is used (see Example 3.2.4) and we consider the **position space**, $H_P := \ell_2(V)$, with an orthonormal basis, $\{|v\rangle\}_{v\in V}$, indexed by V. To add internal degrees of freedom to the vertices, the **coin space** H_C is used, which is an at most countably-dimensional Hilbert space. In general, a UQRW is given by the unitary transformation over the tensored Hilbert space $H = H_C \otimes H_P$. The most common UQRWs are the so-call coined UQRWs. To define these we must first fix an orthonormal basis, $\{|c, v\rangle\}_{(c,v)\in C\times V}$ on H, sometimes referred to as the computational basis, for some index set C, where $|c, v\rangle = |c\rangle \otimes |v\rangle$. We say that a UQRW is **coined** if it is the unitary transformation of an operator U of the form

$$U = S(\sum_{v\in V} U_v \otimes |v\rangle\langle v|), \tag{4.1}$$

where U_v is a unitary operator on H_C for each $v \in V$, and S is a permutation operator which is referred to as the **shift operator**. By a "permutation operator", we mean that S has the form

$$S = \sum_{(c,v)\in C\times V} |\sigma(c, v)\rangle\langle c, v|, \tag{4.2}$$

for some permutation σ of $C \times V$. For each $v \in V$, the unitary operator U_v, referred to as the coin operator at v, changes the coin state at the vertex v in a deterministic way while the shift operator S moves the random walker from one site to another. In the sequel we will only consider coined UQRWs and we will drop the adjective. We say that a UQRW is **space homogeneous** if there exists a unitary operator W on H_C such that $W = U_v$ for every $v \in V$. For a space homogeneous UQRW the form of U in Equation (4.1) simplifies to

$$U = S(W \otimes \mathbb{1}_{H_P}).$$

We say that a shift operator S is **coin-preserving** if, for each $c \in C$, there exists a permutation σ_c of V such that

$$S = \sum_{(c,v) \in C \times V} |c, \sigma_c(v)\rangle\langle c, v|.$$

Furthermore, we say that a UQRW is coin-preserving if its shift operator is coin-preserving. Notice that a coin-preserving shift operator moves the random walker from site to site without affecting the internal state of the walker.

Let U be a unitary operator of the form Equation (4.1) with shift operator, S, given by Equation (4.2). To draw a connection to classical random walks it is helpful to visualize a random walker on the directed graph $G = (V, E)$ where E is the edge set determined by the shift operator S. That is to say, for all $u, v \in V$, $(u, v) \in E$ if and only if there exists $c_1, c_2 \in C$ such that $\sigma(c_1, u) = (c_2, v)$, where σ is the map appearing in Equation (4.2), or, equivalently, $P_v S P_u \neq 0$, where $P_v = \mathbb{1}_{H_C} \otimes |v\rangle\langle v|$; i.e. P_v is the projection from H to $H_C \otimes \text{span}(v)$.

We are specifically interested in the Hadamard walk, which has been studied extensively in the literature (e.g. [8, 29, 48]) and is defined below. Consider the vertex set $V = \{0, \ldots, N-1\}$, for some $N \in \mathbb{N}$ with $N \geq 2$, and set $H_P = \mathbb{C}^N$. Let $H_C = \mathbb{C}^2$, with orthonormal basis $\{|R\rangle, |L\rangle\}$. Define the (coin-preserving) integer shift operator by

$$S = \sum_{n=0}^{N-1} |R, n+1\rangle\langle R, n| + |L, n-1\rangle\langle L, n|,$$

where addition on the integers is done modulo N. Throughout the rest of this section addition (on V) will be done modulo N. Notice that $|R\rangle$ now corresponds to a shift right on the integers and $|L\rangle$ corresponds to a shift left on the integers. In this case the directed graph $G = (V, E)$ has edge set which is given by $E = \{(n, n+1), (n, n-1)\}_{n=0}^{N-1}$. The unitary operator

$$h := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

on H_C is referred to as the Hadamard matrix (or Hadamard coin/gate). The **Hadamard walk on $\boldsymbol{V}$** is the map $\Theta : X \to X$, where (X, K) is the state space defined in Example 3.2.4, given by

$$\Theta(\rho) = U\rho U^*, \text{ for each } \rho \in X, \text{ where } U = S(h \otimes \mathbb{1}_{H_P}). \tag{4.3}$$

It is clear that the Hadamard walk is coin-preserving and space homogeneous.

Remark 4.1.1. *The Hadamard walk can easily be extended to $V = \mathbb{Z}$ as opposed to the finite cycle $V = \{0, \ldots, N-1\}$ and it is often viewed in this manner. (e.g. [48, Section 5.1])*

4.1.2 Open quantum random walks

In this subsection we recall the definitions of open quantum random walks (OQRWs) introduced by Attal, et. al in [12].

Again, we begin with the tensored Hilbert space $H = H_C \otimes H_P$, where the coined Hilbert space, $H_C = \mathbb{C}^d$, is meant to represent the, $d \in \mathbb{N}$, internal degrees of freedom (or chirality) for a walker and the position Hilbert space, $H_P = \mathbb{C}^V$ (or more generally $\ell_2(V)$ if V is countably infinite), is meant to represent the position of a random walker on an at most countable vertex set V. We will retain all notations from the previous subsection.

A completely positive, trace-preserving (CP-TP) map (or quantum channel) $\mathcal{M} : S_1(H) \to S_1(H)$, where $S_1(H)$ is the trace-class operators on H, is an *open quantum random walk* if it has the following Kraus decomposition:

$$\mathcal{M}(\rho) := \sum_{i,j \in V} M_{i,j} \rho M_{i,j}^*, \quad \text{for all } \rho \in S_1(H), \tag{4.4}$$

where $M_{i,j} = B_{i,j} \otimes |i\rangle\langle j|$ for some $B_{i,j} \in B(H_C)$ for each $i, j \in V$. The operator $B_{i,j}$ is meant to describe the change in the coin-state degrees of freedom when the random walker moves from site j to site i.

It is clear that an OQRW, given by Equation (4.4), must satisfy

$$\sum_{i,j \in V} M_{i,j}^* M_{i,j} = \mathbb{1}_H, \tag{4.5}$$

or equivalently

$$\sum_{i \in V} B_{i,j}^* B_{i,j} = \mathbb{1}_{H_C}, \quad \text{for each } j \in V. \tag{4.6}$$

Note that for any OQRW $\mathcal{M}$ and state $\rho \in S_1(H)$ the output state is of the form

$$\mathcal{M}(\rho) = \sum_{i \in V} \rho_i \otimes |i\rangle\langle i|, \tag{4.7}$$

where $\rho_i \in S_1(H_C)$ for each $i \in V$.

4.1.3 A connections between open and unitary quantum random walks

In their seminal paper on OQRWs ([12]), the authors show that unitary and open QRWs differ only by a single step in their realization procedure. Namely, the step that requires decoherence and hence simulation of interaction with an environment. If we recall that measurements can be thought of as an interaction with an environment, we find a special case of OQRWs as UQRWs with measurement, as you will see. In what follows, we consider only space-homogeneous, coin-preserving UQRWs for simplicity, but the result extends easily to general coined UQRWs.

Suppose Θ is a space-homogeneous, coin-preserving UQRW governed by a unitary operator $U = S(W \otimes \mathbb{1}_{H_P})$, let $\gamma = (P_v)_{v\in V}$, where the projections defining $\mathcal{T}$ are given by

$$P_v = \mathbb{1}_{H_C} \otimes |v\rangle\langle v|, \text{ for each } v \in V,$$

and let $\mathcal{T} =$ be the corresponding Lüders - von Neumann instrument governed by γ. Consider the shift operator's decomposition

$$S = \sum_{(c,v)\in C\times V} |c, \sigma_c(v)\rangle\langle c, v|$$

for some permutations σ_c of V. For each $c \in C$, let B_c be the matrix whose c^{th} row is equal to the c^{th} row of W and has every other entry equal to 0. Now, for each $u, v \in V$, set $B_{u,v} = \sum_{\substack{c\in C\\ \sigma_c(v)=u}} B_c$.

Then, for each $v \in V$, notice that

$$\sum_{u\in V} B^*_{u,v}B_{u,v} = \sum_{c\in C} B^*_c B_c = \mathbb{1}_{H_C},$$

by the definition of the B_c's since W is unitary. Thus,

$$\mathcal{M}(\cdot) = \sum_{u,v\in V} (B_{u,v} \otimes |u\rangle\langle v|) \cdot (B^*_{u,v} \otimes |v\rangle\langle u|) \tag{4.8}$$

is an OQRW. Moreover, given any $\rho =\in S_1(H)$ of the form Equation (4.7), which will be its form after one application of $\mathcal{M}$, we have

$$\begin{aligned}
\mathcal{T} \circ \Theta(\rho) &= \mathcal{T}(S(\sum_{v\in V} W\rho_v W^* \otimes |v\rangle\langle v|)S^*) \\
&= \mathcal{T}(\sum_{c,d\in C}\sum_{v\in V} B_c\rho_v B^*_d \otimes |\sigma_c(v)\rangle\langle\sigma_d(v)|) = \mathcal{M}(\rho).
\end{aligned}$$

Thus, $\mathcal{M} = \mathcal{T} \circ \Theta$ defines a special case of OQRWs.

4.2 Quantum Dynamical Systems

Let $(\mathcal{A}, \Sigma(\mathcal{A}))$ be a von-Neumann algebraic system (also referred to as an algebraic probability space in the literature) with $\Sigma(\mathcal{A})$ the set of all normal states on the von

Neumann algebra $\mathcal{A}$. Throughout this section we will, for simplicity, ignore the GNS construction and assume that $\mathcal{A} \subseteq B(H)$ for some separable Hilbert space H and we will identify each normal state, $\omega \in \Sigma(\mathcal{A})$ with its density operator $\rho \in S_1(H)$, the space of trace class operators on H, through the identification $\omega(\cdot) = \mathrm{tr}(\rho\ \cdot)$ (see Remark 3.1.1). An algebraic probability space together with an automorphism Θ and an initial state $\rho \in S_1(H)$ will be denoted by the triple $(\mathcal{A}, \Theta, \rho)$ and referred to as a **quantum dynamical system** (QDS). We will be mainly interested in **stationary** quantum dynamical systems; i.e. $\omega \circ \Theta = \omega$ or equivalently $\Theta^*(\rho) = \rho$, where Θ^* is the dual of Θ.

Fix a QDS $(\mathcal{A}, \Theta, \rho)$. Similar to classical dynamical systems in Section 2.2 we wish to define joint probabilities associated to course-grained measurements determined by a fixed partition. Two key differences in QDSs are that, instead of pmfs (as in the classical case), probabilities in quantum mechanics are determined by density operators and we are allowed more general partitions; i.e. operational partitions of unity given in Example 3.2.1. To each operational partition of unity $\gamma = (\gamma_k)_{k=1}^d$ of $\mathcal{A}$ we define the associated density operator $\rho[\gamma] \in M_d$ having (i,j)-entry

$$\rho[\gamma]_{i,j} := \omega(\gamma_j^* \gamma_i) = \mathrm{tr}(\gamma_i \rho \gamma_j^*), \quad \text{for each } i, j \in \{1, \ldots, d\}. \tag{4.9}$$

$\rho[\gamma]$ can be thought of as the density matrix describing the initial state ρ with the measurement determined by γ.

To describe the joint probabilities associated to the QDS $(\mathcal{A}, \Theta, \rho)$ with respect to γ, we alternate evolving with Θ and measuring with γ and arrive at a density operator $\rho^{(n)}[\gamma] \in M_d^{\otimes n} = M_{d^n}$ whose $(\bar{i}, \bar{j})$-entry is given by

$$\begin{aligned} \rho^{(n)}[\gamma]_{\bar{i},\bar{j}} &= \omega(\gamma_{j_1}^* \Theta(\cdots \Theta(\gamma_{j_n}^* \gamma_{i_n}) \cdots) \gamma_{i_1}) \\ &= \mathrm{tr}(\gamma_{i_n} \Theta^*(\cdots \Theta^*(\gamma_{i_1} \rho \gamma_{j_1}^*) \cdots) \gamma_{j_n}^*), \end{aligned} \tag{4.10}$$

for each $n \in \mathbb{N}$, where $\bar{i} = (i_1, \ldots, i_n)$ and $\bar{j} = (j_1, \ldots, j_n)$. We will refer to density matrices $\rho^{(n)}[\gamma] \in M_d^{\otimes n}$ as the **joint densities of $(\mathcal{A}, \Theta, \rho)$ with respect to γ**.

See also [38, Equation 3.20].

Next we define the join of operational partitions of unity in order to describe the correlated probabilities of the QDS across time. Given any two operational partitions of unity $\gamma = (\gamma_k)_{k=1}^{d}$ and $\lambda = (\lambda_l)_{l=1}^{m}$, we define the **join of $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda}$** to be the operational partition of unity

$$\gamma \circ \lambda = \{\gamma_k \lambda_l : k \in \{1, \dots, d\} \text{ and } l \in \{1, \dots, m\}\}. \tag{4.11}$$

Note that, due to the non-commutativity of $\mathcal{A}$, the join of operational partitions of unity is not commutative; i.e. $\gamma \circ \lambda \neq \lambda \circ \gamma$. The join of any finite number of operational partitions of unity can be defined using Equation (4.11) recursively.

Given any stationary QDS $(\mathcal{A}, \Theta, \rho)$, operational partition of unity $\gamma = (\gamma_k)_{k=1}^{d}$ and integer time $n \in \mathbb{N}$, let $\Theta^n(\gamma) = (\Theta^n(\gamma_k))_k$ and consider the density operator $\rho[\Theta^{n-1}(\gamma) \circ \cdots \circ \Theta(\gamma) \circ \gamma]$. If Θ is a $*$-automorphism; i.e. $\Theta(a^*) = \Theta(a)^*$ for each $a \in \mathcal{A}$ and $\Theta(ab) = \Theta(a)\Theta(b)$ for each $a, b \in \mathcal{A}$, then Equation (4.10) simplifies to

$$\begin{aligned} \rho^{(n)}[\gamma]_{\bar{i},\bar{j}} &= \omega(\gamma_{j_1}^* \Theta(\gamma_{j_2})^* \cdots \Theta^{n-1}(\gamma_{j_n})^* \Theta^{n-1}(\gamma_{i_n}) \cdots \Theta(\gamma_{i_2})\gamma_{i_1}) \\ &= \operatorname{tr}(\gamma_{i_n} \Theta^*(\gamma_{i_{n-1}}) \cdots \Theta^{n-1*}(\gamma_{i_1}) \rho \Theta^{n-1*}(\gamma_{j_1}^*) \cdots \Theta^*(\gamma_{j_{n-1}})^* \gamma_{j_n}^*), \end{aligned} \tag{4.12}$$

for each $n \in \mathbb{N}$ and $\bar{i}, \bar{j} \in \{1, \dots, d\}^n$, where the last equality holds since ρ is invariant with respect to Θ^*. Therefore, in this case,

$$\rho^{(n)}[\gamma] = \rho[\Theta^{n-1}(\gamma) \circ \cdots \circ \Theta(\gamma) \circ \gamma] \quad \text{for each } n \in \mathbb{N}. \tag{4.13}$$

We finish off this section by showing that for any classical DS (Ω, Σ, f, μ), the associated (commutative) QDS on $L^\infty(\Omega)$ with dynamics given by the Koopman operator (see Section 2.4) has probabilities that are exactly equal to the joint probabilities of (Ω, Σ, μ, f). Fix a DS (Ω, Σ, f, μ) and let T_f be the associated Koopman operator given in Equation (2.12). To each partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ we can associate a partition of unity given by the collection of characteristic functions $\gamma = \{\mathbb{1}_C\}_{C \in \mathcal{C}} \subseteq L^\infty(\Omega)$.

To the DS (Ω, Σ, μ, f), we associate the QDS $(L^\infty(\Omega), T_f, \mu)$, where μ is a state in the sense of Example 3.2.3 which is given more explicitly in Remark 2.4.1.

Proposition 4.2.1. *For any stationary DS* (Ω, Σ, μ, f) *and partition* $\mathcal{C} = \{C_i\}_{i=1}^d \in \mathcal{P}_{ar}(\Omega)$, *the joint probabilities,* $\mu_n^{(f,\mathcal{C})}$, *given in Equation* (2.9) *are equal to the diagonal entries of the joint densities,* $\mu^{(n)}[\gamma]$, *of the associated QDS* $(L^\infty(\Omega), T_f, \mu)$ *given in Equation* (4.12). *Moreover, the off-diagonal entries of* $\mu^{(n)}[\gamma]$ *are equal to 0 for all* $n \in \mathbb{N}$. *Therefore*

$$\mu^{(n)}[\gamma] = \mathrm{diag}(\mu_n^{(f,\mathcal{C})})$$

for each $n \in \mathbb{N}$, *where* $\mathrm{diag}(\nu) \in M_d$ *is the diagonal matrix with entries from* ν *for any probability measure on* $(\{1, \ldots, d\}, \mathcal{P}(\{1, \ldots, d\}))$.

Proof. Fix a DS (Ω, Σ, μ, f) and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$. Let $(L^\infty(\Omega), T_f, \mu)$ be the associated QDS and $\gamma = (\mathbb{1}_C)_{C \in \mathcal{C}}$ be the corresponding partition of unity in $L^\infty(\Omega)$ as described in the preceding paragraph. Note that by Property (i) given immediately after the definition of T_f in Equation (2.12), the joint densities $\mu^{(n)}[\gamma]$ are indeed given by Equation (4.12) for each $n \in \mathbb{N}$. Then

$$\begin{aligned}
\mu^{(n)}[\gamma]_{\bar{i},\bar{j}} &= \mu(\gamma_{j_1}^* T_f(\gamma_{j_2})^* \cdots T_f^{n-1}(\gamma_{j_n}^*) T_f^{n-1}(\gamma_{i_n}) \cdots T_f(\gamma_{i_2})\gamma_{i_1}) \text{ Equation (4.12)} \\
&= \int_\Omega \prod_{k=0}^{n-1} T_f^k(\mathbb{1}_{C_{i_{k+1}}} \mathbb{1}_{C_{j_{k+1}}}) d\mu \quad \text{Remark 2.4.1} \\
&= \delta_{\bar{i},\bar{j}} \int_\Omega \mathbb{1}_{C_{i_1}} \cdot f(\mathbb{1}_{C_{i_2}}) \cdot \cdots \cdot f^{n-1}(\mathbb{1}_{C_{i_n}}) d\mu \quad \text{Equation (2.12)} \\
&= \delta_{\bar{i},\bar{j}} \int_{C_{i_1}} f(\mathbb{1}_{C_{i_2}}) \cdot \cdots \cdot f^{n-1}(1_{C_{i_n}}) d\mu \\
&\;\;\vdots \\
&= \delta_{\bar{i},\bar{j}} \int_{\cap_{k=0}^{n-1} f^{-k}(C_{i_{k+1}})} \mathbb{1}_\Omega d\mu \\
&= \delta_{\bar{i},\bar{j}} \mu(\cap_{k=0}^{n-1} f^{-k}(C_{i_{k+1}})) = \delta_{\bar{i},\bar{j}} \mu_n^{(f,\mathcal{C})}(C_{i_1}, \ldots, C_{i_n}), \quad \text{Equation (2.9)}
\end{aligned}$$

as desired, where equality 6 holds since $\int_A f(\mathbb{1}_B) d\mu = \int_{A \cap f^{-1}(B)} d\mu$ for any $A, B \in \Sigma$. Therefore $\mu^{(n)}[\gamma] = \mathrm{diag}(\mu_n^{(f,\mathcal{C})})$ for each $n \in \mathbb{N}$. $\square$

In this section we recall the definition of quantum Markov chains (QMCs). The QMC approach to non-commutative dynamics was first introduced by Accardi in [1], was developed further for the Accardi-Ohya-Watanabe (AOW) entropy in [3], and can be thought of as a symbolic dynamics for QDSs which utilizes spin chains from quantum statistical mechanics. Another QMC approach was introduced by Tuyls in [61] for the study of the Alicki-Fannes (AF) entropy, which was introduced in [6]. Finally, a generalization of both QMC approaches was given in [38], where the authors introduced the Kossakowski-Ohya-Watanabe (KOW) entropy. Throughout this section, we will follow mainly the terminology and notations of [3] and [38], but we will follow the construction given in [61], which is most suitable for our purposes.

Fix a stationary QDS $(\mathcal{A}, \Theta, \rho)$. We will refer to any completely positive, unital map $\mathcal{E} : M_d \otimes \mathcal{A} \to \mathcal{A}$ as a **transition expectation**, for any $d \in \mathbb{N}$. Let $\gamma = (\gamma_i)_{i=1}^d$ be an operational partition of unity. Following [61, Page 413] (see also [38, Equation 3.14]), we will consider the transition expectation $\mathcal{E}_\gamma : M_d \otimes \mathcal{A} \to \mathcal{A}$ given by the equation

$$\mathcal{E}_\gamma([a_{i,j}]) = \sum_{i,j=1}^{d} \gamma_i^* a_{i,j} \gamma_j \quad \text{for all } [a_{i,j}] \in M_d \otimes \mathcal{A}. \tag{4.14}$$

Remark 4.3.1. *In the original paper, [3], the authors only considered γ that were partitions of unity and considered a simplified (as compared to Equation* (4.14)*) transition expectation given by*

$$E_\gamma([a_{i,j}]) = \sum_{i=1}^{d} \gamma_i a_{i,i} \gamma_i \quad \textit{for all } [a_{i,j}] \in M_d \otimes \mathcal{A}.$$

Notice that there is no need for a γ_i^* in the above equation since $\gamma_i^* = \gamma_i$, for each i, whenever γ is a partition of unity. We will consider this transition expectation further in Section 7.2.

To include the dynamics of the QDS, we will also make use of the transition expectation

$$\mathcal{E}_{\gamma,\Theta} = \Theta \circ \mathcal{E}_{\gamma}. \tag{4.15}$$

A **quantum Markov chain** (QMC) is a pair $\{\rho, \mathcal{E}\}$ where ρ is an initial state and $\mathcal{E}$ is a transition expectation. We will be specifically interested in QMCs whose transition expectation is given by Equation (4.15). Given a QMC, we define the **quantum Markov state** $\psi \in \Sigma(M_d^{\otimes \mathbb{N}})$ by the equation

$$\psi(a_1 \otimes \cdots \otimes a_n) = \mathrm{tr}(\rho \mathcal{E}(a_1 \otimes \mathcal{E}(a_2 \otimes \mathcal{E}(\cdots \mathcal{E}(a_n \otimes \mathbb{1}) \cdots)))), \tag{4.16}$$

for all $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in M_d$. For notational convenience, we will often write $\psi = \{\rho, \mathcal{E}\}$ whenever ψ is the quantum Markov state obtained from the QMC $\{\rho, \mathcal{E}\}$ as defined in Equation (4.16).

The **joint densities for $\boldsymbol{\psi}$** are given by the density matrices $\rho_n \in M_d^{\otimes n}$ satisfying

$$\psi(a_1 \otimes \cdots \otimes a_n) = \mathrm{tr}(\rho_n a_1 \otimes \cdots \otimes a_n), \tag{4.17}$$

for all $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in M_d$. For any stationary QDS $(\mathcal{A}, \Theta, \rho)$, operational partition of unity γ, and associated quantum Markov chain and state $\{\rho, \mathcal{E}_{\gamma,\Theta}\}$ and ψ, respectively, the joint densities of ψ given in Equation (4.17) are equal to the joint densities are equal to the joint densities of $(\mathcal{A}, \Theta, \rho)$ with respect to γ given in Equation (4.10), as we will see in Section 7.

In a similar vein to the coupled classical system in Section 2.4, we can think of the coupled system $M_d \otimes \mathcal{A}$ here. In this case we only have access to the measurements with values in M_d; notice that when defining the joint correlations in Equation (4.16) we have assumed the state $\mathbb{1}$ on $\mathcal{A}$. We can think of the output of the transition expectation $\mathcal{E}_{\gamma}(a \otimes \mathbb{1})$ as the most likely state of $\mathcal{A}$ to have produced the measurement outcome $a \in M_d$ with respect to the partition γ. In practice, we will usually apply the lifting $\mathcal{E}^*_{\gamma,\Theta} : \Sigma(\mathcal{A}) \to \Sigma(M_d \otimes \mathcal{A})$, in the sense of [2], to the initial state ρ iteratively to obtain the joint densities ρ_n. The iterative applications of the lifting $\mathcal{E}^*_{\gamma,\Theta}$ can be thought of in the Schrödinger Picture as providing, in the limit, the state ψ which contains all the correlations of a classical stochastic process!

We will finish this section by giving the QMC representation for a classical dynamical system. Fix a DS (Ω, Σ, μ, f), a finite partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ of size d and let $(L^\infty(\Omega), T_f, \mu)$ and γ be the associated QDS and partition of unity, respectively. For each $k \in \{1, \ldots, d\}$, let $e_k = |k\rangle\langle k|$ in $L^\infty(\{1, \ldots, d\})$, where we identify $L^\infty(\{1, \ldots, d\})$ with the diagonal matrices in M_d, which we denote by $\mathrm{diag}(M_d)$.

The transition expectation $\mathcal{E}_\gamma : L^\infty(\{1, \ldots, d\}) \otimes L^\infty(\Omega) \to L^\infty(\Omega)$ for γ, given in Equation (4.14), simplifies to

$$\mathcal{E}_\gamma(\sum_{k=1}^{d} e_k \otimes f_k) = \sum_{k=1}^{d} \mathbb{1}_{C_k} \cdot f_k \quad \text{for any } f_1, \ldots, f_d \in L^\infty(\Omega). \tag{4.18}$$

Notice that since we have identified $L^\infty(\{1, \ldots, d\})$ with $\mathrm{diag}(M_d)$, there are no off-diagonal entries to consider and Equation (4.18) is of the form introduced in Remark 4.3.1.

The QMC representing the DS (Ω, Σ, μ, f) with respect to $\mathcal{C}$ is then given by the pair $\{\mu, \mathcal{E}_\gamma\}$ on the spin chain $\mathrm{diag}(M_d)^{\otimes \mathbb{N}}$ with quantum Markov state given by Equation (4.16). Recall that, in the symbolic dynamics picture for a classical DS, we define a measure $\hat{\mu}$ on $\Omega^* = \oplus_{\mathbb{N}} \Omega$ (see Remark 2.3.3) by

$$\hat{\mu}(C \left(\begin{smallmatrix} A_1 & \cdots & A_n \\ 1 & \cdots & n \end{smallmatrix} \right)) = \mu(\cap_{k=1}^{n} f^{-(k-1)}(\mathcal{C})), \tag{4.19}$$

for any cylinder set in Σ^*, and extend uniquely to Σ^*. Again, we will denote $\hat{\mu}(C \left(\begin{smallmatrix} A_1 & \cdots & A_n \\ 1 & \cdots & n \end{smallmatrix} \right))$ by $\mu^{(s,\hat{\mathcal{C}})}(A_1, \ldots, A_n)$ as in Equation (2.9). On the other hand, the quantum Markov state ψ plays the role of $\mu^{(s,\hat{\mathcal{C}})}$ in the spin chain. Identifying each A_k in Equation (4.19) with its representation in $\mathrm{diag}(M_d)$; i.e. $A_k = \sum_{j: C_j \subseteq A_k} e_j$, we have

$$\begin{aligned} \psi(A_1 \otimes \cdots \otimes A_n) &= \mu(\mathcal{E}_{\gamma,T_f}(A_1 \otimes \mathcal{E}_{\gamma,T_f}(A_2 \otimes \mathcal{E}_{\gamma,T_f}(\cdots \mathcal{E}_{\gamma,T_f}(A_n \otimes \mathbb{1}) \cdots)))) \quad (4.20) \\ &= \mu(\mathcal{E}_{\gamma,T_f}(A_1 \otimes \mathcal{E}_{\gamma,T_f}(A_2 \otimes \mathcal{E}_{\gamma,T_f}(\cdots \mathcal{E}_{\gamma,T_f}(A_{n-1} \otimes (f \circ \mathbb{1}_{A_n}))))) \end{aligned}$$

$$= \mu(f \circ \mathbb{1}_{A_1} \circ f^2 \cap \cdots \cap f^n \circ \mathbb{1}_{A_n})$$
$$= \mu(\cap_{k=1}^n f^{-k}(A_k)).$$

Therefore, if f is μ-invariant; i.e. (Ω, Σ, μ, f) is stationary, we have that

$$\psi(A_1 \otimes \cdots \otimes A_n) = \mu^{(s,\hat{\mathcal{C}})}(A_1, \ldots, A_n),$$

where $\mu^{(s,\hat{\mathcal{C}})}(A_1, \ldots, A_n)$ is the notation given just beneath Equation (4.19). Hence ψ can only take values depending on the partition γ (or $\mathcal{C}$), similar to what we have seen before with classical symbolic dynamics.

Chapter 5

Entropy in Classical Systems

Dynamical entropy in classical systems can be seen from two distinct viewpoints: The information theoretic viewpoint (see e.g. [19]) which uses entropy rate of stochastic processes and the dynamical systems viewpoint (see e.g. [22]) which uses the Kolmogorov-Sinai (KS) dynamical entropy. We prove that the connection between entropy rate and KS entropy is seen through the symbolic dynamics of a stochastic process, which is a dynamical system with KS entropy equal to the entropy rate of the original stochastic process (see Section 5.3). On the other hand, entropy rate and KS entropy are inherently different as the former is probabilistic in nature and the latter is deterministic (see Section 5.4).

5.1 Entropy in Dynamical Systems

Let (Ω, Σ, μ) be a probability space and $\mathcal{P}_{ar}(\Omega)$ be the collection of all finite or countably infinite measurable partitions of Ω as in Section 2.2. Given any partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ we define the **entropy of $\mathcal{C}$** by

$$H_\mu(\mathcal{C}) := \sum_{C \in \mathcal{C}} \eta(\mu(C)),$$

where $\eta : [0, \infty) \to [0, \infty)$ is given by $\eta(x) = -x \ln x$, for $x > 0$ and we agree that $\eta(0) = 0$. When there is no confusion about the probability measure in question, we will simply write $H(\mathcal{C})$ instead of $H_\mu(\mathcal{C})$.

Remark 5.1.1 ([22, Page 23]). *It is well known that η is countably subadditive; i.e. $\eta(\sum_n a_n) \le \sum_n \eta(a_n)$ for any nonnegative sequence $\{a_n\}_n$. This gives that for any* *probability space* (Ω, Σ, μ) *and any two partitions* $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$ *satisfying* $\mathcal{D} \leq \mathcal{C}$*, we have that* $H(\mathcal{D}) \leq H(\mathcal{C})$*.*

Fix a probability space (Ω, Σ, μ). Recall that, for any two sets $C, D \in \Sigma$, conditional probability of C given D is given by $\mu(C|D) := \mu(C \cap D)/\mu(D)$. Given two partitions, $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$, the **conditional entropy of $\mathcal{C}$ given $\mathcal{D}$** is given by

$$H(\mathcal{C}|\mathcal{D}) := \sum_{D \in \mathcal{D}} \mu(D) \sum_{C \in \mathcal{C}} \eta(\mu(C|D)) = -\sum_{\substack{C \in \mathcal{C} \\ D \in \mathcal{D}}} \mu(C \cap D) \ln(\mu(C|D)). \tag{5.1}$$

The so-called chain rule follows.

Theorem 5.1.2 (Chain Rule, [22, Equation 1.4.3])**.** *Let* (Ω, Σ, μ) *be a probability space and* $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$*. Then*

$$H(\mathcal{C} \vee \mathcal{D}) = H(\mathcal{D}) + H(\mathcal{C}|\mathcal{D}).$$

More generally, given a finite collection of partitions $\mathcal{C}_0, \ldots, \mathcal{C}_n \in \mathcal{P}_{ar}(\Omega)$*, we have*

$$H(\vee_{k=0}^{n} \mathcal{C}_k) = H(\mathcal{C}_0) + \sum_{k=1}^{n} H(\mathcal{C}_k | \vee_{\ell=0}^{k-1} \mathcal{C}_\ell).$$

Also, from the definition of conditional entropy (Equation (5.1)) and the countable subadditivity of η (Remark 5.1.1) we have, for any $\mathcal{B}, \mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$ satisfying $\mathcal{B} \leq \mathcal{D}$, that

$$0 \leq H(\mathcal{C}|\mathcal{D}) \leq H(\mathcal{C}|\mathcal{B}). \tag{5.2}$$

See [22, Section 1.4] for more details on conditional entropy of partitions. The following theorem will be used throughout this manuscript. In the proof we will use the well known Césaro mean Theorem which states that, for any sequence of real numbers $\{a_n\}_{n=1}^{\infty}$ converging to some element, $a \in \mathbb{R} \cup \{\infty\}$, the sequence $\{b_n\}_{n=1}^{\infty}$ given by $b_n = \frac{1}{n} \sum_{k=1}^{n} a_k$, for each $n \in \mathbb{N}$, also converges to a.

Theorem 5.1.3. *Let* (Ω, Σ, μ) *be a probability space and* $\{\mathcal{C}_n\}_{n=1}^{\infty}$ *be a sequence in* $\mathcal{P}_{ar}(\Omega)$*. If* $\lim_{n\to\infty} H(\mathcal{C}_n | \vee_{k=1}^{n-1} \mathcal{C}_k) = a$*, then* $\lim_{n\to\infty} \frac{1}{n} H(\vee_{k=1}^{n} \mathcal{C}_k) = a$*.*

Proof. Set $a_1 = H(\mathcal{C}_1)$ and $a_n = H(\mathcal{C}_n | \vee_{k=1}^{n-1} \mathcal{C}_k)$, for each $n \in \mathbb{N}$ with $n \geq 2$. Then, by assumption, a_n converges to $a \in \mathbb{R} \cup \{\infty\}$. For each $n \in \mathbb{N}$, set $b_n = \frac{1}{n}\sum_{k=1}^{n} a_k$. Then, by Theorem 5.1.2, $b_n = \frac{1}{n}H(\vee_{k=1}^{n}\mathcal{C}_k)$ which converges to a by the Césaro mean Theorem. □

Next we wish to define the Kolmogorov-Sinai (KS) dynamical entropy. Fix a DS (Ω, Σ, μ, f) and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$. The **Kolmogorov-Sinai (KS) entropy of (Ω, Σ, μ, f) with respect to $\mathcal{C}$** is given by

$$h^{KS}(f, \mathcal{C}) = \lim_{n\to\infty} \frac{1}{n} H(\vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})), \tag{5.3}$$

whenever this limit exists.

From Theorem 5.1.3, we have

$$h^{KS}(f, \mathcal{C}) = \lim_{n\to\infty} H(f^{-n}(\mathcal{C}) | \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})), \tag{5.4}$$

whenever this limit exists.

Corollary 5.1.4. *Let (Ω, Σ, μ, f) be a stationary DS and $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$. Then the limit in Equation* (5.4)*, and hence the limit in Equation* (5.3)*, exists and*

$$h^{KS}(f, \mathcal{C}) = \lim_{n\to\infty} H(f^{-n}(\mathcal{C}) | \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})).$$

Proof. For each $n \in \mathbb{N}$ with $n \geq 2$, we have

$$\begin{aligned} H(f^{-n}(\mathcal{C}) | \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})) &\leq H(f^{-n}(\mathcal{C}) | \vee_{k=2}^{n} f^{-(k-1)}(\mathcal{C})) \quad \text{by (5.2)} \\ &= H(f^{-(n-1)}(\mathcal{C}) | \vee_{k=1}^{n-1} f^{-(k-1)}(\mathcal{C})), \end{aligned}$$

where the last equality holds since (Ω, Σ, μ, f) is stationary. Therefore $H(f^{-n}(\mathcal{C})| \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C}))$ is a decreasing sequence which is bounded below by zero and hence converges. By Theorem 5.1.3,

$$h^{KS}(f, \mathcal{C}) = \lim_{n\to\infty} H(f^{-n}(\mathcal{C}) | \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})).$$

□

Remark 5.1.5. *As mentioned before it is common in the literature to only refer to* (Ω, Σ, μ, f) *as a DS whenever* μ *is* f*-invariant. Although this convention has its benefits, as evidenced by Corollary 5.1.4, we find it restrictive and do not adopt it here.*

Finally, the **KS entropy of (Ω, Σ, μ, f)** is given by

$$h^{KS}(f) = \sup_{\substack{\mathcal{C} \in \mathcal{P}_{ar}(\Omega) \\ H(\mathcal{C}) < \infty}} h^{KS}(f, \mathcal{C}). \tag{5.5}$$

Remark 5.1.6. *Fix a dynamical system* (Ω, Σ, μ, f)*. In many instances, KS entropy is taken as the* sup *over only finite partitions. However, the two definitions are equivalent (see [22, Page 102]). Furthermore, it is remarked in [22, Page 61] that the restriction of the* sup *in Equation* (5.5) *to include only those partitions,* $\mathcal{C}$*, satisfying* $H(\mathcal{C}) < \infty$ *is natural because otherwise it is possible to obtain infinite KS entropy for the identity transformation. This is due to the fact that* $H(f, \mathcal{C}) = \infty$ *whenever* $H(\mathcal{C}) = \infty$*.*

For a more detailed exposition on dynamical entropy and classical dynamical systems (with invariant measures), we refer the reader to the book of Walters [64]. For extensions of the results of Walters to include infinite partitions with finite entropy, the reader is referred to the book of Downarowicz [22].

5.2 Entropy in Probability Theory

Next we look at entropy of random variables and stochastic processes. We will stick to to discrete output spaces, although more general definitions are well known. Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ be a discrete measurable space and X an (Ω, E) random variable with pmf $p = p_X$. The **entropy X** is given by the equation

$$H_\mu(X) := \sum_{x \in E} \eta(p(x)). \tag{5.6}$$

When there is no confusion about the probability measure in question, we will simply write $H(X)$ instead of $H_\mu(X)$. Note that, by Equation (2.6), $H(X^{\mathcal{C}}) = H(\mathcal{C})$ for any partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ on a probability space (Ω, Σ, μ), where $X^{\mathcal{C}}$ is given just before Equation (2.6).

Any finite collection, $(X_k)_{k=1}^n$, of (Ω, E) discrete random variables can also be viewed as a discrete (Ω, E^n) random variable (or random vector) and the entropy of $(X_1, \dots, X_n)$ is given by Equation (5.6) and is related to its joint pmf by the equation

$$H(X_1, \dots, X_n) := \sum_{\substack{x_k \in E \\ 1 \le k \le n}} \eta(p_{X_1,\dots,X_n}(x_1, \dots, x_n)). \tag{5.7}$$

Notice that, by Equation (2.7), for any finite collection of partitions $(\mathcal{C}_k)_{k=1}^n$ in $\mathcal{P}_{ar}(\Omega)$, where (Ω, Σ, μ) is a probability space, we have $H(X^{\mathcal{C}_1}, \dots, X^{\mathcal{C}_n}) = H(\vee_{k=1}^n \mathcal{C}_k)$.

The **conditional entropy of $\boldsymbol{X_{n+1}}$ given $\boldsymbol{(X_1, \dots, X_n)}$** is given by the equation

$$H(X_{n+1}|X_1, \dots, X_n) := \sum_{\substack{x_k \in E \\ 1 \le k \le n}} p(x_1, \dots, x_n) \sum_{x_{n+1} \in E} \eta(p(x_{n+1}|x_1, \dots, x_n)), \tag{5.8}$$

where $p(x_{n+1}|x_1, \dots, x_n)$ is given in Equation (2.2). Notice that, by Equation (2.8), for any finite collection of partitions $(\mathcal{C}_k)_{k=1}^n \in \mathcal{P}_{ar}(\Omega)$ of a probability space (Ω, Σ, μ), we have $H(X^{\mathcal{C}_n}|X^{\mathcal{C}_1}, \dots, X^{\mathcal{C}_{n-1}}) = H(\mathcal{C}_n| \vee_{k=1}^{n-1} \mathcal{C}_k)$. See [19, Section 2.2] for more details on conditional entropy of random variables.

Remark 5.2.1. *The entropy and conditional entropy of non-discrete random variables can be defined similarly to Equations* (5.6) *and* (5.8), *respectively, by using integration and probability distribution functions instead of sums and pmfs. However, we are mainly interested in discrete random variables here.*

Next we turn to entropy rate of stochastic processes. Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ be a discrete measurable space, and $\mathbf{X} = (X_n)_{n=1}^\infty$ be an (Ω, E) stochastic process. Given a stochastic process of discrete random variables, the entropy of a finite initial subsequence is given by Equation (5.7) and the conditional entropy of the n^{th} term given all the previous ones is given by Equation (5.8).

The **entropy rate of a stochastic process** $\mathbf{X} = (X_n)_{n=1}^{\infty}$ is given by

$$H(\mathbf{X}) := \lim_{n\to\infty} \frac{1}{n} H(X_1, \ldots, X_n), \tag{5.9}$$

whenever this limit exists.

Another quantity, which is often equal to the entropy rate, is given by

$$H'(\mathbf{X}) := \lim_{n\to\infty} H(X_{n+1}|X_1, \ldots, X_n), \tag{5.10}$$

whenever this limit exists. The two quantities $H(\mathbf{X})$ and $H'(\mathbf{X})$ correspond to two different interpretations of entropy rate. The first is interpreted as the average entropy of the first n random variables and the second as the entropy of the last random variable given the past. The following result shows the relationship between $H(\mathbf{X})$ and $H'(\mathbf{X})$.

Corollary 5.2.2. *Let $\mathbf{X} = (X_n)_{n=1}^{\infty}$ be a stochastic process. If the limit in Equation (5.10) exists, then the limit in Equation (5.9) also exists and $H(\mathbf{X}) = H'(\mathbf{X})$.*

Proof. By the definitions of $H(X_1, \ldots, X_n)$ and $H(X_n|X_1, \ldots, X_{n-1})$ in Equations (5.7) and (5.8), respectively, (and the sentence following each of those equations) this is simply a restatement of Theorem 5.1.3. $\square$

The following is another corollary for stationary stochastic processes which is also proved in [19, Theorem 4.2.2].

Corollary 5.2.3. *Let $\mathbf{X}$ be a stationary stochastic process. Then the limits in Equations (5.9) and (5.10) both exist and $H(\mathbf{X}) = H'(\mathbf{X})$.*

Proof. The proof is similar to the proof of Corollary 5.1.4. For each $n \in \mathbb{N}$ with $n \geq 2$, we have

$$H(X_n|X_1, \ldots, X_{n-1}) \leq H(X_n|X_1, \ldots, X_{n-1}) \quad \text{by (5.2)}$$

$$= \quad H(X_{n-1}|X_1, \ldots, X_{n-2}) \quad \text{since } \mathbf{X} \text{ is stationary.}$$

Therefore $H(X_n|X_1, \ldots, X_{n-1})$ is a decreasing sequence which is bounded below by zero and hence converges. By Equations (5.7) and (5.8), and Theorem 5.1.3, $H(\mathbf{X}) = H'(\mathbf{X})$. $\square$

Next we look at the entropy rate of discrete Markov processes governed by a transition matrix; i.e. the discrete Markov processes satisfying Equation (2.4). Recall the conventions for representing transition probabilities as stochastic matrices, probability measures on a discrete measurable space $(E, \mathcal{P}(E))$ as probability vectors, and defining their product by matrix multiplication at the end of Section 2.1. Also recall that μ is P-invariant whenever $P\mu = \mu$ and the convention of setting the initial measure μ to be p_{X_1}. The following theorem gives a simplification of the entropy rate for Markov processes governed by a transition matrix.

Theorem 5.2.4. *Let $\mathbf{X}$ be a discrete (Ω, E) Markov process governed by the transition matrix P and set $\mu = p_{X_1}$. Then*

$$H(\mathbf{X}) = \lim_{n\to\infty} \sum_{y\in E} (P^n\mu)_y \sum_{x\in E} \eta(p_{x,y}),$$

whenever the limit exists. Moreover, if $\mathbf{X}$ is stationary, then

$$H(\mathbf{X}) = \sum_{y\in E} \mu_y \sum_{x\in E} \eta(p_{x,y}).$$

Proof. Since $\mathbf{X}$ is a Markov process governed by the transition matrix P we have

$$H(X_{n+1}|X_1, \ldots, X_n) = H(X_{n+1}|X_n) = \sum_{y\in E} p_{X_n}(y) \sum_{x\in E} \eta(p_{x,y}),$$

for each $n \in \mathbb{N}$, where the second equality follows from Equation (5.8). Then from the definition of matrix multiplication, for each $n \in \mathbb{N}$ and $e_1, \ldots, e_n \in E$, we have

$$p_{X_n}(e_n) = \sum_{\substack{e_k\in E\\ 1\le k\le n-1}} p_{X_1}(e_1) \prod_{k=2}^{n} p_{e_k,e_{k-1}} = (P^n\mu)_{e_n}.$$

Therefore

$$H(\mathbf{X}) = \lim_{n\to\infty} \sum_{y\in E} (P^n\mu)_y \sum_{x\in E} \eta(p_{x,y}),$$

whenever the limit exists. The moreover statement is immediate because $P^n\mu = \mu$ for all $n \in \mathbb{N}$ whenever $\mathbf{X}$ is stationary. $\square$

Remark 5.2.5. *Certain cases of Theorem 5.2.4 appear frequently in the literature, but to the best of our knowledge we have not seen it presented in the generality of above. For instance, it can be seen for the case where μ is P-invariant in [19, Theorem 4.2.4] or [64, Theorem 4.26].*

In the literature, given a transition matrix P on a discrete measurable space $(E, \mathcal{P}(E))$ with a unique invariant probability vector μ, it is common to set

$$H(P) := \sum_{y\in E} \mu_y \sum_{x\in E} \eta(p_{x,y}), \tag{5.11}$$

and refer to $H(P)$ as the **entropy of $\boldsymbol{P}$.** As it is shown in Theorem 5.2.4, the entropy, $H(P)$, of P is equal to the entropy rate, $H(\mathbf{X})$, of any stationary Markov process, $\mathbf{X}$, governed by the transition matrix P such that $p_{X_1} = \mu$.

5.3 Entropy in Symbolic Dynamics: The connection between entropy rate and KS entropy

Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ a (not necessarily discrete) measurable space, $\mathbf{X} = (X_n)_{n=0}^\infty$ an (Ω, E) stochastic process, and $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$ the corresponding symbolic dynamics. Recall that, for any $\mathcal{C} \in \mathcal{P}_{ar}(E)$, we defined the partition $\hat{\mathcal{C}} \in \mathcal{P}_{ar}(E^*)$, by

$$\hat{\mathcal{C}} := \{C\left(\begin{smallmatrix} A \\ 1 \end{smallmatrix}\right)\}_{A\in\mathcal{C}}$$

and the set

$$\hat{\mathcal{P}_{ar}}(E) := \{\hat{\mathcal{C}} : \mathcal{C} \in \mathcal{P}_{ar}(E)\} \subset \mathcal{P}_{ar}(E^*).$$

Since $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$ is a DS, its partition dependent and independent KS entropies are given by Equations (5.3) and (5.5), respectively.

Of particular interest is the KS entropy of $(E^*, \mathcal{E}^*, \mu^{\mathbf{X}}, s)$ with respect to the partitions in $\hat{\mathcal{P}_{ar}}(E)$. For each $\mathcal{C} \in \mathcal{P}_{ar}(E)$, recall the $(\Omega, \mathcal{C})$ stochastic process $\mathbf{X}_{\mathcal{C}} = (X_n^{\mathcal{C}})_{n=0}^{\infty}$ defined in the paragraph preceding Proposition 2.3.2. Since the values of $\mathbf{X}_{\mathcal{A}}$ are singletons, it is clear that $\mathbf{X}$ can be identified with $\mathbf{X}_{\mathcal{A}}$ whenever $\mathcal{A}$ is the atomic partition of the discrete space E. The following corollary shows that the KS entropy of $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$ with respect to $\hat{\mathcal{C}}$ and the entropy rate of $\mathbf{X}_{\mathcal{C}}$ are equal.

Corollary 5.3.1. *Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ a (not necessarily discrete) measurable space, $\mathbf{X}$ an (Ω, E) stochastic process and $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$ the symbolic dynamics of $\mathbf{X}$. Then for each $\mathcal{C} \in \mathcal{P}_{ar}(E)$, $H(\mathbf{X}_{\mathcal{C}}) = h^{KS}(s, \hat{\mathcal{C}})$. In particular, whenever E is a discrete space, $H(\mathbf{X}) = h^{KS}(s, \hat{\mathcal{A}})$, where $\mathcal{A}$ is the atomic partition of E.*

Proof. This is an immediate consequence of Proposition 2.3.2 and the definitions of entropy introduced in the two preceding sections. □

An important tool for computing the KS entropy of a DS is the Kolmogorov-Sinai Theorem. First, given a DS (Ω, Σ, μ, f) and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, we say that **$\mathcal{C}$ is a generating partition for (Ω, Σ, μ, f)** if

$$\sigma(\cup_{n=1}^{\infty} \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})) = \Sigma.$$

Notice that the definition of a generating partition does not depend on μ, but for simplicity of notation we keep the full DS.

Theorem 5.3.2 (Kolmogorov-Sinai Theorem)**.** *Let (Ω, Σ, μ, f) be a DS and $\mathcal{C}, \mathcal{D} \in \mathcal{P}_{ar}(\Omega)$. If $\sigma(\mathcal{D}) \subseteq \sigma(\cup_{n=1}^{\infty} \vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C}))$, then*

$$h^{KS}(f, \mathcal{C}) \geq h^{KS}(f, \mathcal{D}).$$

In particular, if $\mathcal{C}$ is a generating partition and $H(\mathcal{C}) < \infty$ then $h^{KS}(f) = h^{KS}(f, \mathcal{C})$.

A proof of Theorem 5.3.2 can be found in [22, Theorem 4.2.2].

Corollary 5.3.3. *Let (Ω, Σ, μ) be a probability space, $(E, \mathcal{E})$ be a discrete measurable space, $\mathcal{A}$ be the atomic partition of E, $\mathbf{X}$ be an (Ω, E) stochastic process and $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$ be the symbolic dynamics of $\mathbf{X}$. Then $H(\mathbf{X}) = h^{KS}(s) = h^{KS}(s, \hat{\mathcal{A}})$ whenever X_1 has finite entropy.*

Proof. Since

$$\vee_{k=1}^{n-1} s^{-(k-1)}(\hat{\mathcal{A}}) = \{C\left(\begin{smallmatrix}\{e_1\} & \cdots & \{e_n\} \\ 1 & \cdots & n\end{smallmatrix}\right) : e_1, \ldots, e_n \in E\},$$

we see that $\hat{\mathcal{A}}$ is a generating partition for $(E^*, \Sigma^*, \mu^{\mathbf{X}}, s)$. Then Proposition 5.3.1 and Theorem 5.3.2 give that $h^{KS}(s) = h^{KS}(s, \hat{\mathcal{A}}) = H(\mathbf{X})$, whenever $\hat{\mathcal{A}}$ has finite entropy. Noticing that $H(\hat{\mathcal{A}}) = H(X^{\hat{\mathcal{A}}}) = H(X_1) < \infty$, the result follows. $\square$

Remark 5.3.4. *Notice that the results in Proposition 5.3.1 and Corollary 5.3.3 look nearly identical except that the condition $H(X_1) < \infty$ has been added to the latter. This assumption is necessary due to the fact that $h^{KS}(s, \hat{\mathcal{A}})$ is defined regardless of whether $H(\hat{\mathcal{A}})$ is finite or infinite, but is only considered in the supremum of Equation* (5.5) *when $H(\hat{\mathcal{A}})$ is finite.*

5.4 Differences between entropy rate and KS entropy

In this section we give the differences between entropy rate and KS entropy. The first thing to notice is that dynamics of a stochastic process are probabilistic in nature, whereas the dynamics of a DS are deterministic in nature. This fact will be exploited to establish the differences in the two entropies in this section and again in Section 6.3 to establish differences between quantum dynamical entropy and KS entropy. The following two propositions give properties of KS entropy whose analogous statements do not hold true for entropy rate. The first proposition will use the well known fact (see e.g. [22, Equation (1.3.2)]) that for any probability space (Ω, Σ, μ) and partition

$\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, we have

$$H(\mathcal{C}) \leq \ln |\mathcal{C}| \leq \ln |\Omega|. \tag{5.12}$$

Proposition 5.4.1. *Let (Ω, Σ, μ, f) be a DS such that $|\Omega| < \infty$. Then $h^{KS}(f) = 0$.*

Proof. For any partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, Equation (5.12) gives that

$$H(\vee_{k=1}^{n} f^{-(k-1)}(\mathcal{C})) \leq \ln |\Omega|$$

, for each $n \in \mathbb{N}$. Therefore $h^{KS}(f, \mathcal{C}) = 0$ for every $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ and thus $h^{KS}(f) = 0$. □

Proposition 5.4.2 ([22, Fact 4.1.14])**.** *Let (Ω, Σ, μ, f) be a DS. Then the KS entropy of f is linear in time; i.e.*

$$h^{KS}(f^n) = n h^{KS}(f), \quad \textit{for all } n \in \mathbb{N}.$$

The example of a stationary Markov process governed by the unbiased random walk on a cycle (which is defined below) is enough to show that entropy rate does not have the analogous properties of KS entropy given in Propositions 5.4.1 and 5.4.2. Let $V = \{0, \ldots, N-1\}$, for some $N \in \mathbb{N}$ with $N \geq 3$, let μ be the uniform distribution on V and consider the discrete probability space $(V, \mathcal{P}(V), \mu)$. The unbiased random walk on the N-cycle, V, is governed by the transition matrix P with entries $p_{v+1,v} = p_{v-1,v} = 1/2$, where addition is done modulo N, for all $v \in V$, and $p_{u,v} = 0$ if $u \neq v \pm 1$.

Proposition 5.4.3. *Let $(V, \mathcal{P}(V), \mu)$ be the discrete probability space with $V = \{0, \ldots, N-1\}$, for some $N \in \mathbb{N}$ with N odd and $N \geq 3$, μ be the uniform distribution on V and P be the transition matrix governing the unbiased random walk on V. Then $H(P) = \ln 2$ and $H(P^2) = \frac{3}{2} \ln 2$.*

Proof. Clearly μ is the unique probability measure that is P-invariant. Therefore Equation (5.11) gives that

$$H(P) = \sum_{v \in V} \mu_v \sum_{u \in V} \eta(p_{u,v}) = \sum_{v \in V} \frac{1}{N} 2\eta(\frac{1}{2}) = \ln 2.$$

Also notice that, for all $v \in V$, P^2 has entries $p^{(2)}_{v\pm 2,v} = \frac{1}{4}$, $p^{(2)}_{v,v} = \frac{1}{2}$ and $p^{(2)}_{u,v} = 0$ in all other cases, where addition is done modulo N. Again μ is the unique probability measure that is P^2-invariant and Equation (5.11) gives that

$$H(P^2) = \sum_{v\in V} \mu_v \sum_{u\in V} \eta(p^{(2)}_{u,v}) = \sum_{v\in V} \frac{1}{N}(2\eta(\frac{1}{4}) + \eta(\frac{1}{2})) = \frac{3}{2}\ln 2.$$

□

Proposition 5.4.3 and Corollary 5.3.3 establish that the KS entropy of the symbolic dynamics of a stochastic process with range in a finite measurable space need not be zero, whereas Proposition 5.4.1 states that the KS entropy of a finite DS must be 0. Propositions 5.4.3 and 5.4.1 do not contradict Proposition 5.3.1 since the cardinality of E^*, in the symbolic dynamics of a stochastic process, is not finite unless the range, E, of the stochastic process is a singleton. Also, Proposition 5.4.3 says that entropy rate is not linear in time whereas Proposition 5.4.2 says that KS entropy is linear in time. Again these two propositions are not contradictory. We will elaborate a bit further for clarity. In what follows, we will denote the KS entropy of a DS (Ω, Σ, μ, f) by $h^{KS}(f, \mu)$ instead of $h^{KS}(f)$ to distinguish between different measures. We will denote the partition dependent KS entropy similarly.

Let $(V, \mathcal{P}(V), \mu)$ be the finite discrete probability space with $V = \{0, \ldots, N-1\}$, for some $N \in \mathbb{N}$ with N odd and $N \geq 3$, μ be the uniform distribution on V, P be the transition matrix governing the unbiased random walk on the N-cycle, V, $\mathbf{X} = (X_n)_{n=0}^{\infty}$ be any stationary Markov process governed by the transition matrix P, $\mathcal{A}_V$ the atomic partition of V and $(V^*, \mathcal{P}(V)^*, \mu^{\mathbf{X}}, s_1)$ be the symbolic dynamics of $\mathbf{X}$, where s_1 denotes the shift map on V^*. Since $\hat{\mathcal{A}}_V$ is a generating partition for $(V^*, \mathcal{P}(V)^*, \mu^{\mathbf{X}}, s_1)$, Corollary 5.3.3 shows that $H(\mathbf{X}) = h^{KS}(s_1, \mu^{\mathbf{X}}) = h^{KS}(s_1, \mu^{\mathbf{X}}, \hat{\mathcal{A}}_V)$. Since $\hat{\mathcal{A}}_V \vee s_1^{-1}(\hat{\mathcal{A}}_V)$ is a generating partition for $(V^*, \mathcal{P}(V)^*, \mu^{\mathbf{X}}, s_1^2)$ with finite entropy, the KS Theorem gives that $h^{KS}(s_1^2, \mu^{\mathbf{X}}) = h^{KS}(s_1^2, \mu^{\mathbf{X}}, \hat{\mathcal{A}}_V \vee s_1^{-1}(\hat{\mathcal{A}}_V))$. Next, consider the stationary Markov process $\mathbf{Y} = ((X_{2n-1}, X_{2n}))_{n=1}^{\infty}$ and let $((V \times V)^*, \mathcal{P}(V \times V)^*, \mu^{\mathbf{Y}}, s_2)$ be the symbolic dynamics of $\mathbf{Y}$ and $\mathcal{A}_{V\times V}$ be the atomic partition of $V \times V$, where s_2 denotes the shift map on $(V \times V)^*$. Since $\hat{\mathcal{A}_{V\times V}}$ is a generating partition for $((V \times V)^*, \mathcal{P}(V \times V)^*, \mu^{\mathbf{Y}}, s_2)$ with finite entropy, Corollary 5.3.3 gives that $H(\mathbf{Y}) = h^{KS}(s_2, \mu^{\mathbf{Y}}) = h^{KS}(s_2, \mu^{\mathbf{Y}}, \hat{\mathcal{A}_{V\times V}})$. Notice that

$$\mu^{\mathbf{Y}}\left(C\begin{pmatrix} (e_1,e_2) & \cdots & (e_{2n-1},e_{2n}) \\ 1 & \cdots & n \end{pmatrix}\right) = \mu^{\mathbf{X}}\left(C\begin{pmatrix} e_1 & \cdots & e_{2n} \\ 1 & \cdots & 2n \end{pmatrix}\right),$$

for all $e_1, \ldots, e_{2n} \in E$ and $n \in \mathbb{N}$. Thus

$$H_{\mu^{\mathbf{Y}}}(\vee_{k=1}^{n} s_2^{-(k-1)}(\hat{\mathcal{A}_{V\times V}})) = H_{\mu^{\mathbf{X}}}(\vee_{k=1}^{n} (s_1^2)^{-(k-1)}(\hat{\mathcal{A}}_V \vee s_1^{-1}(\hat{\mathcal{A}}_V)),$$

for all $n \in \mathbb{N}$, and therefore

$$H(\mathbf{Y}) = h^{KS}(s_2, \mu^{\mathbf{Y}}, \hat{\mathcal{A}_{V\times V}}) = h^{KS}(s_1^2, \mu^{\mathbf{X}}, \hat{\mathcal{A}}_V \vee s_1^{-1}(\hat{\mathcal{A}}_V)) = 2H(\mathbf{X}).$$

In other words, the KS entropy of $(V^*, \mathcal{P}(V)^*, \mu^{\mathbf{X}}, s_1^2)$ is equal to the KS entropy of $((V \times V)^*, \mathcal{P}(V \times V)^*, \mu^{\mathbf{Y}}, s_2)$ and corresponds to the entropy rate of $\mathbf{Y}$.

Next consider the stochastic process $\mathbf{Z} = (X_{2n-1})_{n=1}^{\infty}$ and let $(V^*, \mathcal{P}(V)^*, \mu^{\mathbf{Z}}, s_1)$ be the symbolic dynamics of $\mathbf{Z}$. Then $\mathbf{Z}$ is the stationary and invariant Markov process governed by the transition matrix P^2 and, from Proposition 5.4.3 and Corollary 5.3.3, $H(\mathbf{Z}) = h^{KS}(s_1, \mu^{\mathbf{Z}}) = h^{KS}(s_1, \mu^{\mathbf{Z}}, \hat{\mathcal{A}}_V) = \frac{3}{2}\ln 2$. Thus Propositions 5.4.2 and 5.4.3 are not contradictory as $2H(P) = 2H(\mathbf{X}) = h^{KS}(s_1^2, \mu^{\mathbf{X}})$ corresponds to the entropy rate of $\mathbf{Y}$, whereas $H(P^2) = H(\mathbf{Z}) = h^{KS}(s_1, \mu^{\mathbb{Z}})$ corresponds to the entropy rate of $\mathbf{Z}$.

5.5 Entropy in Data Compression

Next we revisit the entropy of random variables for the purpose of data compression. We will discuss extensions to quantum data compression in Subsection 7.3.1. In this section, all codings will be done into strings of bits. The extensions to d-bits can be done easily.

Let S be a finite or countable set equipped with the power set σ-algebra $\mathcal{P}(S)$, and let X be a random variable with values in S. The set S will be referred to as the **symbols set** that we wish to encode. In the literature, the set S is referred to as the set of object, the message set, or sometimes even the index set. For any set Y, we will denote by Y^+ the set $\cup_{\ell=0}^{\infty} Y^\ell$ which can be thought of as the collection of all possible strings from Y, where Y^0 denotes the empty set (or empty string). Lastly, let $A = \{0, 1\}$ be the binary alphabet. A **code**, also referred to as a source code, $C : S \to A^+$ is a mapping from S to A^+, strings with letters in the binary alphabet A. For each $x \in S$, we refer to $C(x)$ as the **codeword of the symbol $\boldsymbol{x}$**. We define $\ell : A^+ \to \mathbb{N}_0$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, by $\ell(a) = m$ whenever $a \in A^m$ and refer to $\ell(a)$ as the **length** of a.

The **expected length** of a code C on a symbol set S is given by

$$EL(C) := \sum_{x \in S} p(x)\ell(x) = \mathbb{E}[\ell(C(X))], \tag{5.13}$$

where p is the pmf of the random variable X and the expectation $\mathbb{E}$ is taken with respect to p.

A code C is said to be **non-singular** whenever

$$x \neq y \Rightarrow C(x) \neq C(y), \quad \text{for all } x, y \in S; \tag{5.14}$$

i.e. whenever C is an injective map and hence the codewords are pairwise distinguishable. We extend the code C to the **extended code**, also called the extension of C, $C^+ : S^+ \to A^+$ by concatenation. That is to say

$$C^+(x_1 x_2 \cdots x_n) = C(x_1)C(x_2)\cdots C(x_n), \tag{5.15}$$

for all $x_1 x_2 \cdots x_n \in S^n$ and $n \in \mathbb{N}$, and we define $C^+(\emptyset) = \emptyset$. We call the code C **uniquely decodable** whenever its extension C^+ is non-singular; i.e. C is uniquely decodable whenever all strings of symbols from S are pairwise distinguishable.

An extremely useful class of uniquely decodable codes are the so-called **instantaneous (or prefix-free)** codes. A code is said to be prefix-free if no codeword is the prefix of another; i.e. for every distinct pair $x, y \in S$ there is no $a \in A^+$ such that $C(x)a = C(y)$. Prefix-free codes are called instantaneous because the decoder is able to read out each codeword from a string of codewords, instantaneously, as soon as she sees that word appear in a string and without waiting for the entirety of the string.

The Kraft-McMillan Inequality is fundamental in classical data compression.

Theorem 5.5.1. *(Kraft-McMillan Inequality, [19, Theorems 5.2.1 and 5.5.1]) For any uniquely decodable code over a symbol set S with cardinality $|S| = m \in \mathbb{N}$, the codeword lengths $\ell_1, \ell_2, \ldots, \ell_m$ must satisfy the inequality*

$$\sum_{i=1}^{m} 2^{-\ell_i} \leq 1.$$

Conversely, given a set of codeword lengths that satisfies this inequality, there exists an instantaneous code with these code lengths.

Remark 5.5.2. *The Kraft-McMillan Inequality is sometimes referred to only as the Kraft Inequality. This is due to the fact that Kraft was the first to prove the inequality in [39], although his original theorem refers only to instantaneous codes. McMillan later extended Kraft's work to include all uniquely decodable codes in [42]. Furthermore, it is worth noting that the Kraft-McMillan inequality can be extended to a countable set of symbols (see Theorem 5.2.2 and the corollary following Theorem 5.5.1 in [19]). When including countable sets of symbols, the inequality is referred to as the Extended Kraft-McMillan Inequality.*

An immediate corollary to the Kraft-McMillan Inequality is the following:

Corollary 5.5.3. *If $\ell_1, \ell_2, \ldots, \ell_m$ are the codeword lengths for any uniquely decodable code, then there exists an instantaneous code with these same code lengths.*

We call a uniquely decodable code C **optimal** whenever the expected length $EL(C)$ is minimized; i.e. the optimal uniquely decodable code is given by

$$C^{\text{opt}} := \text{argmin}_C \{EL(C) : \sum_i 2^{-\ell_i} \leq 1\}.$$

We set $EL^*(X) := EL(C^{\text{opt}})$ the **optimal expected length** of X. The results for the optimal expected length are summarized in the following:

Theorem 5.5.4. *([19, Theorem 5.4.1]) Let X be a random variable with range in the symbol set S. Then the optimal expected length of X satisfies the inequality*

$$H(X) \leq EL^*(X) < H(X) + 1,$$

where $H(X)$ is the Shannon entropy of X, i.e. $H(X) = -\sum_{i=1}^m p_i \log p_i$ where $(p_i)_{i \in S}$ is the pmf of X.

Well known examples of codes which satisfy the inequality of Theorem 5.5.4 are the so-called Huffman codes and Shannon-Fano codes.

In the above theorem, we are only interested in the compressability of single codewords. Suppose instead that we wish to compress strings of codewords with code distributions given by a stochastic process $\mathbf{X} = (X_i)_{i=1}^\infty$. Then, for each $n \in \mathbb{N}$, Theorem 5.5.4 holds for the random vector $(X_1, X_2, \ldots, X_n)$, giving

$$H(X_1, X_2, \ldots, X_n) \leq EL^*(X_1, X_2, \ldots, X_n) < H(X_1, X_2, \ldots, X_n) + 1.$$

For each $n \in \mathbb{N}$, we set

$$EL_n^*(\mathbf{X}) := \frac{1}{n} EL^*(X_1, X_2, \ldots, X_n) \tag{5.16}$$

to be the **optimal expected codeword length per symbol** for the first n symbols. We can then express the optimal expected codeword length per symbol (over all symbols) in terms of the entropy rate (Equation (5.9) or (5.10)).

Theorem 5.5.5. *([19, Theorem 5.4.2]) The minimum expected codeword length per symbol for a stochastic process* $\mathbf{X} = (X_i)_{i=1}^{\infty}$ *satisfies*

$$\frac{H(X_1, X_2, \ldots, X_n)}{n} \leq EL_n^*(\mathbf{X}) < \frac{H(X_1, X_2, \ldots, X_n)}{n} + \frac{1}{n}.$$

Moreover, if $\mathbf{X}$ *is such that the limit defining entropy rate exists (e.g.* $\mathbf{X}$ *is a stationary stochastic process), then*

$$EL_n^*(\mathbf{X}) \to H(\mathbf{X}) \quad \textit{as } n \to \infty.$$

In particular, if $\mathbf{X}$ *consists of independent identically distributed (i.i.d.) copies of a random variable* X*, then*

$$EL_n^*(\mathbf{X}) \to H(X) \quad \textit{as } n \to \infty.$$

The fact that $H(\mathbf{X}) = H(X)$ for any stochastic process $\mathbf{X}$ consisting of i.i.d. copies of a random variable X is a simple consequence of Theorem 5.2.4. This finishes our brief overview of data compression in classical information theory. For a more detailed exposition see [19, Chapter 5].

Chapter 6

Słomczyński-Życzkowski Entropy

The remaining chapters are devoted to the presentation of dynamical entropy in quantum systems. Unlike classical systems, where KS entropy and entropy rate have established themselves as *the* dynamical entropies, there are many competing definitions for quantum dynamical entropy (QDE) (all of which are valid generalizations of KS entropy). Moreover, the relationship between the different definitions is not fully understood, although some work in this direction has been done (e.g. [4, 61, 46]).

In this chapter we consider the QDE introduced by Słomczyński and Życzkowski in [59]. In Section 6.1 we provide the definitions for the Słomczyński-Życzkowski dynamical entropy. In Theorem 6.2.5 at the end of Section 6.2 we show that SZ entropy is not linear in the time interval between successive measurements which answers an open problem posed in [59, page 5692 Question (2)]. This result is in contrast to KS entropy which is linear in time (see Proposition 5.4.2). Moreover, since entropy rate is nonlinear in time (see Proposition 5.4.3), the result gives further evidence that measurements of a deterministic quantum system produce properties that are probabilistic in nature. In Section 6.3, we apply SZ entropy to the Hadamard walk and its square with a variety of Lüders - von Neumann instruments. The results in Theorems 6.3.2 and 6.3.3 show explicitly the nonlinearity of SZ entropy established in Theorem 6.2.5. Moreover, by comparing the results of Theorems 6.3.2 and 6.3.3, we provide further evidence of the sensitivity of quantum systems to measurement.

6.1 SZ Entropy: Definition

Słomczyński and Życzkowski introduced their version of a QDE in 1994 in [59]. In that paper, the authors use a semi-classical approach to develop a QDE using the general notions of state space, phase space, observables and instruments introduced in Chapter 3. A key benefit of the QDE introduced in [59] is that it is not guaranteed to be zero for finite systems, unlike some of the others. This is because their approach, at least as developed here, draws on the intuitions gained by study probabilistic random walks (Section 2.1). Compared to some of the other approaches, which are guaranteed to be zero for finite systems and have a dynamical systems type viewpoint, we already see parallels to classical dynamical entropies (see Section 5.4).

Let (X, K) be a state space and $u \in K$ be a state. Let (Ω, Σ) be a phase space, $\mathcal{T}$ an instrument and Θ a τ-preserving automorphism of X; i.e. $\tau(\Theta v) = \tau(v)$ for all $v \in X$. Let (Ω^*, Σ^*) be the measurable space defined for symbolic dynamics in Section 5.3. We will define an instrument and state-dependent probability measure, $\mu^{(\Theta,\mathcal{T},u)}$, on (Ω^*, Σ^*) which is similar, but not the same, as the one we used for symbolic dynamics. First, we define the values of $\mu^{(\Theta,\mathcal{T},u)}$ on the cylinder sets in Σ^* with an initial interval of time sequences, $\{k\}_{k=1}^{n}$ for some $n \in \mathbb{N}$, by

$$\mu^{(\Theta,\mathcal{T},u)}\left(C\begin{pmatrix} A_1 & \cdots & A_n \\ 1 & \cdots & n \end{pmatrix}\right) = \tau(\mathcal{T}(A_n) \circ \Theta \circ \cdots \circ \mathcal{T}(A_2) \circ \Theta \circ \mathcal{T}(A_1)u), \tag{6.1}$$

for all $A_1, \ldots, A_n \in \Sigma$. Since the collection of cylinder sets with an initial interval of time sequences form a π-system which generates Σ^*, there is a unique extension of $\mu^{(\Theta,\mathcal{T},u)}$ to (Ω^*, Σ^*) by the π-λ Theorem.

Notice that, for a stochastic process $\mathbf{X}$, we defined the measure $\mu^{\mathbf{X}}$ first on cylinder sets with arbitrary time sequences (Equation (2.10)), whereas the measure, $\mu^{(\Theta,\mathcal{T},u)}$ in Equation (6.1), was defined first on the cylinder sets with an initial interval of time sequences. By defining $\mu^{(\Theta,\mathcal{T},u)}$ in this way, we have that, for $A_1, A_3 \in \Sigma$,

$$\mu^{(\Theta,\mathcal{T},u)}\left(C\begin{pmatrix} A_1 & A_3 \\ 1 & 3 \end{pmatrix}\right) = \mu^{(\Theta,\mathcal{T},u)}\left(C\begin{pmatrix} A_1 & \Omega & A_3 \\ 1 & 2 & 3 \end{pmatrix}\right)$$

$$= \tau(\mathcal{T}(A_3) \circ \Theta \circ \mathcal{T}(\Omega) \circ \Theta \circ \mathcal{T}(A_1)u),$$

which is not necessarily equal to $\tau(\mathcal{T}(A_3) \circ \Theta^2 \circ \mathcal{T}(A_1)u)$. Therefore we interpret $\mu^{(\Theta,\mathcal{T},u)}(C\left(\begin{smallmatrix} A_1 & A_3 \\ 1 & 3 \end{smallmatrix}\right))$ as the probability that a system in initial state u will be measured at times $1, 2, 3$ and record the measurement sequence (A_1, A_3) at times 1 and 3. In other words, we must assume that the instrument $\mathcal{T}$ is interacting with the system at all integer times, regardless of whether or not we record a measurement. Just as in Section 2.3, we will use the notation

$$\mu^{(\Theta,\mathcal{T},u)}(A_1, \ldots, A_n) := \mu^{(\Theta,\mathcal{T},u)}(C\left(\begin{smallmatrix} A_1 & \cdots & A_n \\ 1 & \cdots & n \end{smallmatrix}\right)), \tag{6.2}$$

for all $A_1, \ldots, A_n \in \Sigma$ and $n \in \mathbb{N}$, whenever we consider initial time sequences $1, \ldots, n$.

Define the (Ω^*, Ω) stochastic process $\mathbf{X}^{(\Theta,\mathcal{T},u)} = (X_n^{(\Theta,\mathcal{T},u)})_{n=1}^{\infty}$ by setting $X_n^{(\Theta,\mathcal{T},u)}(x) = x_n$ for each $x = (x_m)_{m\in\mathbb{N}} \in \Omega^*$ and, for each $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, define the $(\Omega^*, \mathcal{C})$ stochastic process $\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)} = (X_n^{(\Theta,\mathcal{T},u,\mathcal{C})})_{n=1}^{\infty}$ by $X_n^{(\Theta,\mathcal{T},u,\mathcal{C})} = i_{\mathcal{C}} \circ X_n^{(\Theta,\mathcal{T},u)}$, where $i_{\mathcal{C}} : \Omega \to \mathcal{C}$ is the natural map that assigns to each $x \in \Omega$ the unique $A \in \mathcal{C}$ such that $x \in A$. Note that even though the formulas of $\mathbf{X}^{(\Theta,\mathcal{T},u)}$ and $\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)}$ do not depend on Θ, $\mathcal{T}$ and u, the measure $\mu^{(\Theta,\mathcal{T},u)}$ on their domain, Ω^*, depends on Θ, $\mathcal{T}$ and u.

We define the **Słomczyński-Życzkowski (SZ) entropy of $(\Theta, \mathcal{T}, u)$ with respect to $\mathcal{C}$** to be the entropy rate of the stochastic process $\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)}$. It is related to $\mu^{(\Theta,\mathcal{T},u)}$ by the equation

$$h^{SZ}(\Theta, \mathcal{T}, u, \mathcal{C}) := H(\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)}) = \limsup_{n\to\infty} \frac{1}{n} \sum_{\substack{A_k \in \mathcal{C} \\ 1\le k\le n}} \eta(\mu^{(\Theta,\mathcal{T},u)}(A_1, \ldots, A_n)). \tag{6.3}$$

The second equality follows from Equations (5.7) and (5.9).

Remark 6.1.1. *Let $(\Omega^*, \Sigma^*, \mu^{(\Theta,\mathcal{T},u)}, s)$ be the symbolic dynamics for the system $(\Theta, \mathcal{T}, u)$, where s is the shift transformation given in Equation* (2.11)*. From Proposition 5.3.1 and the definition of $\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)}$ it is clear that*

$$h^{SZ}(\Theta, \mathcal{T}, u, \mathcal{C}) = H(\mathbf{X}_{\mathcal{C}}^{(\Theta,\mathcal{T},u)}) = h^{KS}(s, \hat{\mathcal{C}}).$$

Next we split the SZ entropy of $(\Theta, \mathcal{T}, u)$ with respect to $\mathcal{C}$ into two different causes for randomness. The first cause of randomness is that caused by the choice of instrument, is referred to as the **measurement SZ entropy** and is given by

$$h^{SZ}_{\text{meas}}(\mathcal{T}, u, \mathcal{C}) := h^{SZ}(\mathbb{1}, \mathcal{T}, u, \mathcal{C}). \tag{6.4}$$

The second cause of randomness is given by the dynamics; i.e. the automorphism Θ, is referred to as the **dynamical SZ entropy**, and is given by the difference

$$h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, u, \mathcal{C}) := h^{SZ}(\Theta, \mathcal{T}, u, \mathcal{C}) - h^{SZ}_{\text{meas}}(\mathcal{T}, u, \mathcal{C}).$$

Finally, we define the **dynamical SZ entropy of $(\Theta, \mathcal{T}, u)$** by

$$h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, u) := \sup_{\mathcal{C} \in \mathcal{P}_{ar}(\Omega) H(\hat{\mathcal{C}}) < \infty} h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, u, \mathcal{C}). \tag{6.5}$$

In the remainder of this section, we consider the dynamical SZ entropy of the commutative QDS associated to a classical DS with sharp measurements (Example 3.2.3). We begin by showing that the measurement SZ entropy for classical sharp measurement instruments is equal to zero.

Lemma 6.1.2. *Let $(\Omega, \mathcal{B})$, (X, K), τ and $\mathcal{T}$ be as in Example 3.2.3. Then, for any state $\mu \in K$ and partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ with finite entropy, we have $h^{SZ}_{meas}(\mathcal{T}, \mu, \mathcal{C}) = 0$.*

Proof. Fix a state $\mu \in K$ and a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ with finite entropy. Then, for any $n \in \mathbb{N}$ and $A_1, \ldots, A_n \in \mathcal{C}$, we have that

$$\mu^{(\mathbb{1}, \mathcal{T}, \mu)}(A_1, \ldots, A_n) = \begin{cases} \mu(A_1) & \text{if } A_1 = \cdots = A_n \\ 0 & \text{else} \end{cases}$$

Therefore

$$\begin{aligned} h^{SZ}_{\text{meas}}(\mathcal{T}, \mu, \mathcal{C}) &= \lim_{n \to \infty} \frac{1}{n} \sum_{A \in \mathcal{C}} \eta(\mu(A)) \\ &= \lim_{n \to \infty} \frac{1}{n} H(\mathcal{C}) = 0. \end{aligned}$$

□

The following result shows that the dynamical SZ entropy is a valid generalization of KS entropy (Equation (5.5)) by showing that the KS entropy of a dynamical system with sharp measurements is equal to the dynamical SZ entropy of that system. The result is claimed without proof in [59, Proposition 4(A)]; we provide a proof for completeness.

Proposition 6.1.3. *Let $(\Omega, \mathcal{B})$, (X, K), τ and $\mathcal{T}$ be as in Example 3.2.3. Let $\mu \in K$ be a state; i.e. a probability measure on $(\Omega, \mathcal{B})$, and $f : \Omega \to \Omega$ a measurable map so that $(\Omega, \mathcal{B}, \mu, f)$ is a DS. Let $T_f : X \to X$ be the Koopman operator given in Equation (2.12). Then for each $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$, $h^{KS}(f, \mathcal{C}) = h^{SZ}_{dyn}(T_f, \mathcal{T}, \mu, \mathcal{C})$.*

Proof. Fix a partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$. For all $n \in \mathbb{N}_0$ and $A_0, \ldots, A_n \in \mathcal{C}$ we see that

$$\begin{aligned}
\mu^{(T_f,\mathcal{T},\mu)}(A_1, \ldots, A_n) &= \tau(\mathcal{T}(A_n) \circ T_f \circ \cdots \circ T_f \circ \mathcal{T}(A_0)\mu) \quad \text{by (6.1)} \\
&= (\mathcal{T}(A_n) \circ T_f \circ \cdots \circ T_f \circ \mathcal{T}(A_0))\mu(X) \\
&= (\mathcal{T}(A_n) \circ T_f \circ \cdots \circ \mathcal{T}(A_1) \circ T_f)\mu(A_0) \quad \text{by (3.2)} \\
&= (\mathcal{T}(A_n) \circ T_f \circ \cdots \circ T_f \circ \mathcal{T}(A_1))\mu(f^{-1}(A_0)) \\
&= \vdots \\
&= \mu(A_{i_n} \cap f^{-1}(A_{i_1}) \cap \cdots \cap f^{-n}(A_{i_0})).
\end{aligned}$$

where we used Equation (2.12) in equality 4. Using Remark 2.2.3 and Lemma 6.1.2, we get

$$h^{KS}(f, \mathcal{C}) = h^{SZ}(T_f, \mathcal{T}, \mu, \mathcal{C}) = h^{SZ}_{\text{dyn}}(T_f, \mathcal{T}, \mu, \mathcal{C}).$$

□

If we compare Proposition 6.1.3 to Proposition 4.2.1 (and Corollary 7.1.2 given in the next chapter) we see the versatility of the SZ entropy. Namely, we are able to compare the SZ and KS entropies directly in the classical mechanics picture without resorting to an associated QDS. It is worth noting, however, that the dynamical SZ entropy of the associated QDS is still equal to the KS entropy of the original DS.

In this section, we restrict our attention to the SZ entropy of quantum systems measured with Lüders - von Neumann instruments (see Example 3.2.4). The following lemma, stating that the measurement SZ entropy is zero for Lüders - von Neumann instruments, is claimed in [58]. For completeness we provide the proof.

Lemma 6.2.1. *Let* $(\Omega, \mathcal{P}(\Omega))$, (X, K) *and* H *be as in Example 3.2.4 and let* $\mathcal{T}$ *be a Lüders-von Neumann instrument. Then* $h^{SZ}_{meas}(\mathcal{T}, \rho, \mathcal{C}) = 0$ *for any state* $\rho \in K$ *and any* $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ *with finite entropy; i.e.* $H(\hat{\mathcal{C}}) < \infty$.

Proof. Let $(P_i)_{i\in\Omega}$ be the family of pairwise orthogonal projections that governs $\mathcal{T}$ and fix a state $\rho \in K$. Since the family, $(P_i)_{i\in\Omega}$, is pairwise orthogonal we have, for any $n \in \mathbb{N}$ and $A_1, \dots, A_n \in \mathcal{P}(\Omega)$, that

$$\mu^{(\mathbb{1},\mathcal{T},\rho)}(A_1, \dots, A_n) = \begin{cases} \sum_{a\in A_1} \operatorname{tr}(P_a \rho P_a) = \mu^{(\mathbb{1},\mathcal{T},\rho)}(A_1), & \text{if } A_1 = \cdots = A_n \\ 0, & \text{else} \end{cases}$$

Therefore, for any $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ with $H(\hat{\mathcal{C}}) < \infty$, we have

$$\begin{aligned} h^{SZ}_{\text{meas}}(\mathcal{T}, \rho, \mathcal{C}) &= \lim_{n\to\infty} \frac{1}{n} \sum_{A\in\mathcal{C}} \eta(\mu^{(\mathbb{1},\mathcal{T},\rho)}(A)) \quad \text{by (6.3) and (6.4)} \\ &= \lim_{n\to\infty} \frac{1}{n} H(\hat{\mathcal{C}}) = 0 \quad \text{by the definition of } \hat{\mathcal{C}}. \end{aligned}$$

□

Remark 6.2.2. *It is natural to consider only the partitions* $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ *with finite entropy in Lemma 6.2.1, because these are the only partitions considered in Equation* (6.5).

Fix a discrete phase space (Ω, Σ) with $|\Omega| = N$. Then Lemma 6.2.1, together with Equation (5.12), implies that $h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, \rho, \mathcal{C}) = h^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{C})$ for any unitary transformation Θ, partition $\mathcal{C}$, state ρ, and coherent states instrument $\mathcal{T}$. Recall that, given a unitary U on a Hilbert space H, the **unitary transformation, $\Theta : X \to X$, of U** is given by

$$\Theta(\cdot) = U \cdot U^*. \tag{6.6}$$

Next we examine the properties of SZ entropy for coherent states instruments. The following lemma gives a simplification of Equation (6.1) for coherent states instruments. The moreover statement of the following lemma is mentioned in [58, page 3] without proof.

Lemma 6.2.3. *Let H, (X, K), τ and $(\Omega, \mathcal{P}(\Omega))$ be as in Example 3.2.4. Let $\gamma = (P_i)_{i\in\Omega}$ be a partition of unity on H such that each P_i is a rank-1 projection, and let $a_i \in H$ be such that $P_i = |a_i\rangle\langle a_i|$, for each $i \in \Omega$. Let $\mathcal{T}$ be the coherent states instrument governed by γ, U a unitary operator on H, Θ the unitary transformation of U and $\rho \in K$ a state. Then, for all $n \in \mathbb{N}$ and $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$,*

$$\mu^{(\Theta,\mathcal{T},\rho)}(A_1, \ldots, A_n) = \sum_{\substack{a_k\in A_k \\ 1\leq k\leq n}} \langle a_1|\rho|a_1\rangle \prod_{k=2}^{n} |\langle a_k|U|a_{k-1}\rangle|^2. \tag{6.7}$$

Moreover, $\mathbf{X}^{(\Theta,\mathcal{T},\rho)}$ is a Markov process governed by the transition matrix P on Ω with (i, j)-entry given by $|\langle a_i|U|a_j\rangle|^2$, for all $i, j \in \Omega$.

Proof. By direct calculation, Equation (6.1) simplifies to

$$\begin{aligned}
\mu^{(\Theta,\mathcal{T},\rho)}(A_1, \ldots, A_n) &= \tau(\mathcal{T}(A_n) \circ \Theta \circ \cdots \circ \Theta \circ \mathcal{T}(A_1)\rho) \\
&= \sum_{\substack{a_k\in A_k \\ 1\leq k\leq n}} \mathrm{tr}(\mathcal{T}(\{a_n\}) \circ \Theta \circ \cdots \circ \Theta \circ \mathcal{T}(\{a_1\})\rho) \\
&= \sum_{\substack{a_k\in A_k \\ 1\leq k\leq n}} \mathrm{tr}(P_{a_n} U \cdots U P_{a_1} \rho P_{a_1} U^* \cdots U^* P_{a_n}) \\
&= \sum_{\substack{a_k\in A_k \\ 1\leq k\leq n}} \mathrm{tr}(|a_n\rangle\langle a_n|U \cdots U|a_1\rangle\langle a_1|\rho|a_1\rangle \times \\
&\qquad\qquad \langle a_1|U^* \cdots U^*|a_n\rangle\langle a_n|) \\
&= \sum_{\substack{a_k\in A_k \\ 1\leq k\leq n}} \langle a_1|\rho|a_1\rangle \prod_{k=2}^{n} |\langle a_k|U|a_{k-1}\rangle|^2,
\end{aligned}$$

where the second to last equality follows from writing $P_{a_k} = |a_k\rangle\langle a_k|$, for all $1 \leq k \leq n$ and the last equality follows since $\overline{\langle a_{k-1}|U^*|a_k\rangle} = \langle a_k|U|a_{k-1}\rangle$, for all $2 \leq k \leq n$. It is immediately clear that $\mathbf{X}^{(\Theta,\mathcal{T},\rho)}$ is a stationary Markov process governed by the transition matrix P. $\square$

It is worth noting that Equation (6.7) is a simplification of the probabilities in [59, Equations (27)-(29)] for Lüders-von Neumann coherent states instruments.

Corollary 6.2.4. *Let H, (X, K), τ and $(\Omega, \mathcal{P}(\Omega))$ be as in Example 3.2.4. $\gamma = (P_i)_{i\in\Omega}$ be a partition of unity on H such that each P_i is a rank-1 projection, and let $a_i \in H$ be such that $P_i = |a_i\rangle\langle a_i|$, for each $i \in \Omega$, $\mathcal{T}$ the coherent states instrument governed by γ, U a unitary operator on H, Θ the unitary transformation of U, $\rho \in K$ a state and P the transition matrix defined in Lemma 6.2.3. Then*

$$h^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = \lim_{n\to\infty} \sum_{y\in\Omega} (P^n\mu)_y \sum_{x\in\Omega} \eta(|\langle a_x|U|a_y\rangle|^2),$$

where $\mu = p_{X_1^{(\Theta,\mathcal{T},\rho)}}$ and $\mathcal{A}$ is the atomic partition of Ω. Moreover, whenever $\mu = (\mu_y)_{y\in\Omega}$ is P-invariant, we have

$$h^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = \sum_{y\in\Omega} \mu_y \sum_{x\in\Omega} \eta(|\langle a_x|U|a_y\rangle|^2).$$

Proof. This follows immediately from Lemma 6.2.3 and Theorem 5.2.4. $\square$

In [59, Section IV] the authors require that the state ρ is invariant in the sense that

$$\Theta(\mathcal{T}(\Omega)\rho) = \rho \tag{6.8}$$

when defining SZ entropy for coherent states instruments. This seems to be due to the fact that, in [59, Proposition 2(B)], the authors show that, under Assumption (6.8), for a general coherent states instrument, the stochastic process $X^{(\Theta,\mathcal{T},\rho)}$ is stationary and hence, by the "moreover" part of Corollary 6.2.4, $h^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = \sum_{y\in\Omega} \mu_y \sum_{x\in\Omega} \eta(|\langle a_x|U|a_y\rangle|^2)$. We find Assumption (6.8) restrictive and do not adopt it here. It is also worth mentioning that another invariance condition often imposed on quantum dynamical systems is that $\Theta(\rho) = \rho$. For instance, for AF entropy in [6, Page 76] which is formulated for general C^*-algebras and defined in Chapter 7, the authors require that a state ω satisfies $\omega \circ \Theta = \omega$ which is equivalent to $\Theta(\rho) = \rho$ in the Hilbert space quantum mechanics picture whenever ω is a normal state given by $\omega(\cdot) = \mathrm{tr}(\rho\cdot)$. Also, given a Hilbert space H, a unitary operator U on H and a norm-1 eigenvector, $x \in H$, of U, the pure (or vector) state $\rho = |x\rangle\langle x|$ satisfies $\Theta(\rho) = \rho$, where Θ is the unitary transformation of U. There has been a lot of interest in finding these pure, invariant states in the literature for UQRWs (see e.g. [36, 24]). Therefore $\Theta(\rho) = \rho$ seems another natural definition of invariance. However, we will show in Proposition 6.3.1 that $\Theta(\rho) = \rho$ does not imply that $X^{(\Theta,\mathcal{T},\rho)}$ is a stationary stochastic process.

The following result states that SZ entropy is not linear in the time interval between successive measurements which answers an open problem posed in [59, page 5692 Question (2)]. This result is in contrast to KS entropy which is linear in time (see Proposition 5.4.2). Moreover, since entropy rate is nonlinear in time (see Proposition 5.4.3), the result gives further evidence that measurements of a deterministic quantum system produce properties that are probabilistic in nature.

Theorem 6.2.5. *Let (X, K) and τ be as in Example 3.2.4. Let $(\Omega, \mathcal{P}(\Omega))$ be a discrete phase space with $|\Omega| = N$ for some $N \in \mathbb{N}$, $\mathcal{T}$ a Lüders-von Neumann instrument, Θ a τ-preserving automorphism and $\rho \in K$ a state. Then $h^{SZ}_{dyn}(\Theta^n, \mathcal{T}, \rho) \leq N$ for all $n \in \mathbb{N}$. Therefore, if $h^{SZ}_{dyn}(\Theta, \mathcal{T}, \rho) \neq 0$, then $h^{SZ}_{dyn}(\Theta^n, \mathcal{T}, \rho) \neq nh^{SZ}_{dyn}(\Theta, \mathcal{T}, \rho)$ for all sufficiently large $n \in \mathbb{N}$.*

Proof. Let $(\Omega^*, \mathcal{P}(\Omega)^*, \mu^{(\Theta^n,\mathcal{T},\rho)}, s)$ be the symbolic dynamics of $(\Theta^n, \mathcal{T}, \rho)$ for each $n \in \mathbb{N}$ and let $\mathcal{A}$ be the atomic partition of Ω. Using Equation (5.12), we have that $H_{\mu^{(\Theta^n,\mathcal{T},\rho)}}(\vee_{k=1}^m s^{-k}(\hat{\mathcal{A}})) \leq \ln(|\vee_{k=1}^m s^{-k}(\hat{\mathcal{A}})|) \leq m \ln N$, for all $n, m \in \mathbb{N}$. Therefore $h^{SZ}_{\mathrm{dyn}}(\Theta^n, \mathcal{T}, \rho, \mathcal{A}) \leq \ln N$, for each $n \in \mathbb{N}$. If $h^{SZ}_{\mathrm{dyn}}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = k \neq 0$, then we have, for all $n > \frac{\ln N}{k}$, that

$$h_{\text{dyn}}^{SZ}(\Theta^n, \mathcal{T}, \rho, \mathcal{A}) \leq \ln N < nh_{\text{dyn}}^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{A})$$

Since $\hat{\mathcal{A}}$ is a generating partition for $(\Omega^*, \mathcal{P}(\Omega)^*, \mu^{(\Theta^n, \mathcal{T}, \rho)}, s)$, we have ${}_{\text{dyn}}^{SZ}(\Theta^n, \mathcal{T}, \rho)$ $= h_{\text{dyn}}^{SZ}(\Theta^n, \mathcal{T}, \rho, \mathcal{A})$ and $nh_{\text{dyn}}^{SZ}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = nh_{\text{dyn}}^{SZ}(\Theta, \mathcal{T}, \rho)$. The result follows. $\square$

In [58] the authors establish a class of instruments which have positive dynamical SZ entropy and we give further such examples in Section6.3. Therefore Proposition 6.2.5 does establish the nonlinearity of dynamical SZ entropy in time. In the next section we give examples exhibiting the nonlinearity of dynamical SZ entropy.

6.3 SZ Entropy of the Hadamard walk

This section is dedicated to applying SZ entropy to the Hadamard walk and its square with a variety of instruments. The results in Theorems 6.3.2 and 6.3.3 show explicitly the nonlinearity of SZ entropy established in Theorem 6.2.5. Moreover, by comparing the results of Theorems 6.3.2 and 6.3.3, we provide further evidence of the sensitivity of quantum systems to measurement.

Let $H = H_C \otimes H_P$, $C = \{R, L\}$ and take the phase space to be $(C \times V, \mathcal{P}(C \times V))$. Take (X, K) be the state space $(S_1^{sa}(H), S_1^{+}(H))$ as in Example 3.2.4. Define the parition of unity $\gamma = (P_e)_{e \in C \times V}$, where, for each $(c, v) \in C \times V$,

$$P_e = |c, v\rangle\langle c, v|, \text{ whenever } e = (c, v) \in C \times V. \tag{6.9}$$

Let $\mathcal{T}$ be the coherent states instrument governed by γ. The next proposition states that, for a unitary transformation Θ and a state $\rho \in K$, $\Theta(\rho) = \rho$ does not imply that the associated Markov chain, $X^{(\Theta, \mathcal{T}, \rho)}$, is stationary. The result shows that this natural definition of invariance for ρ is not sufficient for the stationarity of $X^{(\Theta, \mathcal{T}, \rho)}$, whereas Assumption (6.8), imposed by the authors of [59], does guarantee that $X^{(\Theta, \mathcal{T}, \rho)}$ is stationary.

Proposition 6.3.1. *Let Θ be the Hadamard walk on a vertex set V with $|V| = N \geq 2$ defined in Equation* (4.3). *Let $\mathcal{T}$ be the coherent states instrument given by the partition of unity $\gamma = (P_e)_{e \in C \times V}$ (Equation* (6.9)*) and let $\mathcal{A}$ be the atomic partition on $C \times V$. Let $|x\rangle = \frac{1}{\sqrt{N(4+2\sqrt{2})}}((1+\sqrt{2})|R\rangle + |L\rangle) \otimes \sum_{v \in V} |v\rangle$ (which is a unit norm eigenvector for the unitary matrix, U, of the Hadamard walk) and $\rho = |x\rangle\langle x|$. Then the pmf, $p_{X_1^{(\Theta,\mathcal{T},\rho)}}$, of $X_1^{(\Theta,\mathcal{T},\rho)}$ is not P-invariant, where P is the transition matrix defined in Lemma 6.2.3. Furthermore, the dynamical SZ entropy is equal to $h^{SZ}_{dyn}(\Theta, \mathcal{T}, \rho) = \ln 2$.*

Proof. For each $(c, v) \in C \times V$,

$$p_{X_1^{(\Theta,\mathcal{T},\rho)}}(c, v) = \langle c, v|\rho|c, v\rangle = \frac{1}{N(4+2\sqrt{2})}((3+2\sqrt{2})\delta_{c,R} + \delta_{c,L}). \tag{6.10}$$

Also, for each $e = (c, v), f = (d, u) \in C \times V$, a straightforward calculation yields

$$|\langle e|U|f\rangle|^2 = \begin{cases} \frac{1}{2} & c = R \text{ and } u = v - 1 \\ \frac{1}{2} & c = L \text{ and } u = v + 1 \\ 0 & \text{else} \end{cases} \tag{6.11}$$

$$= \frac{1}{2}\delta_{u, v-(-1)^{\delta_{c,L}}}. \tag{6.12}$$

Recall that $|\langle e|U|f\rangle|^2$ is the (e, f)-entry of P, for each $e, f \in C \times V$. Thus, for each $e = (c, v) \in C \times V$,

$$\begin{aligned} (Pp_{X_1^{(\Theta,\mathcal{T},\rho)}})_e &= \sum_{f \in C \times V} p_{X_1^{(\Theta,\mathcal{T},\rho)}}(f)|\langle e|U|f\rangle|^2 \\ &= \frac{1}{2}(p_{X_1^{(\Theta,\mathcal{T},\rho)}}(R, v - (-1)^{\delta_{c,L}}) + p_{X_1^{(\Theta,\mathcal{T},\rho)}}(L, v - (-1)^{\delta_{c,L}})) \\ &= \frac{1}{2}(\frac{3+2\sqrt{2}}{N(4+2\sqrt{2})} + \frac{1}{N(4+2\sqrt{2})}) = \frac{1}{2N}, \quad \text{by (6.10)} \end{aligned}$$

where we used Equation (6.11) in equality 2.

Therefore $Pp_{X_1^{(\Theta,\mathcal{T},\rho)}} \neq p_{X_1^{(\Theta,\mathcal{T},\rho)}}$ and thus $\mathbf{X}^{(\Theta,\mathcal{T},\rho)}$ is not stationary. Continuing to find the dynamical SZ entropy, we see that $Pp_{X_1^{(\Theta,\mathcal{T},\rho)}}$ is the uniform distribution, μ, on $C \times V$, which is invariant with respect to P. Thus Corollary 6.2.4 and Lemma 6.2.1 imply that

$$h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, \rho, \mathcal{A}) = \sum_{f \in C \times V} \mu_f \sum_{e \in C \times V} \eta(|\langle e|U|f\rangle|^2) = \sum_{f \in C \times V} \frac{1}{2N} 2\eta(\frac{1}{2}) = \ln 2.$$

This implies which is equal to $h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, \rho) = \ln 2$ because $\hat{\mathcal{A}}$ is a generating partition for $(\Omega^*, \mathcal{P}(\Omega)^*, \mu^{(\Theta, \mathcal{T}, \rho)}, s)$ by Corollary 5.3.3. □

As the UQRW is a quantum analogue of the classical random walk, it is natural to consider measurements of the position space only. There are two options for how to go about this. One option is to take the phase space to be $(C \times V, \mathcal{P}(C \times V))$, the coherent states instrument $\mathcal{T}$ to be given by the partition of unity $\gamma = (P_e)_{e \in C \times V}$, defined in Equation (6.9), and calculate the dynamical SZ entropy with respect to the partition

$$\mathcal{C}_V = \{C_v\}_{v \in V}, \text{ where } C_v := \{|R, v\rangle, |L, v\rangle\}, \text{ for each } v \in V. \tag{6.13}$$

On the other hand we could take the phase space to be $(V, \mathcal{P}(V))$, define the projections

$$P_v = \mathbb{1}_{H_C} \otimes |v\rangle\langle v|, \text{ for each } v \in V, \tag{6.14}$$

and calculate the dynamical SZ entropy of the Lüders-von Neumann instrument $\mathcal{V}$ with respect to the atomic partition of V, where $\mathcal{V}$ is governed by the partition of unity $\lambda = (P_v)_{v \in V}$. We will calculate the entropies for both these scenarios (with the same initial state) on the Hadamard walk, Θ, and its square, Θ^2. We will see that the two interpretations do not yield the same entropy. This is further evidence to the sensitivity of a closed quantum system to measurement. Furthermore, Theorems 6.3.2 and 6.3.3 provide concrete examples illustrating the fact that dynamical SZ entropy is not linear in time. In fact, one can also see that the dynamical SZ entropy is nonlinear in time by considering Θ^3 in Proposition 6.3.1, but we do not include the calculation here.

Theorem 6.3.2. *Let Θ be the the Hadamard walk on V with $|V| = N \geq 3$. Let $\mathcal{T}$ be the coherent states instrument governed by the partition of unity $\gamma = (P_e)_{e \in C \times V}$ given in Equation* (6.9)*, $\rho = \frac{\mathbb{1}_H}{2N}$ and $\mathcal{C}_V$ the partition given in Equation* (6.13)*. Then $h^{SZ}_{dyn}(\Theta, \mathcal{T}, \rho, \mathcal{C}_V) = \ln 2$ and $h^{SZ}_{dyn}(\Theta^2, \mathcal{T}, \rho, \mathcal{C}_V) = \frac{3}{2} \ln 2$.*

Proof. Notice that $p_{X_1^{(\Theta,\mathcal{T},\rho)}}(c, v) = \langle c, v|\rho|c, v\rangle = \frac{1}{2N}$ for all $(c, v) \in C \times V$ and recall that the transition matrix P, which governs Θ with respect to the coherent states instrument $\mathcal{T}$, has entries given by Equation (6.11). In the following, it will be more convenient to rewrite Equation (6.11) viewing f as the fixed index. In this manner, for each $f = (c, v) \in C \times V$, we have

$$U|f\rangle = \frac{1}{\sqrt{2}}(|R, v+1\rangle + (-1)^{\delta_{c,L}}|L, v-1\rangle), \tag{6.15}$$

and hence

$$|\langle e|U|f\rangle|^2 = \frac{1}{2}(\delta_{e,(R,v+1)} + \delta_{e,(L,v-1)}). \tag{6.16}$$

Also, we have

$$\begin{aligned}
\mu^{(\Theta,\mathcal{T},\rho)}(C_{v_1}, \ldots, C_{v_n}) &= \sum_{\substack{c_k \in \{R,L\} \\ 1 \leq k \leq n}} \langle c_1, v_1|\rho|c_1, v_1\rangle \prod_{k=2}^{n} |\langle c_k, v_k|U|c_{k-1}, v_{k-1}\rangle|^2 \\
&= \sum_{\substack{c_k \in \{R,L\} \\ 1 \leq k \leq n}} \frac{1}{2N} \prod_{k=2}^{n} \left(\frac{\delta_{v_k,v_{k-1}+1}\delta_{c_k,R} + \delta_{v_k,v_{k-1}-1}\delta_{c_k,L}}{2} \right) \\
&= \frac{1}{N} \prod_{k=2}^{n} \left(\frac{\delta_{v_k,v_{k-1}+1} + \delta_{v_k,v_{k-1}-1}}{2} \right),
\end{aligned}$$

for all $v_1, \ldots, v_n \in V$, where we used Equation (6.7) in equality 1 and Equation (6.16) in equality 2. These are exactly the probabilities $p_{\mathbf{X}}(v_1, \ldots, v_n)$ of a stationary Markov chain $\mathbf{X}$ which is governed by the transition matrix, Q, for the unbiased random walk on the N-cycle V. Therefore

$$h^{SZ}_{\text{dyn}}(\Theta, \mathcal{T}, \rho, \mathcal{C}_V) = H(Q) = \ln 2,$$

where the second equality follows from Proposition 5.4.3.

Next we show that $h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{T}, \rho, \mathcal{C}_V) = \frac{3}{2}\ln 2$. For all $f = (c, v) \in C \times V$, we have

$$\begin{aligned} U^2|f\rangle &= \frac{1}{\sqrt{2}} U(|R, v+1\rangle + (-1)^{\delta_{c,L}}|L, v-1\rangle) \quad \text{by (6.15)} \\ &= \frac{1}{2}((-1)^{\delta_{R,c}}|L, v-2\rangle + (-1)^{\delta_{L,c}}|R, v\rangle + |L, v\rangle + |R, v+2\rangle) \end{aligned} \tag{6.17}$$

and hence

$$|\langle e|U^2|f\rangle|^2 = \begin{cases} \frac{1}{4} & e = (R, v) \text{ or } (L, v) \\ \frac{1}{4} & e = (R, v+2) \\ \frac{1}{4} & e = (L, v-2) \\ 0 & \text{else} \end{cases} \tag{6.18}$$

Notice that $|\langle e|U^2|f\rangle|^2$ in Equation (6.18) does not depend on the coin space component of f. Therefore

$$\begin{aligned} &\mu^{(\Theta^2, \mathcal{T}, \rho)}(C_{v_1}, \ldots, C_{v_n}) \\ &= \sum_{\substack{c_k \in \{R,L\} \\ 1 \le k \le n}} \langle c_1, v_1|\rho|c_1, v_1\rangle \prod_{k=2}^{n} |\langle c_k, v_k|U^2|c_{k-1}, v_{k-1}\rangle|^2 \quad \text{by (6.7)} \\ &= \sum_{\substack{c_k \in \{R,L\} \\ 1 \le k \le n}} \frac{1}{2N} \prod_{k=2}^{n} \left(\frac{\delta_{c_k,L}\delta_{v_k,v_{k-1}-2} + \delta_{c_k,L}\delta_{v_k,v_{k-1}}}{4} \right. \\ &\qquad\qquad \left. + \frac{\delta_{c_k,R}\delta_{v_k,v_{k-1}} + \delta_{c_k,R}\delta_{v_k,v_{k-1}+2}}{4} \right) \\ &= \frac{1}{N} \prod_{k=2}^{n} \left(\frac{1}{4}\delta_{v_k,v_{k-1}-2} + \frac{1}{2}\delta_{v_k,v_{k-1}} + \frac{1}{4}\delta_{v_k,v_{k-1}+2} \right), \end{aligned}$$

for all $v_1, \ldots, v_n \in V$, where we used Equation (6.7) in equality 2. These are exactly the probabilities $p_{\mathbf{Y}}(v_1, \ldots, v_n)$ of a stationary Markov chain $\mathbf{Y}$ which is governed by the transition matrix Q^2. Therefore

$$h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{T}, \rho, \mathcal{C}_V) = H(Q^2) = \frac{3}{2}\ln 2,$$

where the second equality follows from Proposition 5.4.3. $\square$

Theorem 6.3.3. *Let Θ be the Hadamard walk on V with $|V| = N \geq 3$ defined in Equation (4.3). Let $\mathcal{V}$ be the Lüders-von Neumann instrument governed by the* *partition of unity* $\lambda = (P_v)_{v\in V}$ *defined in Equation* (6.14) *and* $\rho = \frac{\mathbb{1}_H}{2N}$. *Then* $h^{SZ}_{dyn}(\Theta, \mathcal{V}, \rho) = \ln 2$ *and* $h^{SZ}_{dyn}(\Theta^2, \mathcal{V}, \rho) = \frac{4}{3}\ln 2$.

Proof. Notice that from Equation (6.1), for each $m, n \in \mathbb{N}$ and $v_1, \ldots, v_n \in V$, we have

$$\begin{aligned} \mu^{(\Theta^m,\mathcal{V},\rho)}(v_1,\ldots,v_n) &:= \tau(\mathcal{V}(v_n)\circ\Theta^m\circ\cdots\circ\Theta^m\circ\mathcal{V}(v_1)\rho) \\ &= \operatorname{tr}(P_{v_n}U^m\cdots U^m P_{v_1}\rho P_{v_1}(U^m)^*\cdots(U^m)^*P_{v_n}). \end{aligned} \tag{6.19}$$

Also, notice that $\rho = \frac{1}{2N}\sum_{v\in V}(|R,v\rangle\langle R,v| + |L,v\rangle\langle L,v|)$ and so, for each $m, n \in \mathbb{N}$ and $v_1, \ldots, v_n \in V$, Equation (6.19) becomes

$$\begin{aligned} &\mu^{(\Theta^m,\mathcal{V},\rho)}(v_1,\ldots,v_n) \\ &= \sum_{c\in\{R,L\}}\frac{1}{2N}\operatorname{tr}(P_{v_n}U^m\cdots U^m P_{v_1}|c,v_1\rangle\langle c,v_1|P_{v_1}(U^m)^*\cdots(U^m)^*P_{v_n}) \\ &= \sum_{c,d\in\{R,L\}}\frac{1}{2N}|\langle d,v_n|U^m P_{v_{n-1}}\cdots P_{v_1}U^m|c,v_1\rangle|^2. \end{aligned} \tag{6.20}$$

Let $\mathcal{A}$ be the atomic partition of V. We first show that $h^{SZ}_{\text{dyn}}(\Theta, \mathcal{V}, \rho, \mathcal{A}) = \ln 2$. Notice that for $(c, v) \in C \times V$, $UP_v|c, v\rangle = U|c, v\rangle$ and is given by Equation (6.15). Thus, by direct calculation, we have that

$$\begin{aligned} &\mu^{(\Theta,\mathcal{V},\rho)}(v_1,\ldots,v_n) \\ &= \sum_{c_1,c_n\in\{R,L\}}\frac{1}{2N}|\langle c_n,v_n|UP_{v_{n-1}}\cdots P_{v_2}U|c_1,v_1\rangle|^2 \quad \text{by (6.20)} \\ &= \frac{1}{2N}\sum_{c_\cdot,c_n\in\{R,L\}}|\langle c_n,v_n|UP_{v_{n-1}}\cdots P_{v_1}\times \\ &\qquad (\frac{1}{\sqrt{2}}(|R,v_1+1\rangle + (-1)^{\delta_{c_1,L}}|L,v_1-1\rangle))|^2 \\ &= \frac{1}{2N}\sum_{c_1,c_2,c_n\in\{R,L\}}\frac{1}{2}(\delta_{v_2,v_1+1}\delta_{c_2,R} + \delta_{v_2,v_1-1}\delta_{c_2,L})\times \\ &\qquad |\langle c_n,v_n|UP_{v_{n-1}}\cdots P_{v_3}U|c_2,v_2\rangle|^2 \\ &\quad\vdots \\ &= \sum_{\substack{c_k\in\{R,L\}\\1\le k\le n}}\frac{1}{2N}\prod_{k=2}^{n}\left(\frac{\delta_{v_k,v_{k-1}+1}\delta_{c_k,R} + \delta_{v_k,v_{k-1}-1}\delta_{c_k,L}}{2}\right) \end{aligned}$$

$$= \frac{1}{N}\prod_{k=2}^{n}\left(\frac{\delta_{v_k,v_{k-1}+1}+\delta_{v_k,v_{k-1}-1}}{2}\right),$$

for all $v_1,\ldots,v_n \in V$, which are exactly the probabilities $p_{\mathbf{X}}(v_1,\ldots,v_n)$ of a stationary Markov chain $\mathbf{X}$ which is governed by the transition matrix, Q, for the unbiased random walk on the N-cycle V. Therefore

$$h_{\text{dyn}}^{SZ}(\Theta,\mathcal{V},\rho,\mathcal{A}) = H(Q) = \ln 2,$$

where the second equality follows from Proposition 5.4.3. Moreover, since $\hat{\mathcal{A}}$ is a generating partition for $(V^*,\mathcal{P}(V)^*,\mu^{(\Theta,\mathcal{V},\rho)},s)$, we have that

$$h_{\text{dyn}}^{SZ}(\Theta,\mathcal{V},\rho) = \ln 2$$

by Corollary 5.3.3.

Next we show that $h_{\text{dyn}}^{SZ}(\Theta^2,\mathcal{V},\rho,\mathcal{A}) = \frac{4}{3}\ln 2$ using path counting techniques. To that end, for each $n \in \mathbb{N}$ and n-tuple $v = (v_1,\ldots,v_n) \in V^n$, we set

$$l_v := |\{k : k < n \text{ such that } v_k = v_{k+1} = \cdots = v_n\}|.$$

Then, we define the sets

$$L_{\text{c}}^n := \{\bar{v} \in V^n : \bar{v} = (v,\ldots,v) \text{ for some } v \in V\},$$

$$L_{\text{e}}^n := \{v \in V^n : l_v \text{ is even}\}\backslash L_{\text{c}}^n \text{ and } L_{\text{o}}^n := \{v \in V^n : l_v \text{ is odd}\}\backslash L_{\text{c}}^n.$$

For each $n \in \mathbb{N}$ and $v = (v_1,\ldots,v_n) \in V^n$, we will identify v with the cylinder set $C\left(\begin{smallmatrix} v_1 & \cdots & v_n \\ 1 & \cdots & n\end{smallmatrix}\right)$ and consider $\mathcal{L}^n := \{L_{\text{c}}^n, L_{\text{e}}^n, L_{\text{o}}^n\}$ as a partition of V^*.

With this correspondence, we will show that the conditional probabilities, $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_{n+1}|v)$, for v_{n+1} given $v = (v_1,\ldots,v_n)$ are dependent upon which set $L \in \mathcal{L}^n$ that v belongs to. First, we will determine the change of coin state that occurs after measuring the walker at the same site a number of times in a row. We claim that the resulting coin state, after n measurements at a site v, depends only on the initial coin state $c \in \{R, L\}$ and the congruence class of n modulo 4. Specifically, for all $n \in \mathbb{N}$, $\bar{v} = (v, \ldots, v) \in V^{n+1}$ and $c \in \{R, L\}$, we claim the following:

if $n \equiv 0 \mod 4$, then

$$\underbrace{P_v U^2 \cdots P_v U^2}_{n \text{ times}} |c, v\rangle = a|c, v\rangle \text{ for some } a \in \mathbb{C} \text{ with } |a| = \frac{1}{2^{\lfloor \frac{n+1}{2} \rfloor}}, \tag{6.21}$$

if $n \equiv 1 \mod 4$, then

$$P_v U^2 \cdots P_v U^2 |c, v\rangle = a|L + (-1)^{\delta_{c,L}} R, v\rangle, \text{ for some } a \in \mathbb{C} \text{ with } |a| = \frac{1}{2^{\lfloor \frac{n+1}{2} \rfloor}}, \tag{6.22}$$

if $n \equiv 2 \mod 4$, then

$$P_v U^2 \cdots P_v U^2 |c, v\rangle = a|c^\perp, v\rangle \text{ for some } a \in \mathbb{C} \text{ with } |a| = \frac{1}{2^{\lfloor \frac{n+1}{2} \rfloor}}, \tag{6.23}$$

where we set $R^\perp = L$ and $L^\perp = R$, and, if $n \equiv 3 \mod 4$, then

$$P_v U^2 \cdots P_v U^2 |c, v\rangle = a|L - (-1)^{\delta_{c,L}} R, v\rangle, \text{ for some } a \in \mathbb{C} \text{ with } |a| = \frac{1}{2^{\lfloor \frac{n+1}{2} \rfloor}}, \tag{6.24}$$

where we used the abbreviation $|L \pm R, v\rangle := |L, v\rangle \pm |R, v\rangle$. We will prove the claims by induction on n.

The base case, $n = 0$, is trivial. For the inductive step we will handle the different congruence classes of n separately. To this end, let $m \in \mathbb{N}$ with $m \geq 1$ and suppose that for all $n < m$ Equations (6.21)-(6.24) hold for all $\bar{v} \in V^{n+1}$ and their respective values of n. Fix $\bar{v} \in V^{m+1}$ and $c \in \{R, L\}$. If $m \equiv 1 \mod 4$, then, for some $a \in \mathbb{C}$ with $|a| = \frac{1}{2^{\lfloor \frac{m}{2} \rfloor}}$, we have

$$\begin{aligned} P_v U^2 \cdots P_v U^2 |c, v\rangle &= P_v U^2 a|c, v\rangle \text{ by (6.21)} \\ &= \frac{a}{2}|L + (-1)^{\delta_{c,L}} R, v\rangle \text{ by (6.17)}, \end{aligned}$$

and Equation (6.22) is satisfied since $\frac{1}{2 \cdot 2^{\lfloor \frac{m}{2} \rfloor}} = \frac{1}{2^{\lfloor \frac{m+1}{2} \rfloor}}$. If $m \equiv 2 \mod 4$, then, for some $a \in \mathbb{C}$ with $|a| = \frac{1}{2^{\lfloor \frac{m}{2} \rfloor}}$, we have

$$P_v U^2 \cdots P_v U^2 |c, v\rangle = P_v U^2 a|L + (-1)^{\delta_{c,L}} R, v\rangle \text{ by (6.22)}$$

$$= \quad (-1)^{\delta_{c,L}} a|c^{\perp}, v\rangle,$$

where the second equality holds because

$$U^2|R+L, v\rangle = \sqrt{2}U|R, v+1\rangle = |L, v\rangle + |R, v+2\rangle \text{ and} \tag{6.25}$$
$$U^2|L-R, v\rangle = -\sqrt{2}U|L, v-1\rangle = |L, v-2\rangle - |R, v\rangle,$$

for all $v \in V$. Thus Equation (6.23) is satisfied since $\frac{1}{2^{\lfloor \frac{m}{2} \rfloor}} = \frac{1}{2^{\lfloor \frac{m+1}{2} \rfloor}}$. If $m \equiv 3 \mod 4$, then, for some $a \in \mathbb{C}$ with $|a| = \frac{1}{2^{\lfloor \frac{m}{2} \rfloor}}$, we have

$$\begin{aligned} P_v U^2 \cdots P_v U^2 |c, v\rangle &= P_v U^2 a|c^{\perp}, v\rangle \text{ by (6.23)} \\ &= \frac{a}{2}|L - (-1)^{\delta_{c,L}} R, v\rangle \text{ by (6.17)}, \end{aligned}$$

where we used the fact that $(-1)^{\delta_{c^{\perp},L}} = -(-1)^{\delta_{c,L}}$ in the second equality. Hence Equation (6.24) is satisfied since $\frac{1}{2 \cdot 2^{\lfloor \frac{m}{2} \rfloor}} = \frac{1}{2^{\lfloor \frac{m+1}{2} \rfloor}}$. If $m \equiv 0 \mod 4$, then, for some $a \in \mathbb{C}$ with $|a| = \frac{1}{2^{\lfloor \frac{m}{2} \rfloor}}$, we have

$$\begin{aligned} P_v U^2 \cdots P_v U^2 |c, v\rangle &= P_v U^2 a|L - (-1)^{\delta_{c,L}} R, v\rangle \text{ by (6.24)} \\ &= -(-1)^{\delta_{c,L}} a|c, v\rangle \text{ by (6.25)} \end{aligned}$$

and hence Equation (6.21) is satisfied since $\frac{1}{2^{\lfloor \frac{m}{2} \rfloor}} = \frac{1}{2^{\lfloor \frac{m+1}{2} \rfloor}}$. Therefore the induction is complete and the claims are verified.

Next we claim that for all $v = (v_1, \ldots, v_n) \in V^n \backslash L_{\mathrm{c}}^n$ with $p_{\mathbf{X}^{(\Theta^2, \mathcal{V}, \rho)}}(v) \neq 0$ there exists some $\psi \in H_C$ such that for all $v_{n+1} \in V$ the conditional pmf of $\mathbf{X}^{(\Theta^2, \mathcal{V}, \rho)}$ is given by

$$p_{\mathbf{X}^{(\Theta^2, \mathcal{V}, \rho)}}(v_{n+1}|v_1, \ldots, v_n) = \frac{\sum_{d \in \{R,L\}} |\langle d, v_{n+1}|U^2|\psi, v_n\rangle|^2}{\|\psi\|^2}. \tag{6.26}$$

Indeed, for all $v \in V^n$ with $p_{\mathbf{X}^{(\Theta^2, \mathcal{V}, \rho)}}(v) \neq 0$, we have

$$\begin{aligned} & p_{\mathbf{X}^{(\Theta^2, \mathcal{V}, \rho)}}(v_{n+1}|v_1, \ldots, v_n) \\ = \; & \frac{\mu^{(\Theta^2, \mathcal{V}, \rho)}(v_1, \ldots, v_{n+1})}{\mu^{(\Theta^2, \mathcal{V}, \rho)}(v_1, \ldots, v_n)} \qquad (6.27) \\ = \; & \frac{\sum_{c,d \in \{R,L\}} |\langle d, v_{n+1}|U^2 P_{v_n} \cdots P_{v_2} U^2|c, v_1\rangle|^2}{\sum_{c,d \in \{R,L\}} |\langle d, v_n|U^2 P_{v_{n-1}} \cdots P_{v_2} U^2|c, v_1\rangle|^2} \quad \text{by (6.20)} \end{aligned}$$

$$= \sum_{c\in\{R,L\}} q_c \frac{\sum_{d\in\{R,L\}} |\langle d, v_{n+1}|U^2 P_{v_n}\cdots P_{v_2}U^2|c,v_1\rangle|^2}{\sum_{d\in\{R,L\}} |\langle d, v_n|U^2 P_{v_{n-1}}\cdots P_{v_2}U^2|c,v_1\rangle|^2}, \tag{6.28}$$

where, for each $c \in \{R, L\}$, we set

$$q_c := \frac{\sum_{d\in\{R,L\}} |\langle d, v_n|U^2 P_{v_{n-1}}\cdots P_{v_2}U^2|c,v_1\rangle|^2}{\sum_{c',d\in\{R,L\}} |\langle d, v_n|U^2 P_{v_{n-1}}\cdots P_{v_2}U^2|c',v_1\rangle|^2}. \tag{6.29}$$

Notice that if $q_c = 0$ in Equation (6.29) then the denominator on the right hand side of Equation (6.28) is also equal to 0. In this case, we will use the convention that their product is defined and equal to 0.

For each $c \in \{R, L\}$, we define $\psi_c \in H_C$ to be the unique element satisfying the equation

$$P_{v_n}U^2\cdots P_{v_2}U^2|c,v_1\rangle = |\psi_c, v_n\rangle.$$

Then Equation (6.28) simplifies to

$$p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_{n+1}|v_1,\ldots,v_n) = \sum_{c\in\{R,L\}} q_c \frac{\sum_{d\in\{R,L\}} |\langle d, v_{n+1}|U^2|\psi_c, v_n\rangle|^2}{\|\psi_c\|^2}, \tag{6.30}$$

where equality in the denominator follows by Parseval's identity.

Notice that Equation (6.17) gives

$$\begin{aligned} \mathrm{Ran}(P_{v+2}U^2P_v) &= \mathrm{span}\{|R, v+2\rangle\} \text{ and} \\ \mathrm{Ran}(P_{v-2}U^2P_v) &= \mathrm{span}\{|L, v-2\rangle\}, \text{ for all } v \in V. \end{aligned} \tag{6.31}$$

Moreover, Equation (6.31) implies that, for any operator $A \in B(H)$,

$$\begin{aligned} \mathrm{Ran}(P_{v+2}U^2P_vA) &\subseteq \mathrm{span}\{|R, v+2\rangle\} \text{ and} \\ \mathrm{Ran}(P_{v-2}U^2P_vA) &\subseteq \mathrm{span}\{|L, v-2\rangle\}, \text{ for all } v \in V. \end{aligned}$$

Hence, if $v \in V^n \backslash L_c^n$, then for each $c \in \{R, L\}$ we have

$$P_{v_{n-l_v}}U^2\cdots P_{v_2}U^2|c,v_1\rangle = a_c|d, v_{n-l_v}\rangle, \text{ for some } a_c \in \mathbb{C}, \tag{6.32}$$

where $d = R$ whenever $v_{n-l_v} = v_{n-l_v-1} + 2$ and $d = L$ when $v_{n-l_v} = v_{n-l_v-1} - 2$. Thus d does not depend on the initial coin state c. For each $d \in \{R, L\}$, we also have that

$$P_{v_n}U^2\cdots P_{v_{n-l_v+1}}U^2|d, v_{n-l_v}\rangle = a|\psi, v_n\rangle, \text{ for some } a \in \mathbb{C} \text{ with } |a| = \frac{1}{2^{\lfloor \frac{l_v+1}{2}\rfloor}}, \tag{6.33}$$

where ψ is the coin state given by Equations (6.21)-(6.24) depending on the congruence class of l_v modulo 4 and we used the fact that $v_{n-l_v} = v_{n-l_v+1} = \cdots = v_n$ by definition of l_v. We combine Equations (6.32) and (6.33) to get that, for each $c \in \{R, L\}$,

$$P_{v_n}U^2 \cdots P_{v_2}U^2|c, v_1\rangle = a'_c|\psi, v_n\rangle, \tag{6.34}$$

where $a'_c = a_c \cdot a$ with a_c and a coming from Equations (6.32) and (6.33), respectively. Since $q_R + q_L = 1$ and both ψ_R and ψ_L in Equation (6.30) are equal to ψ which appears in Equation (6.34), we see that Equation (6.30) simplifies to Equation (6.26) as claimed.

Next we claim that, for $n \in \mathbb{N}$, if $v = (v_1, \dots, v_n) \in L_o^n$ and $p_{\mathbf{X}(\Theta^2,\mathcal{V},\rho)}(v) \neq 0$, then

$$p_{\mathbf{X}(\Theta^2,\mathcal{V},\rho)}(v_{n+1}|v_1, \dots, v_n) = \begin{cases} \frac{1}{2} & \text{if } v_{n+1} = v_n \\ \frac{1}{2} & \text{if } v_{n+1} \text{ is exactly one of } v_n \pm 2 \\ 0 & \text{else} \end{cases}, \tag{6.35}$$

where the exactly one value of $v_{n+1} \in \{v_n - 2, v_n + 2\}$ with nonzero conditional probability depends on the given sequence $(v_1, \dots, v_n)$ in the following manner:

- If
  - (i) $v_{n-l_v} = v_{n-l_v-1} + 2$ and $l_v = 1 \bmod 4$, or
  - (ii) $v_{n-l_v} = v_{n-l_v-1} - 2$ and $l_v = 3 \bmod 4$,

  then $v_{n+1} = v_n + 2$.
- If
  - (iii) $v_{n-l_v} = v_{n-l_v-1} - 2$ and $l_v = 1 \bmod 4$, or
  - (iv) $v_{n-l_v} = v_{n-l_v-1} + 2$ and $l_v = 3 \bmod 4$,

  then $v_{n+1} = v_n - 2$.

In addition we claim that, for $n \in \mathbb{N}$ and $v = (v_1, \dots, v_n) \in V^n$, if $v \in L_e^n \cup L_c^n$ and

$p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v) \neq 0$, then

$$p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_{n+1}|v_0,\ldots,v_n) = \begin{cases} \frac{1}{2} & \text{if } v_{n+1} = v_n \\ \frac{1}{4} & \text{if } v_{n+1} = v_n + 2 \\ \frac{1}{4} & \text{if } v_{n+1} = v_n - 2 \\ 0 & \text{else} \end{cases} \tag{6.36}$$

In order to see Equation (6.35), let $v \in L_{\mathrm{o}}^n$ with $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v) \neq 0$ and suppose v satisfies the conditions for Case (i); i.e. $v_{n-l_v} = v_{n-l_v-1} + 2$ and $l_v \equiv 1 \mod 4$. Since $v_{n-l_v} = v_{n-l_v-1} + 2$, the coin state, d, on the right hand side of Equation (6.32) is $d = R$. Using this, the fact that $l_v \equiv 1 \mod 4$ and Equation (6.22), we see that the coin state, ψ, on the right hand sides of Equations (6.33) and (6.34) is given by $\psi = R + L$. Plugging into Equation (6.26) and using (6.25), we have

$$p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_{n+1}|v_1,\ldots,v_n) = \begin{cases} \frac{1}{2} & \text{if } v_{n+1} = v_n \\ \frac{1}{2} & \text{if } v_{n+1} = v_n + 2 \\ 0 & \text{else} \end{cases}$$

in this case. The other three cases can be done similarly and thus we obtain that Equation (6.35) is satisfied for all $v \in L_{\mathrm{o}}^n$.

Next, for the proof of Equation (6.36), let $v \in L_{\mathrm{e}}^n$ with $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v) \neq 0$. By Equations (6.21) and (6.23), we can see that the coin state ψ in Equation (6.34) is given by $\psi = c$, for some $c \in \{R, L\}$. Plugging this value of ψ into Equation (6.26) and using (6.18) we can see that the conditional pmf, $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}$, of $\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}$ is given by Equation (6.36), for all $v \in L_{\mathrm{e}}^n$.

It remains only to show that Equation (6.36) is valid for all $\bar{v} = (v,\ldots,v) \in L_{\mathrm{c}}^n$. Since the modulus of a in Equations (6.21)-(6.24) is independent of $c \in \{R, L\}$, we have $q_c = \frac{1}{2}$ in Equation (6.29) for both values of c. Note that by Equations (6.21)-(6.24) we have that the vector $\psi_c \in H_C$ which appears in Equation (6.30) is given by

$$\psi_c = \begin{cases} c & \text{if } n-1 \equiv 0 \bmod 4 \\ L + (-1)^{\delta_{c,L}} R & \text{if } n-1 \equiv 1 \bmod 4 \\ c^{\perp} & \text{if } n-1 \equiv 2 \bmod 4 \\ L - (-1)^{\delta_{c,L}} R & \text{if } n-1 \equiv 3 \bmod 4 \end{cases}$$

Thus if $n-1$ is even, then by Equations (6.30) and (6.18) we obtain immediately Equation (6.36). If $n-1$ is odd, then we must examine the cases $n-1 \equiv 1 \mod 4$ and $n-1 \equiv 3 \mod 4$ separately. If $n-1 \equiv 1 \mod 4$ then $\psi_R = L+R$, $\psi_L = L-R$ and Equations (6.30) and (6.25) give Equation (6.36). The case of $n-1 \equiv 3 \mod 4$ can be verified similarly.

We are now set to show $h^{SZ}_{\text{dyn}}(\Theta^2, \mathcal{V}, \rho, \mathcal{A}) = \frac{4}{3} \ln 2$. By direct calculation, we have that

$$\begin{aligned} & H_{\mu^{(\Theta^2,\mathcal{V},\rho)}}(X^{(\Theta^2,\mathcal{V},\rho)}_{n+1} | (X^{(\Theta^2,\mathcal{V},\rho)}_1, \dots, X^{(\Theta^2,\mathcal{V},\rho)}_n)) \\ = \quad & \sum_{\substack{v_k \in V \\ 1 \leq k \leq n}} p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_1, \dots, v_n) \sum_{v_{n+1} \in V} \eta(p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_{n+1} | v_1, \dots, v_n)) \quad \text{by (5.8)} \\ = \quad & \mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_{\text{e}} \cup L^n_{\text{c}})(2\eta(\frac{1}{4}) + \eta(\frac{1}{2})) + \mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_{\text{o}})(2\eta(\frac{1}{2})) \\ = \quad & \mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_{\text{e}} \cup L^n_{\text{c}})\frac{3}{2} \ln 2 + \mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_{\text{o}}) \ln 2, \end{aligned} \tag{6.37}$$

where we used Equations (6.35) and (6.36) in equality 2 and

$$\mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_x) = \sum_{(v_1,\dots,v_n) \in L^n_x} p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v_1, \dots, v_n) \text{ for each } x \in \{\text{c}, \text{e}, \text{o}\}.$$

It remains only to solve for

$$\lim_{n\to\infty} \mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_x) \text{ for each } x \in \{\text{c}, \text{e}, \text{o}\}.$$

Notice that, by definition of L^n_{c}, we have that

$$\mu^{(\Theta^2,\mathcal{V},\rho)}(L^n_{\text{c}}) = \sum_{v \in V} p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(\underbrace{v, \dots, v}_{n \text{ times}}), \text{ for each } n \in \mathbb{N}.$$

Using Equation (6.36) $n-1$ times and Equation (6.20) to see that $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v) = \frac{1}{N}$, we obtain $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(\underbrace{v,\ldots,v}_{n \text{ times}}) = \frac{1}{2^{n-1}N}$ and hence

$$\mu^{(\Theta^2,\mathcal{V},\rho)}(L_{\mathrm{c}}^n) = \frac{1}{2^{n-1}}, \text{ for each } n \in \mathbb{N}. \tag{6.38}$$

For ease of notation, set $e_n = \mu^{(\Theta^2,\mathcal{V},\rho)}(L_{\mathrm{e}}^n)$, $o_n = \mu^{(\Theta^2,\mathcal{V},\rho)}(L_{\mathrm{o}}^n)$ and $c_n = \mu^{(\Theta^2,\mathcal{V},\rho)}(L_{\mathrm{c}}^n)$, for each $n \in \mathbb{N}$. Notice that $e_1 = o_1 = 0$, $c_1 = 1$ and $c_n = \frac{1}{2^{n-1}}$, for all $n \in \mathbb{N}$, by Equation (6.38). For each $v = (v_1, \ldots, v_{n-1}) \in V^{n-1}$ and $v_n \in V$ let $v \circ v_n := (v_1, \ldots, v_n) \in V^n$. Suppose $v \in L_{\mathrm{o}}^{n-1}$. If $v_n = v_{n-1}$ then $l_{v \circ v_n} = l_v + 1$ and if $v_n \neq v_{n-1}$ then $l_{v \circ v_n} = 0$. Thus $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v \circ v_n \in L_{\mathrm{e}}^n | v \in L_{\mathrm{o}}^{n-1}) = 1$ and $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v \circ v_n \in L_x^n | v \in L_{\mathrm{o}}^{n-1}) = 0$ for $x \in \{\mathrm{o}, \mathrm{c}\}$. Suppose $v \in L_{\mathrm{e}}^{n-1} \cup L_{\mathrm{c}}^{n-1}$. Then $v \circ v_n \in L_{\mathrm{e}}^n$ exactly when $v_n \neq v_{n-1}$. Therefore Equation (6.36) gives $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v \circ v_n \in L_{\mathrm{e}}^n | v \in L_x^{n-1}) = \frac{1}{2}$ for each $x \in \{\mathrm{e}, \mathrm{c}\}$. Therefore

$$e_n = o_{n-1} + \frac{1}{2}(e_{n-1} + c_{n-1}), \text{ for all } n \geq 2. \tag{6.39}$$

Equation (6.36) also gives that $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v \circ v_n \in L_{\mathrm{o}}^n | v \in L_{\mathrm{e}}^{n-1}) = \frac{1}{2}$; using this together with the fact that $p_{\mathbf{X}^{(\Theta^2,\mathcal{V},\rho)}}(v \circ v_n \in L_{\mathrm{o}}^n | v \in L_{\mathrm{c}}^{n-1}) = 0$, we have

$$o_n = \frac{1}{2} e_{n-1}, \text{ for all } n \geq 2. \tag{6.40}$$

Therefore

$$e_n = \frac{1}{2}(e_{n-1} + e_{n-2} + c_{n-1}), \text{ for all } n \geq 3. \tag{6.41}$$

We claim that the limits $e := \lim_{n\to\infty} e_n$ and $o := \lim_{n\to\infty} o_n$ both exist. It is enough, by Equation (6.40), to show that the limit e exists. To see this we show that $(e_n)_{n\in\mathbb{N}}$ is increasing and bounded. We show that $(e_n)_{n\in\mathbb{N}}$ is increasing by induction. Since $e_1 = 0$ and $e_2 = \frac{1}{2}$ (by Equation (6.39)), the base case is done. Next, fix $n \in \mathbb{N}$ with $n \geq 3$, and suppose that $e_{m-1} < e_m$ for all $m \in \{2, \ldots, n-1\}$. Then, by Equation (6.41), it is enough to show that $e_{n-2} + c_{n-1} > e_{n-1}$. We see that,

$$e_{n-1} = \frac{1}{2}(e_{n-2} + e_{n-3} + c_{n-2}) < e_{n-2} + c_{n-1},$$

where the inequality follows by the inductive hypothesis and the fact that $\frac{1}{2}c_{n-2} = c_{n-1}$. Therefore $(e_n)_{n\in\mathbb{N}}$ is increasing and trivially bounded by 1, and both the limits e and o exist. Furthermore, for all $n \in \mathbb{N}$, $1 = e_n + o_n + c_n$ because $\mathcal{L}^n$ is a partition of V^* and $\lim_{n\to\infty} c_n = 0$ by Equation (6.38). Hence $1 = e + o = \frac{3e}{2}$, implying that $e = \frac{2}{3}$ and $o = \frac{1}{3}$. Taking the limit in Equation (6.37), we see that

$$\lim_{n\to\infty}(e_n + c_n)\frac{3}{2}\ln 2 + o_n \ln 2 = \frac{4}{3}\ln 2.$$

Therefore $h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{V}, \rho) = h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{V}, \rho, \mathcal{A}) = \frac{4}{3}\ln 2$ as desired. □

The fact that $h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{V}, \rho, \mathcal{A}) \neq h_{\text{dyn}}^{SZ}(\Theta^2, \mathcal{T}, \rho, \mathcal{C}_V)$ provides further evidence of the sensitivity of quantum systems to measurement.

Remark 6.3.4. *The results of this section are also submitted for publication in [9].*

Chapter 7

Quantum Markov Chain Entropy

In this section we recall the quantum Markov chain (QMC) approach to quantum dynamical entropy. This QMC approach was first introduced in [3] in terms of the Accardi-Ohya-Watanabe (AOW) entropy. Another QMC approach was introduced by Tuyls in [61] for the study of the Alicki-Fannes (AF) entropy, which was introduced in [6] and often referred to as ALF entropy to emphasize Lindblad's contributions. Finally, a generalization of both QMC approaches was given in [38], where the authors introduced the Kossakowski-Ohya-Watanabe (KOW) entropy. Throughout this chapter, we will follow mainly the terminology and notations of [3] and [38].

7.1 QMC Entropy: Definition

Fix a stationary QDS $(\mathcal{A}, \Theta, \rho)$ and an operational partition of unity $\gamma = (\gamma_i)_{i=1}^d$ on $\mathcal{A}$. Let $\mathcal{E}_\gamma$ and $\mathcal{E}_{\gamma,\Theta}$ be the transition expectations defined in Equations (4.14) and (4.15), respectively, which are both mappings from $M_d \otimes \mathcal{A}$ to $\mathcal{A}$. Also, let $\psi = \{\rho, \mathcal{E}_{\gamma,\Theta}\}$ be the corresponding quantum Markov state and $\{\rho_n\}_{n=1}^\infty$ be the joint correlations for ψ given in Equation (4.17). Recall that ψ is given by

$$\psi(a_1 \otimes \cdots \otimes a_n) = \mathrm{tr}(\rho \mathcal{E}_{\gamma,\Theta}(a_1 \otimes \mathcal{E}_{\gamma,\Theta}(a_2 \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(a_n \otimes \mathbb{1}) \cdots)))),$$

for all $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in M_d$.

We will show that the joint correlations, ρ_n with $n \in \mathbb{N}$, can be given explicitly by applying the lifting $\mathcal{E}^*_{\gamma,\Theta} : \Sigma(\mathcal{A}) \to M_d \otimes \Sigma(\mathcal{A})$ iteratively to the initial state ρ, where $\Sigma(\mathcal{A})$ is the set of normal states on $\mathcal{A}$. The lifting $\mathcal{E}^*_{\gamma,\Theta}$ was first described in [2] and is simply the dual of $\mathcal{E}_{\gamma,\Theta}$.

First, recall that $\mathcal{E}_{\gamma,\Theta} = \Theta \circ \mathcal{E}_\gamma$ (by Equation (4.14)) so that $\mathcal{E}^*_{\gamma,\Theta} = \mathcal{E}^*_\gamma \circ \Theta^*$, where Θ^* is the dual of Θ, and $\mathcal{E}^*_\gamma : \Sigma(\mathcal{A}) \to M_d \otimes \Sigma(\mathcal{A})$ is given by

$$\mathcal{E}^*_\gamma(\sigma) = [\gamma_i \sigma \gamma_j^*]_{i,j=1}^d = \sum_{i,j=1}^d E_{i,j} \otimes \gamma_i \sigma \gamma_j^*, \tag{7.1}$$

where $E_{i,j} := |e_i\rangle\langle e_j| \in M_d$ is the standard matrix basis element with (i,j)-entry equal to 1 and every other entry equal to 0.

For each $n \in \mathbb{N}$ and $\bar{i} = (i_1, \ldots, i_n), \bar{j} = (j_1, \ldots, j_n) \in \{1, \ldots, d\}^n$, the $(\bar{i}, \bar{j})$-entry of ρ_n can be calculated in the following manner:

$$\begin{aligned}
\rho_n(\bar{i}, \bar{j}) &= \langle e_{i_1}, \ldots, e_{i_n} | \rho_n | e_{j_1}, \ldots, e_{j_n} \rangle \\
&= \psi(E_{j_1,i_1} \otimes \cdots \otimes E_{j_n,i_n}) \\
&= \operatorname{tr}(\rho \mathcal{E}_{\gamma,\Theta}(E_{j_1,i_1} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A})))) \\
&= \operatorname{tr}(\mathcal{E}^*_{\gamma,\Theta}(\rho) E_{j_1,i_1} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A}))) \\
&= \operatorname{tr}(\mathcal{E}^*_{\gamma}(\rho) E_{j_1,i_1} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A}))) \text{ since } \rho \text{ is invariant} \\
&= \sum_{i,j=1}^d \operatorname{tr}((E_{i,j} \otimes \gamma_i \rho \gamma_j^*)(E_{j_1,i_1} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A})))) \\
&= \operatorname{tr}(\gamma_{i_1} \rho \gamma_{j_1}^* \mathcal{E}_{\gamma,\Theta}(E_{j_2,i_2} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A})))) \\
&= \operatorname{tr}(\gamma_{i_2} \Theta^*(\gamma_{i_1} \rho \gamma_{j_1}^*) \gamma_{j_2}^* \mathcal{E}_{\gamma,\Theta}(E_{j_3,i_3} \otimes \mathcal{E}_{\gamma,\Theta}(\cdots \mathcal{E}_{\gamma,\Theta}(E_{j_n,i_n} \otimes \mathbb{1}_\mathcal{A})))) \\
&\quad \vdots \\
&= \operatorname{tr}(\gamma_{i_n} \Theta^*(\cdots \Theta^*(\gamma_{i_1} \rho \gamma_{j_1}^*) \cdots) \gamma_{j_n}^*),
\end{aligned} \tag{7.2}$$

where we used Equation (7.1) in equality 6. We have proved the following.

Proposition 7.1.1. *Let $(\mathcal{A}, \Theta, \rho)$ be a QDS, γ be an operational partition of unity on $\mathcal{A}$ and $\psi = \{\rho, \mathcal{E}_{\gamma,\Theta}\}$ be the associated quantum Markov state. Then the joint densities, ρ_n, of ψ given in Equation* (7.2) *are exactly equal to the joint densities $\rho^{(n)}[\gamma]$ given in Equation* (4.10).

Finally, the **QMC dynamical entropy of $(\mathcal{A}, \Theta, \rho)$ with respect to γ** is given by

$$
\begin{aligned}
h^{QMC}(\Theta, \rho, \gamma) &= \limsup_{n\to\infty} \frac{1}{n} S(\rho_n) \\
&= \limsup_{n\to\infty} \frac{1}{n} S(\rho^{(n)}[\gamma]),
\end{aligned} \tag{7.3}
$$

where $S(\cdot) := \mathrm{tr}(\eta(\cdot))$ is the von Neumann entropy and $\rho^{(n)}[\gamma]$ is given in Equation (4.10). Notice that the equality between lines 1 and 2 of Equation (7.3) follows from Proposition 7.1.1. Further, given a subalgebra $\mathcal{B}$ of $\mathcal{A}$, the **QMC dynamical entropy of $(\mathcal{A}, \Theta, \rho)$ with respect to $\mathcal{B}$** is given by

$$
h_{\mathcal{B}}^{QMC}(\Theta, \rho) = \sup_{\gamma \subseteq \mathcal{B}} h^{QMC}(\Theta, \rho, \gamma). \tag{7.4}
$$

The QMC dynamical entropy given in Equation (7.3) has been considered by many in different contexts. For the AF entropy ([6]) and AOW entropy ([3]), the authors consider only those Θ's which are $*$-automorphism, and hence the joint densities are given by Equation (4.12). For AOW entropy the authors made the further restriction that γ be simply a partition of unity and considered the transition expectation given in Remark 4.3.1; when considering the transition expectations in Remark 4.3.1, the resulting joint densities in Equation (7.2) are diagonal, as we will see in Section 7.2. In 1999 the KOW entropy was introduced in [38] and the restriction to $*$-automorphisms was not imposed. Moreover, the authors of [38] allow for more generality by introducing an additional Hilbert space to represent noise. Our main result (Theorem 7.3.9) of this chapter makes use of some of the ideas introduced in that paper, but we will not present the full generality here.

We finish this section by stating a corollary relating the KS entropy of a classical DS to the QMC entropy of the associated commutative QDS.

Corollary 7.1.2. *Let (Ω, Σ, μ, f) be a stationary dynamical system and consider the associated commutative quantum dynamical system $(L^\infty(\Omega), T_f, \mu)$ described in*

Section 4.2. Then, for any partition $\mathcal{C} \in \mathcal{P}_{ar}(\Omega)$ with associated partition of unity γ, we have

$$h^{KS}(f, \mu, \mathcal{C}) = h^{QMC}(T_f, \mu, \gamma).$$

Proof. This is a simple consequence of Propositions 4.2.1 and 7.1.1. See also the last paragraph of Section 4.3. □

Remark 7.1.3. *Let (Ω, Σ, μ, f) and $(L^\infty(\Omega), T_f, \mu)$ be as in Corollary 7.1.2 and consider the unital subalgebra $\mathcal{B}$ of $L^\infty(\Omega)$ consisting of all finite linear combinations of characteristic functions of measurable sets. In [62, Section 5.2], Tuyls shows that $\mathcal{B}$ is dense in $L^\infty(\Omega)$, in a particular sense, and that the partition-independent KS and QMC entropies are equal; i.e.*

$$h^{KS}(f, \mu) = h^{QMC}_{\mathcal{B}}(T_f, \mu).$$

7.2 Accardi-Ohya-Watanabe Entropy

In this section we specialize the QMC approach to the Accardi-Ohya-Watanabe (AOW) dynamical entropy as it was originally introduced in [3]. The only differences between this and the previous section is that in the present section we restrict our attention to *only* partitions of unity and we consider the transition expectation given in Remark 4.3.1. We will not require that our dynamics are $*$-automorphisms (as in [3]), but will consider this case separately. We define this transition expectation below for ease of reading.

Remark 7.2.1. *It is worth noting that we could have included this in the previous section and simply given the QMC dynamical entropy (Equation* (7.3)*) an additional parameter for the transition expectation. However, we decided to present the AOW entropy separately to minimize confusion.*

Let $(\mathcal{A}, \Theta, \rho)$ be a QDS and $\gamma = (\gamma_i)_{i=1}^d$ be a partition of unity on $\mathcal{A}$. Define the transition expectations $E_\gamma, E_{\gamma,\Theta} : M_d \otimes \mathcal{A} \to \mathcal{A}$ by

$$E_\gamma([a_{i,j}]) = \sum_{i=1}^{d} \gamma_i a_{i,i} \gamma_i \quad \text{for all } [a_{i,j}] \in M_d \otimes \mathcal{A} \tag{7.5}$$

and

$$E_{\gamma,\Theta} = \Theta \circ E_\gamma. \tag{7.6}$$

Lastly, let $\psi = \{\rho, E_{\gamma,\Theta}\}$ be the corresponding quantum Markov state. Both ψ and its joint correlations, ρ_n, are given by Equations (4.16) and (4.17), respectively.

The liftings $E_\gamma^*, E_{\gamma,\Theta}^* : \Sigma(\mathcal{A}) \to M_d \otimes \Sigma(\mathcal{A})$, where $\Sigma(\mathcal{A})$ is the collection of normal states on $\mathcal{A}$, are given by $E_{\gamma,\Theta}^* := E_\gamma^* \circ \Theta^*$ and

$$E_\gamma^*(\sigma) := \sum_{i=1}^{d} E_{i,i} \otimes \gamma_i \sigma \gamma_i, \tag{7.7}$$

where $E_{i,j} := |e_i\rangle\langle e_j| \in M_d$ is the standard basis element with (i,j)-entry equal to one and all other entries zero, for all $1 \leq i, j \leq d$. In a similar manner to the work leading up to Equation (7.2), we will give explicitly the entries of the joint correlations, ρ_n. To that end, for each $n \in \mathbb{N}$ and $\bar{i}, \bar{j} \in \{1, \ldots, d\}^n$, we have

$$\begin{aligned}
\rho_n(\bar{i}, \bar{j}) &= \langle e_{i_1}, \ldots, e_{i_n} | \rho_n | e_{j_1}, \ldots, e_{j_n} \rangle \\
&= \psi(E_{j_1,i_1} \otimes \cdots \otimes E_{j_k,i_k}) \\
&= \operatorname{tr}(\rho E_{\gamma,\Theta}(E_{j_1,i_1} \otimes E_{\gamma,\Theta}(\cdots E_{\gamma,\Theta}(E_{j_k,i_k} \otimes \mathbb{1})))) \\
&= \operatorname{tr}(E_{\gamma,\Theta}^*(\rho) E_{j_1,i_1} \otimes E_{\gamma,\Theta}(\cdots E_{\gamma,\Theta}(E_{j_k,i_k} \otimes \mathbb{1}))) \\
&= \delta_{i_1,j_1} \operatorname{tr}(\gamma_{i_1} \Theta^*(\rho) \gamma_{i_1} E_{\gamma,\Theta}(E_{j_2,i_2} \otimes E_{\gamma,\Theta}(\cdots E_{\gamma,\Theta}(E_{j_k,i_k} \otimes \mathbb{1})))) \\
&\;\;\vdots \\
&= \delta_{\bar{i},\bar{j}} \operatorname{tr}(\gamma_{i_n} \Theta^*(\cdots \Theta^*(\gamma_{i_1} \Theta^*(\rho) \gamma_{i_1}) \cdots) \gamma_{i_n}).
\end{aligned} \tag{7.8}$$

Therefore each ρ_n is a diagonal matrix with entries given by Equation (7.8). If ρ is invariant with respect to Θ^*, then Equation (7.8) simplifies to

$$\rho_n(\bar{i}, \bar{j}) = \delta_{\bar{i},\bar{j}} \operatorname{tr}(\gamma_{i_n} \Theta^*(\cdots \Theta^*(\gamma_{i_1} \rho \gamma_{i_1}) \cdots) \gamma_{i_n}) \tag{7.9}$$

and if Θ^* is also a $*$-automorphism we have

$$\rho_n(\bar{i},\bar{j}) = \delta_{\bar{i},\bar{j}} \operatorname{tr}(\gamma_{i_n} \cdots \Theta^{(n-1)*}(\gamma_{i_1})\rho\Theta^{(n-1)*}(\gamma_{i_1}) \cdots \gamma_{i_n}). \tag{7.10}$$

The **AOW dynamical entropy of $(\mathcal{A}, \Theta, \rho)$ with respect to γ** is given by

$$h^{AOW}(\Theta, \rho, \gamma) = \limsup_{n\to\infty} \frac{1}{n} S(\rho_n) = \limsup_{n\to\infty} \frac{1}{n} \sum_{\substack{1\le i_k \le d \\ 1 \le k \le n}} \rho_n(\bar{i},\bar{i}). \tag{7.11}$$

Given a subalgebra $\mathcal{B} \subseteq \mathcal{A}$ the *AOW entropy of* $\boldsymbol{\Theta}$ *with respect to* $\boldsymbol{\mathcal{B}}$ is given by

$$h^{AOW}_{\mathcal{B}}(\Theta) = \sup_{\gamma\subseteq\mathcal{B}} h^{AOW}(\Theta, \gamma). \tag{7.12}$$

Notice that Equations (7.11) and (7.12) are identical to Equations (7.3) and (7.4), respectively.

We finish up this section by presenting an example of AOW entropy. This example was also considered in [3, Model 2], but we present it here for completeness.

Example 7.2.2. *Let $\mathcal{A}$ be a matrix algebra M_d acting on a Hilbert space $H = \mathbb{C}^d$. Let U be a unitary on H, Θ be the corresponding unitary transformation (Equation (6.6)), and suppose that $\rho \in S_1(H) = \Sigma(\mathcal{A})$ is invariant with respect to Θ^*. Fix an orthonormal basis, $\{e_i\}_{i=1}^d$, of H and let $\gamma = (\gamma_i)_{i=1}^d$ be the partition of unity with $\gamma_i = |e_i\rangle\langle e_i|$, for each $1 \le i \le d$. Then Θ is a $*$-automorphism such that $\Theta^n(\gamma_i) = |U^n e_i\rangle\langle U^n e_i|$. Hence, by Equation (7.10), we see that the diagonal entries of the joint correlations, ρ_n, are given by*

$$\begin{aligned}
\rho_n(\bar{i},\bar{i}) &= \operatorname{tr}(\gamma_{i_n} \cdots \Theta^{(n-1)*}(\gamma_{i_1})\rho\Theta^{(n-1)*}(\gamma_{i_1}) \cdots \gamma_{i_n}) \\
&= \prod_{k=1}^{n-1} |\langle U e_{i_{k+1}}, e_{i_k}\rangle|^2 \langle e_{i_1}, \rho e_{i_1}\rangle,
\end{aligned} \tag{7.13}$$

where the equality follows in an analogous manner to that of Lemma 6.2.3. Thus, the diagonal entries of the joint correlations are given by a homogeneous Markov process with initial distribution $\mu = (\mu_i)_{i=1}^d$ such that $\mu_i = \langle e_i, \rho e_i\rangle$, for each $1 \le i \le d$, and *transition matrix* $P = (p_{i,j})_{i,j=1}^d$ *with entries* $p_{i,j} = |\langle Ue_i, e_j\rangle|^2$*, for each* $1 \leq i, j \leq d$*. Hence*

$$h^{AOW}(\Theta, \rho, \gamma) = \lim_{n\to\infty} \sum_{i=1}^{d} (P^n\mu)_i \sum_{j=1}^{d} \eta(|\langle Ue_j, e_i\rangle|^2)$$

by Theorem 5.2.4. Moreover, whenever μ *is* P*-invariant, we have*

$$h^{AOW}(\Theta, \rho, \gamma) = \sum_{i=1}^{d} \langle e_i, \rho e_i\rangle \sum_{j=1}^{d} \eta(|\langle Ue_j, e_i\rangle|^2).$$

Notice the similarities between SZ entropy with coherent states instruments in Corollary 6.2.4 and Example 7.2.2. We will revisit AOW entropy in Subsection 7.3.3.

7.3 Quantum Data Compression using QMC Dynamical Entropy

In this section, we give a QMC dynamical entropy representation for optimal quantum data compression similar to Theorem 5.5.5 in the classical case. In Subsection 7.3.1, we recall the notions of indeterminate length codes for lossless quantum data compression and introduce the notions of stochastic ensembles. In Subsection 7.3.2 we introduce an open quantum random walk representation for stochastic ensembles before applying the QMC dynamical entropy to the OQRW, obtaining the desired extension of Theorem 5.5.5. The QMC representation of the OQRW introduced in Subsection 7.3.2 is new, to the best of my knowledge.

7.3.1 Quantum Data Compression

We will introduce the notions of quantum data compression, extending the notions of Section 5.5. Similar to that section, all codings will be done into strings of (binary) qubits. The extensions to d-qubits can easily be done.

We begin with the description of indeterminate-length quantum codes, whose preliminary investigation began with Schumacher ([52]) and Braunstein et. al in [17]. We may think of the codes introduced in the previous section as being varying-length codes; the term indeterminate-length is used to draw attention to the fact that a quantum code must allow for superpositions of codewords, including those superpositions containing codewords with varying lengths. We will follow mainly the formalisms in [13] as opposed to the zero-extended forms of [53]. A description of the connection between these two formalisms can be found in [16].

For any Hilbert space H, we will denote by $H^{\oplus} := \oplus_{\ell=0}^{\infty} H^{\otimes \ell}$ the **free Fock space** of H, where $H^{\otimes 0} = \mathbb{C}$. We will denote $1 \in H^{\otimes 0} = \mathbb{C}$ by $|\emptyset\rangle$ and refer to it as the **empty string**. Further, we set $H^{\oplus r}$ equal to $\oplus_{\ell=0}^{r} H^{\otimes \ell}$. Let $\mathcal{S} = \{p_n, |s_n\rangle\}_{n=1}^{N}$ be an **ensemble of pure states**, or simply ensemble, where $p = \{p_n\}_{n=1}^{N}$ is the pmf of a random variable X and $|s_n\rangle \in H_{\mathcal{S}} = \mathbb{C}^d$ such that $H_{\mathcal{S}} = \operatorname{span}\{|s_n\rangle\}$. The collection $\{|s_n\rangle\}$ will be referred to as the **symbol states** of the ensemble $\mathcal{S}$. An **(indeterminate-length) quantum code**, U, over a quantum binary alphabet $\mathcal{A} := \{|0\rangle, |1\rangle\}$, which is an orthonormal basis for $H_{\mathcal{A}} = \mathbb{C}^2$, is a linear isometry $U : H_{\mathcal{S}} \to H_{\mathcal{A}}^{\oplus}$. The **extended quantum code of $\boldsymbol{U}$** is the linear mapping $U^+ : H_{\mathcal{S}}^{\oplus} \to H_{\mathcal{A}}^{\oplus}$ given by

$$U^+(|x_1 x_2 \cdots x_n\rangle) = U(|x_1\rangle)U(|x_2\rangle)\cdots U(|x_n\rangle),$$

for all $|x_1 x_2 \cdots x_n\rangle \in H_{\mathcal{S}}^{\otimes n}$ and $n \in \mathbb{N}$, and we set $U^+(|\emptyset\rangle) = |\emptyset\rangle$.

The quantum code U is said to be **uniquely decodable** if the extended quantum code U^+ is an isometry. Throughout this section, we will restrict ourselves only to the situation where $\operatorname{Ran} U \subseteq H_{\mathcal{A}}^{\oplus \ell_{\max}}$ for some $\ell_{\max} \in \mathbb{N}$; i.e. there is a finite upper bound $\ell_{\max}$ on the length of all codewords.

Remark 7.3.1. *The uniquely decodable quantum codes considered in this section can be considered lossless because we guarantee in the definition that the corresponding extended quantum codes are linear isometries; i.e. the codings for strings of symbol states are pairwise distinguishable and hence the original strings can be recovered.*

Remark 7.3.2. *The authors of [16] allow non-empty strings to map to the empty string. In their paper, the authors send along a classical side channel to give the* *lengths of the codewords and so that convention is possible. Without the classical side channel (as is the approach here) allowing non-empty strings to map to the empty string will cause the quantum code to not be uniquely decodable.*

Let $\mathcal{S} = \{p_n, |s_n\rangle\}_{n=1}^N$ be an ensemble whose symbol states span a Hilbert space $H_{\mathcal{S}}$ of dimension d. Consider a classical uniquely decodable code, C, on a symbol set, $S = \{x_i\}_{i=1}^d$, with d-many symbols. We will construct a corresponding uniquely decodable quantum code, U, from C by identifying the classical binary alphabet $A = \{0, 1\}$ with the quantum binary alphabet $\mathcal{A} = \{|0\rangle, |1\rangle\} \subseteq \mathbb{C}^2$ and the symbol set, S, with any orthonormal basis $\{|e_i\rangle\}_{i=1}^d$ of $H_{\mathcal{S}}$; this construction is given in [13].

Fix an orthonormal basis $\{|e_i\rangle\}_{i=1}^d$ of $H_{\mathcal{S}}$ and define the quantum code $U : H_{\mathcal{S}} \to H_{\mathcal{A}}^{\oplus}$ by the equation

$$U = \sum_{i=1}^{d} |C(x_i)\rangle\langle e_i|. \tag{7.14}$$

It is clear that $|C(x_i)\rangle \in H_{\mathcal{A}}^{\otimes \ell_i} \subseteq H_{\mathcal{A}}^{\oplus}$, where ℓ_i is the length of $C(x_i)$, and that $\{|C(x_i)\rangle\}_{i=1}^d$ is an orthonormal set, so that U is a linear isometry. Furthermore, since C is uniquely decodable, the map $U^\ell : H_{\mathcal{S}}^{\otimes \ell} \to H_{\mathcal{A}}^{\oplus}$ defined by the equation

$$U^\ell = \sum_{i_1=1}^{d} \cdots \sum_{i_\ell=1}^{d} |C(x_{i_1})C(x_{i_2}) \cdots C(x_{i_\ell})\rangle\langle e_{i_1} e_{i_2} \cdots e_{i_\ell}| \tag{7.15}$$

is a linear isometry for each $\ell \in \mathbb{N}_0$. Since the extended quantum code $U^+ : H_{\mathcal{S}}^{\oplus} \to H_{\mathcal{A}}^{\oplus}$ is given by

$$U^+ = \sum_{\ell=0}^{\infty} U^\ell, \tag{7.16}$$

we see that U^+ is a linear isometry and hence U is uniquely decodable. We will refer to quantum codes constructed from classical ones by Equation (7.14) as **classical-quantum encoding schemes (c-q schemes)**. (In [13] c-q schemes are referred to as *lossless quantum encoding schemes*.)

Remark 7.3.3. *Notice that the symbol states* $\{|s_n\rangle\}_{n=1}^N$ *of the ensemble* $\mathcal{S}$ *are not directly encoded by the* $|C(x_i)\rangle$*'s unless* $N = d$ *and there exists a permutation* σ *of*

$\{1, \ldots, d\}$ *such that* $|s_{\sigma(i)}\rangle = |e_i\rangle$ *for every* $i \in \{1, \ldots, d\}$. *Indeed* $U|s_n\rangle$ *need not belong to* $H_{\mathcal{S}}^{\otimes \ell}$ *for any* $\ell \in \mathbb{N}$, *but can in general be in a superposition of different lengths. (Hence the term indeterminate-length quantum codes.)*

The Kraft-McMillan Inequality (Theorem 5.5.1) was initially extended to the quantum domain in [53] and subsequently in [13]. Before presenting (a slightly different) Quantum Kraft-McMillan Inequality, we will first introduce the length observable and quantum codes with length eigenstates. The **length observable** Λ acting on $H_{\mathcal{A}}^{\oplus}$ is given by

$$\Lambda := \sum_{\ell=0}^{\infty} \ell \Pi_\ell, \tag{7.17}$$

where Π_ℓ is the orthogonal projection onto the subspace $H_{\mathcal{A}}^{\otimes \ell}$ of $H_{\mathcal{A}}^{\oplus}$.

We say that a quantum code $U : H_{\mathcal{S}} \to H_{\mathcal{A}}^{\oplus}$ has **length eigenstates** if there is an orthonormal basis $\{|e_i\rangle\}_{i=1}^{d}$ of $H_{\mathcal{S}}$ and a sequence $\{|\psi_i\rangle\}_{i=1}^{d} \subseteq H_{\mathcal{A}}^{+}$ with $|\psi_i\rangle \in H_{\mathcal{A}}^{\otimes \ell_i}$ for some $\ell_i \leq \ell_{\max}$, for each i, such that U has the form

$$U = \sum_{i=1}^{d} |\psi_i\rangle\langle e_i|. \tag{7.18}$$

Note that the $|\psi_i\rangle$'s are orthogonal due to U being a linear isometry. It is easy to see that every c-q scheme is a quantum code with length eigenstates. Lastly, for each $\ell \in \mathbb{N} \cup \{0\}$, we will refer to the elements of the set $\{\psi_i : i \in \{1, \ldots, d\},\ \psi_i \in H_{\mathcal{A}}^{\otimes \ell}\}$ as the **length $\boldsymbol{\ell}$ eigenstates of $\boldsymbol{U}$**.

Remark 7.3.4. *The quantum Kraft Inequality proved in [53, Section IIC] is more general than the same proved in [13, Theorem 1], although the formalisms are quite different. Our version of the quantum Kraft Inequality, presented below, is a generalization of [13, Theorem 1], but is not quite in the full generality of [53, Section IIC] (in the forward direction) because we only consider uniquely decodable codes (as opposed to the more general notion, called condensable codes, considered in [53]). However, our version does have a converse statement, similar to the classical Kraft Inequality, which is missing from the aforementioned versions.*

Theorem 7.3.5. *(Quantum Kraft-McMillan Inequality) Any uniquely decodable quantum code U with length eigenstates over a binary alphabet must satisfy the inequality*

$$\mathrm{tr}(U^{\dagger} 2^{-\Lambda} U) \leq 1.$$

Conversely, if $U : H_{\mathcal{S}} \to H_{\mathcal{A}}^{\oplus}$ is a linear isometry with length eigenstates satisfying the above inequality, then there exists a c-q scheme $\tilde{U}$ with the same number of length ℓ eigenstates for each $\ell \in \mathbb{N}$.

Proof. Let U be a uniquely decodable quantum code with length eigenstates of the form

$$U = \sum_{i=1}^{d} |\psi_i\rangle\langle e_i|.$$

For each $n, N \in \mathbb{N}$, let

$$C_n^N = \{|\psi\rangle \in H_{\mathcal{A}}^{\otimes N} : |\psi\rangle = |\psi_{i_1}\rangle|\psi_{i_2}\rangle \cdots |\psi_{i_n}\rangle \text{ with } i_1, \ldots, i_N \in \{1, \ldots, d\}\}$$

be the collection of length N strings consisting of n codewords and let

$$d_\ell = |\{i \in \{1, \ldots, d\} : \psi_i \in H_{\mathcal{A}}^{\otimes \ell}\}|$$

be the number of length ℓ eigenstates of U, for each $\ell \in \mathbb{N}$. Then, by the unique decodability of U, each element of C_n^N has a unique representation as a string of n codewords and the elements of C_n^N are pairwise orthogonal, and hence we have

$$|C_n^N| = \sum_{\ell_1 + \cdots + \ell_n = N} d_{\ell_1} d_{\ell_2} \cdots d_{\ell_n} \leq 2^N.$$

Thus

$$2^{-N} \sum_{\ell_1 + \cdots + \ell_n = N} d_{\ell_1} d_{\ell_2} \cdots d_{\ell_n} = \sum_{\ell_1 + \cdots + \ell_n = N} (2^{-\ell_1} d_{\ell_1})(2^{-\ell_2} d_{\ell_2}) \cdots (2^{-\ell_n} d_{\ell_n}) \leq 1.$$

Since each ℓ_i is at most $\ell_{\max}$, where $\ell_{\max} = \max\limits_{1 \leq i \leq d}\{\ell_i\}$, we have $N \leq n\ell_{\max}$. So if we sum the above inequality over N we obtain

$$\sum_{\ell_1, \ell_2, \cdots, \ell_n = 1}^{\ell_{\max}} (2^{-\ell_1} d_{\ell_1})(2^{-\ell_2} d_{\ell_2}) \cdots (2^{-\ell_n} d_{\ell_n}) = \left(\sum_{\ell=1}^{\ell_{\max}} (2^{-\ell} d_\ell) \right)^n \leq n\ell_{\max}.$$

Notice that the left-hand side of this inequality is exponential whereas the right-hand side is linear. This implies that the left-hand side is bounded above by 1. Hence we must have that

$$\operatorname{tr}(U^\dagger 2^{-\Lambda} U) = \sum_{\ell=1}^{\ell_{\max}} 2^{-\ell} \operatorname{tr}(U^\dagger \Pi_\ell U) = \sum_{\ell=1}^{\ell_{\max}} 2^{-\ell} d_\ell \leq 1. \tag{7.19}$$

Notice that the inequality in Equation (7.19) is simply a restatement of the classical Kraft-McMillan Inequality. Conversely, suppose that U is a linear isometry with length eigenstates satisfying the quantum Kraft-McMillan Inequality. Then, by Equation (7.19), the classical Kraft-McMillan Inequality is also valid. Thus, by the converse of the classical Kraft-McMillan Theorem, one can find a classical uniquely decodable code C which has exactly d_ℓ-many codewords of length ℓ, for each $\ell \in \mathbb{N}$. The c-q scheme $\tilde{U}$ constructed from this classical code C has the desired properties. □

We would like to find a quantum code which minimizes the amount of resources required. Unfortunately there are numerous ways to define the length of a codeword for an indeterminate-length quantum code (e.g. base length [16], exponential length [13, Definition 6], etc.). Here, we define the **length of a codeword** $|\boldsymbol{\omega}\rangle$, where $|\omega\rangle = U|s\rangle \in H_{\mathcal{A}}^{\oplus}$, for a unique $|s\rangle \in \{|s_n\rangle\}_{n=1}^N$, as the expectation with respect to the length observable in Equation (7.17). Explicitly, the length of a codeword $|\omega\rangle = U|s\rangle$ will be given by a function $\ell : H_{\mathcal{A}}^{\oplus} \to \mathbb{R}^+$, defined as follows:

$$\ell(|\omega\rangle) := \langle \omega|\Lambda|\omega\rangle = \langle Us, \Lambda Us\rangle = \langle s, U^\dagger \Lambda Us\rangle = \sum_{i=1}^{d} |\langle e_i|s\rangle|^2 \ell_i. \tag{7.20}$$

For any ensemble $\mathcal{S} = \{p_n, |s_n\rangle\}_{n=1}^N$, define the **ensemble state** $\rho_{\mathcal{S}}$ of $\mathcal{S}$ by

$$\rho_{\mathcal{S}} = \sum_{n=1}^{N} p_n |s_n\rangle\langle s_n|. \tag{7.21}$$

If U is a quantum code on $H_{\mathcal{S}}$ define the **average codeword length** with respect to the ensemble $\mathcal{S}$ by

$$EL(U) = \operatorname{tr}(\rho_{\mathcal{S}} U^\dagger \Lambda U).$$

We denote by U_{opt} the **optimal quantum code** with length eigenstates for the ensemble $\mathcal{S}$ if

$$U_{\text{opt}} := \text{argmin}_U \{EL(U) : \text{tr}(U^\dagger 2^{-\Lambda} U) \leq 1\}, \tag{7.22}$$

The **optimal average codeword length** for the ensemble $\mathcal{S}$ is given by

$$EL^*(\rho_{\mathcal{S}}) := EL(U_{\text{opt}}) = \text{tr}(\rho_{\mathcal{S}} U_{\text{opt}}^\dagger \Lambda U_{\text{opt}}). \tag{7.23}$$

It is shown in [13, Theorem 2] that the optimal c-q scheme (and hence optimal quantum code with length eigenstates by the converse of Theorem 7.3.5) is given by the classical Huffman codes. The bounds on $EL^*(\rho_{\mathcal{S}})$ in terms of the von-Neumann entropy follow immediately.

Theorem 7.3.6. *The minimum average codeword length for an ensemble $\mathcal{S}$ is bounded as follows,*

$$S(\rho_{\mathcal{S}}) \leq EL^*(\rho_{\mathcal{S}}) < S(\rho_{\mathcal{S}}) + 1.$$

Proof. See [13, Theorem 3]. □

Next, we wish to consider the optimal average codeword length per symbol for a collection of ensembles $\{\mathcal{S}^k\}_{k=1}^\infty$, where $\mathcal{S}^k = \{p_{n_1,\ldots,n_k}, |s_1 s_2 \cdots s_k\rangle\}_{n_1,\ldots,n_k=1}^N$ and probabilities given by the pmf of a stochastic process $\mathbf{X}$. We will refer to such a collection of ensembles as a **stochastic ensemble**. Note that, by the definitions of a stochastic process, stochastic ensembles $\mathcal{S}^k$ must be compatible in the following sense:

$$\sum_{n_{k+1}=1}^{N} p_{n_1,\ldots,n_k,n_{k+1}} = p_{n_1,\ldots,n_k}, \tag{7.24}$$

for all $n_1, \ldots, n_k \in \{1, \ldots, N\}$ and $k \in \mathbb{N}$. Notice that we allow for the possibility that preparations of the ensemble at each time be dependent upon previous preparations. If the preparations of the ensemble are independent and identically prepared copies of an ensemble $\mathcal{S} = \{p_n, |s_n\rangle\}_{n=1}^N$; i.e. the stochastic process $\mathbf{X}$ is made up of i.i.d.

copies of a random variable X, then $p_{n_1,\ldots,n_k} = p_{n_1}p_{n_2}\cdots p_{n_k}$ and $\rho_{\mathcal{S}^k} = \rho^{\otimes k}$, where $\rho_{\mathcal{S}^k} = \sum_{n_1,\ldots,n_k=1}^{N} p_{n_1,\ldots,n_k}|s_{n_1}\cdots s_{n_k}\rangle$. For each $k \in \mathbb{N}$, let

$$EL_k^*(\rho_{\mathcal{S}^k}) = \frac{1}{k}EL_k^*(\rho_{\mathcal{S}^k}) \tag{7.25}$$

be the **optimal average codeword length per symbol** for the first k symbols with respect to the ensemble $\mathcal{S}^k$. Notice that the optimal average codeword length per symbol is defined analogously to the classical case in Equation (5.16). Then, from Theorem 7.3.6, we have

$$\frac{1}{k}S(\rho_{\mathcal{S}^k}) \leq EL_k^*(\rho_{\mathcal{S}^k}) < \frac{1}{k}S(\rho_{\mathcal{S}^k}) + \frac{1}{k}. \tag{7.26}$$

In the following subsection, we introduce an OQRW representation for a stochastic ensemble before applying the QMC dynamical entropy to the OQRW in Subsection 7.3.3 with the ultimate goal of extending Theorem 5.5.5.

7.3.2 An open quantum random walk associated with a stationary Markov ensemble

Consider a Markov process $\mathbf{X}$ with values in $\{x_n\}_{n=1}^N$ and with pmf $p_{\mathbf{X}}$. We will refer to the stochastic ensemble $\{\mathcal{S}^k\}_{k=1}^{\infty}$, with $\mathcal{S}^1 = \{p_{\mathbf{X}}(x_n), |s_n\rangle\}_{n=1}^N$ spanning $H_{\mathcal{S}} = H_{\mathcal{S}^1}$ and $\mathcal{S}^k = \{p_{\mathbf{X}}(x_{n_1},\ldots,x_{n_k}), |s_{n_1}\cdots s_{n_k}\rangle\}_{n_1,\ldots,n_k=1}^N$ spanning $H_{\mathcal{S}}^{\otimes k} = H_{\mathcal{S}^k}$ for each $k \in \mathbb{N}$, as the **Markov ensemble governed by X**. Whenever the Markov process is stationary or homogeneous we will refer to the Markov ensemble as being stationary or homogeneous, respectively.

Below we consider only stationary and homogeneous Markov ensembles. To this end, let $\mathbf{X}$ be a stationary, homogeneous Markov process governed by the transition matrix $P = (p_{n,m})$ and has initial distribution $p = \{p_n\}_{n=1}^N$. Recall that, since $\mathbf{X}$ is stationary, we have that p is invariant with respect to P. Moreover, the pmf, $p_{\mathbf{X}}$, of $\mathbf{X}$ is given by $p_{\mathbf{X}}(x_{n_1},\ldots,x_{n_k}) = p_{n_1}\prod_{l=2}^{k} p_{n_l,n_{l-1}}$, for each $k \in \mathbb{N}$ and $1 \leq n_1,\ldots,n_k \leq$

N. Setting $d = \dim(H_{\mathcal{S}})$, so that $d^k = \dim(H_{\mathcal{S}^k})$ for each $k \in \mathbb{N}$, we define the following sequence of states which represents this collection of ensembles by

$$\rho_{S^1} = \sum_{n=1}^{N} p_n |s_n\rangle\langle s_n| \in M_d = S_1(H_{\mathcal{S}^1}) \tag{7.27}$$

and, for each $k \in \mathbb{N}$ with $k \geq 2$,

$$\begin{aligned}\rho_{S^k} &= \sum_{n_1,\ldots,n_k=1}^{N} p_{n_1} \prod_{l=2}^{k} p_{n_l,n_{l-1}} |s_{n_1}\cdots s_{n_k}\rangle\langle s_{n_1}\cdots s_{n_k}| \in M_d^{\otimes k} = S_1(H_{\mathcal{S}^k}) \\ &= \sum_{n_1=1}^{N} p_{n_1}|s_{n_1}\rangle\langle s_{n_1}| \otimes \cdots \otimes \sum_{n_k=1}^{N} p_{n_k,n_{k-1}} |s_{n_k}\rangle\langle s_{n_k}|. \end{aligned} \tag{7.28}$$

For each $n \in \{1,\ldots,N\}$ we set $|s'_n\rangle = |s_n\rangle \otimes |n\rangle \in H_{\mathcal{S}} \otimes \mathbb{C}^N$ and consider the **quantum-classical (q-c) state**

$$\rho := \sum_{n=1}^{N} p_n |s'_n\rangle\langle s'_n| \in M_d \otimes M_N = S_1(H_{\mathcal{S}} \otimes \mathbb{C}^N). \tag{7.29}$$

Let $H_C = H_{\mathcal{S}}$ and $H_P = \mathbb{C}^N$, and define the OQRW over $H = H_C \otimes H_P$ by

$$\mathcal{M}(\cdot) := \sum_{m,n=1}^{N} M_{m,n} \cdot M_{m,n}^*, \tag{7.30}$$

where $M_{m,n} = \sqrt{p_{m,n}} U_{m,n} \otimes |m\rangle\langle n|$, and $U_{m,n}$ is any unitary operator on H_C satisfying $U_{m,n}|s_n\rangle = |s_m\rangle$, for all $m,n = 1,\ldots,N$. It is clear that

$$\sum_{m=1}^{N} p_{m,n} U_{m,n}^* U_{m,n} = \mathbb{1}_{H_C},$$

for each n, and hence $\mathcal{M}$ is an OQRW. Furthermore, notice that

$$\begin{aligned}\mathcal{M}(\rho) &= \sum_{m,n=1}^{N} M_{m,n}\Big(\sum_{k=1}^{N} p_k |s'_k\rangle\langle s'_k|\Big) M_{m,n}^* \\ &= \sum_{m,n=1}^{N} p_n p_{m,n} U_{m,n}|s_n\rangle\langle s_n| U_{m,n}^* \otimes |m\rangle\langle m| \\ &= \sum_{n=1}^{N} p_n |s'_n\rangle\langle s'_n| = \rho,\end{aligned}$$

where the second to last equality holds since p is P-invariant. Hence the q-c state ρ is an invariant state for $\mathcal{M}$. where the last equality holds since p is P-invariant. Hence the q-c state ρ is an invariant state for $\mathcal{M}$.

7.3.3 A quantum Markov chain representation for Markov ensembles

Consider the quantum dynamical system $(B(H_P), \Theta^*, \rho_0)$, with

$$\rho_0 = \mathrm{tr}_{H_C}(\rho) = \sum_{n=1}^{N} p_n |n\rangle\langle n|, \tag{7.31}$$

representative of the initial distribution of $\mathbf{X}$, and Θ^* the dual of the map Θ which satisfies the commutative diagram

$$\begin{tikzcd} S_1(H_P) \arrow[rr, "\Theta"] \arrow[d, "\lambda"] & & S_1(H_P) \\ S_1(H_C \otimes H_P) \arrow[rr, "\mathcal{M}"] & & S_1(H_C \otimes H_P) \arrow[u, "a"'] \end{tikzcd}$$

with

(i) $\lambda(|m\rangle\langle n|) = |s'_m\rangle\langle s'_n|$, for all $m, n \in \{1, \ldots, N\}$, and

(ii) $a(\cdot) = \mathrm{tr}_{H_C}(\cdot)$.

Remark 7.3.7. *The definition of Θ above is a slight modification of an optical communication process (see [46, Page 1202]). The main modification we have made is that we have essentially allowed for a different noise term, $|s_n\rangle\langle s_n|$, for each site $|n\rangle\langle n|$ of the position Hilbert space.*

Consider the spectral decomposition of $\rho_{\mathcal{S}^1}$,

$$\rho_{\mathcal{S}^1} = \sum_{i=1}^{d} \rho_i |\rho_i\rangle\langle\rho_i| \tag{7.32}$$

and fix an operational partition of unity $\gamma = (\gamma_i)_{i=1}^d$ of H_P, where

$$\gamma_i := \sum_{n=1}^{N} \langle\rho_i, s_n\rangle |n\rangle\langle n|, \quad \text{for all } i = 1, \ldots, d. \tag{7.33}$$

Notice that

$$\begin{aligned} \sum_{i=1}^{d} \gamma_i^* \gamma_i &= \sum_{i=1}^{d} \sum_{m,n=1}^{N} \langle\rho_i, s_m\rangle \overline{\langle\rho_i, s_n\rangle} |n\rangle\langle n|m\rangle\langle m| \\ &= \sum_{n=1}^{N} (\sum_{i=1}^{d} |\langle\rho_i, s_n\rangle|^2) |n\rangle\langle n| \end{aligned}$$

$$= \sum_{n=1}^{N} \|s_n\| |n\rangle\langle n| = \mathbb{1}_{H_P},$$

where the second to last equality follows by Parseval's identity, and hence γ is indeed an operational partition of unity. Let $E_\gamma : M_d \otimes M_N \to M_N$; i.e. $E_\gamma : B(H_C) \otimes B(H_P) \to B(H_P)$, be the transition expectation given by Equation (4.14) and let

$$E_{\gamma,\Theta^*} := \Theta^* \circ E_\gamma : B(H_C) \otimes B(H_P) \to B(H_P), \tag{7.34}$$

just as in Equation (4.15). Before proceeding with the construction of the quantum Markov chain, we give a technical lemma which will be helpful later.

Lemma 7.3.8. *Let $\{\mathcal{S}^k\}_{k=1}^{\infty}$ be a stationary, homogeneous Markov ensemble, with ensemble states $\{|s_n\rangle\}_{n=1}^{N}$, which is governed by a stationary, time-homogeneous Markov process $\mathbf{X}$ with transition matrix $P = (p_{n,m})$. Let Θ, γ, E_γ and E_{γ,Θ^*} be defined as above. Then the lifting $E^*_{\gamma,\Theta^*} : S_1(H_P) \to S_1(H_C) \otimes S_1(H_P)$ acts on the diagonal states of $S_1(H_P)$ in the following way.*

$$E^*_{\gamma,\Theta^*}(|n\rangle\langle n|) = \sum_{m=1}^{N} p_{m,n} |s'_m\rangle\langle s'_m|,$$

for each $|n\rangle$ in the orthonormal basis of H_P. Moreover,

$$E^*_{\gamma,\Theta^*}(\rho_0) = \rho.$$

Proof. First, recall that $E_{\gamma,\Theta^*} = \Theta^* \circ E_\gamma$ (by Equation (4.15)) so that $E^*_{\gamma,\Theta^*} = E^*_\gamma \circ \Theta$. Then, for each $|n\rangle\langle n| \in S_1(H_P)$, we have

$$\begin{aligned}
\Theta(|n\rangle\langle n|) &= a \circ \mathcal{M} \circ \lambda(|n\rangle\langle n|) \\
&= a \circ \mathcal{M}(|s'_n\rangle\langle s'_n|) \quad \text{with } |s'_n\rangle \text{ defined above Equation (7.29)} \\
&= a\Big(\sum_{m=1}^{N} p_{m,n} |s'_m\rangle\langle s'_m|\Big) \quad \text{by Equation (7.30)} \\
&= \sum_{m=1}^{N} p_{m,n} |m\rangle\langle m|. \qquad (7.35)
\end{aligned}$$

Next we consider the lifting $E^*_\gamma : S_1(H_P) \to S_1(H_C) \otimes S_1(H_P)$ which is given by the formula (see Equation (7.1))

$$E^*_\gamma(\sigma) = [\gamma_i \sigma \gamma_j^*]_{i,j=1}^d = \sum_{i,j=1}^d |\rho_i\rangle\langle\rho_j| \otimes \gamma_i\sigma\gamma_j^*,$$

where we have identified $S_1(H_C)$ with M_d and given the matrix representation with respect to the vectors $\{|\rho_i\rangle\}_{i=1}^d$ from Equation (7.32). Then, for each $|m\rangle\langle m| \in S_1(H_P)$, we have

$$\begin{aligned} E^*_\gamma(|m\rangle\langle m|) &= \sum_{i,j=1}^d |\rho_i\rangle\langle\rho_j| \otimes \gamma_i|m\rangle\langle m|\gamma_j^* \\ &= \sum_{i,j=1}^d |\rho_i\rangle\langle\rho_j| \otimes \langle\rho_i, s_m\rangle|m\rangle\langle m|\langle s_m, \rho_j\rangle \quad \text{by Equation (7.33)} \\ &= \left(\sum_{i=1}^d \langle\rho_i, s_m\rangle|\rho_i\rangle\right)\left(\sum_{j=1}^d \langle\rho_j, s_m\rangle|\rho_j\rangle\right)^* \otimes |m\rangle\langle m| \\ &= |s_m\rangle\langle s_m| \otimes |m\rangle\langle m| = |s'_m\rangle\langle s'_m|. \end{aligned} \tag{7.36}$$

Combining Equations (7.35) and (7.36), for each $|n\rangle\langle n| \in S_1(H_P)$, we have

$$\begin{aligned} E^*_{\gamma,\Theta^*}(|n\rangle\langle n|) &= E^*_\gamma(\sum_{m=1}^N p_{m,n}|m\rangle\langle m|) \quad \text{by Equation (7.35)} \\ &= \sum_{m=1}^N p_{m,n}|s'_m\rangle\langle s'_m| \quad \text{by Equation (7.36).} \end{aligned} \tag{7.37}$$

For the moreover statement, we have

$$\begin{aligned} E^*_{\gamma,\Theta^*}(\rho_0) &= \sum_{n=1}^N p_n E^*_{\gamma,\Theta^*}(|n\rangle\langle n|) \quad \text{by Equation (7.31)} \\ &= \sum_{n,m=1}^N p_n p_{m,n}|s'_m\rangle\langle s'_m| \quad \text{by Equation (7.37)} \\ &= \sum_{m=1}^N p_m|s'_m\rangle\langle s'_m| = \rho \quad \text{since } \mathbf{X} \text{ is stationary.} \end{aligned} \tag{7.38}$$

□

Next, we will consider the quantum Markov state ψ given by the chain $\{\rho_0, E_{\gamma,\Theta^*}\}$, where ρ_0 is given in Equation (7.31) and E_{γ,Θ^*} is as in Equation (7.34). Then, for each $k \in \mathbb{N}$ and $a_1, \ldots, a_k \in B(H_C) = M_d$, we have

$$\begin{aligned}
\psi(a_1 \otimes \cdots \otimes a_k) &= \operatorname{tr}(\rho_0 E_{\gamma,\Theta^*}(a_1 \otimes E_{\gamma,\Theta^*}(\cdots E_{\gamma,\Theta^*}(a_k \otimes \mathbb{1}_{H_P})))) \\
&= \operatorname{tr}(E^*_{\gamma,\Theta^*}(\rho_0)a_1 \otimes E_{\gamma,\Theta^*}(\cdots E_{\gamma,\Theta^*}(a_k \otimes \mathbb{1}_{H_P}))) \\
&= \operatorname{tr}(\sum_{n_1=1}^{N} p_{n_1}|s'_{n_1}\rangle\langle s'_{n_1}|a_1 \otimes E_{\gamma,\Theta^*}(\cdots E_{\gamma,\Theta^*}(a_k \otimes \mathbb{1}_{H_P}))) \\
&= \sum_{n_1=1}^{N} p_{n_1} \operatorname{tr}(|s_{n_1}\rangle\langle s_{n_1}|a_1) \operatorname{tr}(|n_1\rangle\langle n_1|E_{\gamma,\Theta^*}(a_2 \otimes \\
&\qquad\qquad E_{\gamma,\Theta^*}(\cdots E_{\gamma,\Theta^*}(a_k \otimes \mathbb{1}_{H_P})))) \\
&= \sum_{n_1,n_2=1}^{N} p_{n_1}p_{n_2,n_1} \operatorname{tr}(|s_{n_1}\rangle\langle s_{n_1}|a_1) \operatorname{tr}(|s_{n_2}\rangle\langle s_{n_2}|a_2) \times \\
&\qquad\qquad \operatorname{tr}(|n_2\rangle\langle n_2|E_{\gamma,\Theta^*}(\cdots E_{\gamma,\Theta^*}(a_k \otimes \mathbb{1}_{H_P}))) \\
&\;\;\vdots \\
&= \sum_{n_1,\ldots,n_k=1}^{N} p_{n_1} \prod_{l=2}^{k} p_{n_l,n_{l-1}} \prod_{l=1}^{k} \operatorname{tr}(|s_{n_l}\rangle\langle s_{n_l}|a_l),
\end{aligned}$$

where the "moreover" part of Lemma 7.3.8 was used in the 3$^{\text{rd}}$ equality, the fact $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$ was used in the 4$^{\text{th}}$ equality and Lemma 7.3.8 was used in the 5$^{\text{th}}$ equality.

Thus, for each $k \in \mathbb{N}$, the state ρ_k from Equation (4.17) is given by

$$\rho_k = \sum_{n_1,\ldots,n_k=1}^{N} p_{n_1} \prod_{l=2}^{k} p_{n_l,n_{l-1}} |s_{n_1} \cdots s_{n_k}\rangle\langle s_{n_1} \cdots s_{n_k}| = \rho_{\mathcal{S}^k}. \tag{7.39}$$

Therefore,

$$h(\Theta^*, \rho_0, \gamma) = \limsup_{k\to\infty} \frac{1}{k} S(\rho_k) = \limsup_{k\to\infty} \frac{1}{k} S(\rho_{\mathcal{S}^k}) = \limsup_{k\to\infty} EL^*_k(\rho_{\mathcal{S}^k}),$$

where the first equality holds by the definition of the dynamical entropy in Equation (7.3). We have proved the following theorem.

Theorem 7.3.9. *Given any stationary, homogeneous Markov ensemble $\{\mathcal{S}^k\}_{k=1}^{\infty}$, the optimal average codeword length per symbol converges to the dynamical entropy of the quantum dynamical system $(\Theta^*, \rho_0, \gamma)$, described above, in the following sense:*

$$\limsup_{k\to\infty} EL^*_k(\rho_{\mathcal{S}^k}) = \limsup_{k\to\infty} \frac{1}{k} S(\rho_{\mathcal{S}^k}) = h(\Theta^*, \rho_0, \gamma).$$

Corollary 7.3.10. *Given a Markov process* $\mathbf{X}$ *made up of i.i.d. copies of a random variable* X*, the stationary, homogeneous Markov ensemble* $\{\mathcal{S}^k\}_{k=1}^{\infty}$ *governed by* $\mathbf{X}$ *has optimal codeword length per symbol given by*

$$\lim_{k\to\infty} EL_k^*(\rho_{\mathcal{S}^k}) = S(\rho_{\mathcal{S}^1}).$$

Proof. First notice that $\mathbf{X}$ is governed by the transition matrix $P = (p_{n,m})_{n,m=1}^{N}$ such that $p_{n,m} = p_n$, for every $1 \leq n, m \leq N$, where $p = (p_n)_{n=1}^{N}$ is the initial distribution of $\mathbf{X}$. Therefore

$$\rho_{\mathcal{S}^k} = \rho_{\mathcal{S}^1}^{\otimes k}, \quad \text{for each } k \in \mathbb{N}.$$

Using the construction from above and Equation (7.26), we have that

$$S(\rho_k) = S(\rho_{\mathcal{S}^k}) = S(\rho_{\mathcal{S}^1}^{\otimes k}) = kS(\rho_{\mathcal{S}^1}),$$

where the last inequality follows by additivity of von Neumann entropy (see e.g. [65, Equation 2.8]). Therefore, by Theorem 7.3.9, we have

$$\lim_{k\to\infty} EL_k^*(\rho_{\mathcal{S}^k}) = \lim_{k\to\infty} \frac{1}{k} S(\rho_{\mathcal{S}^k}) = \lim_{k\to\infty} \frac{1}{k} kS(\rho_{\mathcal{S}^1}) = S(\rho_{\mathcal{S}^1}).$$

□

We finish this subsection by showing that the i.i.d. Markov ensemble consider in Corollary 7.3.10 has another representation in terms of AOW entropy. The following result is interesting in its own right, but we were unable to extend it to non-i.i.d. Markov ensembles.

Proposition 7.3.11. *Let* $\{\mathcal{S}^k\}_{k=1}^{\infty}$ *be a stationary, homogeneous Markov ensemble governed by a stochastic process* $\mathbf{X}$ *made up of i.i.d. copies of a random variable* X*. Then*

$$\lim_{k\to\infty} EL_k^*(\rho_{\mathcal{S}^k}) = S(\rho) = h^{AOW}(\Theta, \rho^{\otimes\mathbb{N}}, \gamma),$$

where $\rho = \rho_{\mathcal{S}^1}$*,* $\gamma = (|\rho_i\rangle\langle\rho_i|)_{i=1}^{d}$ *(with* $|\rho_i\rangle$*'s coming from the spectral decomposition of* ρ *in Equation* (7.32)*),* Θ *is the Bernoulli shift on the half-spin chain* $M_d^{\otimes\mathbb{N}}$*, and the AOW dynamical entropy is given in Equation* (7.11)*.*

Proof. Set $\rho = \rho_{\mathcal{S}^1}$. Then

$$\rho = \sum_{n=1}^{N} p_n |s_n\rangle\langle s_n| = \sum_{i=1}^{d} \rho_i |\rho_i\rangle\langle\rho_i|, \tag{7.40}$$

where the last equality is simply the spectral decomposition of ρ. Notice that, for each $k \in \mathbb{N}$,

$$\begin{aligned}
\rho_{\mathcal{S}^k} &= \sum_{n_1,\dots,n_k=1}^{N} p_{n_1}\cdots p_{n_k} |s_{n_1}\cdots s_{n_k}\rangle\langle s_{n_1}\cdots s_{n_k}| \\
&= \sum_{n_1=1}^{N} p_{n_1}|s_{n_1}\rangle\langle s_{n_1}| \otimes\cdots\otimes \sum_{n_k=1}^{N} p_{n_k}|s_{n_k}\rangle\langle s_{n_k}| \\
&= \sum_{i_1=1}^{d} \rho_{i_1}|\rho_{i_1}\rangle\langle\rho_{i_1}| \otimes\cdots\otimes \sum_{i_k=1}^{d} \rho_{i_k}|\rho_{i_k}\rangle\langle\rho_{i_k}| = \rho^{\otimes k}.
\end{aligned}$$

Let $\gamma = (\gamma_i)_{i=1}^{d}$ be the partition of unity with $\gamma_i = |\rho_i\rangle\langle\rho_i| \otimes \mathbb{1}^{[1,\infty)}$ for each $i \in \{1,\dots,d\}$, let Θ be the Bernoulli shift on the half-spin chain $M_d^{\otimes\mathbb{N}}$, and let $\psi = \{\rho^{\otimes\mathbb{N}}, \Theta\}$ be the corresponding quantum Markov state. Also, let $\mathbb{E}_\gamma$ be the transition expectation given in Remark 4.3.1 and let $E_{\gamma,\Theta} = \Theta\circ\mathbb{E}_\gamma$.

Then to calculate $h^{AOW}(\Theta, \rho^{\otimes\mathbb{N}}, \gamma)$, we need only find the diagonal entries of the joint correlations, ρ_k, by Equation (7.11). Noticing that $\rho^{\otimes\mathbb{N}}$ is invariant with respect to Θ^* and that Θ is a $*$-automorphism, we can use Equation (7.10) to get

$$\begin{aligned}
\rho_k(\bar{i},\bar{i}) &= \operatorname{tr}(\gamma_{i_k}\cdots\Theta^{(k-1)*}(\gamma_{i_1})\rho^{\otimes\mathbb{N}}\Theta^{(k-1)*}(\gamma_{i_1})\cdots\gamma_{i_k}) \\
&= \operatorname{tr}(\gamma_{i_k}\rho\gamma_{i_k}\otimes\cdots\otimes\gamma_{i_1}\rho\gamma_{i_1}\otimes\rho^{\otimes\mathbb{N}}) \\
&= \operatorname{tr}(\gamma_{i_1}\rho\gamma_{i_1})\operatorname{tr}(\gamma_{i_2}\rho\gamma_{i_2})\cdots\operatorname{tr}(\gamma_{i_k}\rho\gamma_{i_k}) \\
&= \rho_{i_1}\rho_{i_2}\cdots\rho_{i_k},
\end{aligned}$$

for each $k \in \mathbb{N}$ and $\bar{i} = (i_1,\dots,i_k) \in \{1,\dots,d\}^k$.

Therefore $\rho_k = \rho_{\mathcal{S}^k}$ and hence

$$S(\rho_k) = S(\rho_{\mathcal{S}^k}) = S(\rho^{\otimes k}) = kS(\rho), \tag{7.41}$$

for each $k \in \mathbb{N}$. Taking limits in Equations (7.26), (7.41) and (7.11), we see that

$$\lim_{k\to\infty} EL_k^*(\rho_{\mathcal{S}^k}) = S(\rho) = h^{AOW}(\Theta, \rho^{\otimes\mathbb{N}}, \gamma),$$

as desired. □

For a more detailed description of the AOW entropy of the dynamical system considered in Proposition 7.3.11 see [3, Section 4.1].

Remark 7.3.12. *The results of this section are also submitted for publication in [10].*

Bibliography

[1] L. Accardi. "The noncommutative Markovian property". In: *Func. Anal. Appl.* 9.1 (1975), pp. 1–7.

[2] L. Accardi and M. Ohya. "Compound channels, transition expectations, and liftings". In: *Appl. Math. Opt.* 39.1 (1999), pp. 33–59.

[3] L. Accardi, M. Ohya, and N. Watanabe. "Dynamical entropy through quantum Markov chains". In: *Open Syst. Inf. Dyn.* 4.1 (1997), pp. 71–87.

[4] L. Accardi, M. Ohya, and N. Watanabe. "Note on quantum dynamical entropies". In: *Rep. Math. Phys.* 38.3 (1996), pp. 457–469.

[5] Y. Aharonov, L. Davidovich, and N. Zagury. "Quantum Random Walks". In: *Phys. Rev. A* 48.2 (1993).

[6] R. Alicki and M. Fannes. "Defining Quantum Dynamical Entropy". In: *Lett. Math. Phys.* 64 (1994), pp. 75–82.

[7] A. Ambainis. "Quantum walks and their algorithmic applications". In: *Int. J. Quantum Inf.* 1.04 (2003), pp. 507–518.

[8] A. Ambainis et al. "One-dimensional Quantum Walks". In: *Proc. of the Thirty-third Annual ACM Symposium on Theory of Computing*. STOC '01. New York: ACM, 2001, pp. 37–49.

[9] G. Androulakis and D. Wright. "On the nonlinearity of quantum dynamical entropy". Submitted for publication in 2018.

[10] G. Androulakis and D. Wright. "Optimality in Quantum Data Compression using Dynamical Entropy". Submitted for publication in 2019.

[11] S. Attal. “Markov chains and dynamical systems: the open system point of view”. In: *arXiv preprint arXiv:1010.2894* (2010).

[12] S. Attal et al. “Open Quantum Random Walks”. In: *J. Stat. Phys.* 147 (2012), pp. 832–853.

[13] G. Bellomo et al. “Lossless quantum data compression with exponential penalization: an operational interpretation of the quantum Rényi entropy”. In: *Sci. Rep.* 7.1 (2017), p. 14765.

[14] F. Benatti. *Deterministic chaos in infinite quantum systems*. Springer Science & Business Media, 2012.

[15] L. Boltzmann. “Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie”. In: *Wiener Berichte* 53 (1866), pp. 195–220.

[16] K. Bostroem and T. Felbinger. “Lossless quantum data compression and variable-length coding”. In: *Phys. Rev. A* 65.3 (2002), p. 032313.

[17] S. L. Braunstein et al. “A quantum analog of Huffman coding”. In: *IEEE Trans. Inform. Theory* 46.4 (2000), pp. 1644–1649.

[18] A. Connes, H. Narnhofer, and W. Thirring. “Dynamical Entropy of C* Algebras and von Neumann Algebras”. In: *Comm. Math. Phys.* 112 (1987), pp. 691–719.

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

[20] E. B. Davies. *Quantum Theory of Open Systems*. London: Academic Press Inc., Nov. 1976.

[21] E. B. Davies and J. T. Lewis. “An operational approach to quantum probability”. In: *Comm. Math. Phys.* 17.3 (1970), pp. 239–260.

[22] T. Downarowicz. *Entropy in Dynamical Systems*. New York: Cambridge University Press, 2011.

[23] C. M. Edwards. "The algebraic approach to algebraic quantum theory I". In: *Comm. Math. Phys.* 16.3 (1970), pp. 207–230.

[24] T. Endo and N. Konno. "The stationary measure of a space-inhomogeneous quantum walk on the line". In: *arXiv preprint arXiv:1309.3054* (2014).

[25] G. L. Gerstein and B. Mandelbrot. "Random Walk Models for the Spike Activity of a Single Neuron". In: *Biophys. J.* 4 (1964), pp. 41–68.

[26] N. W. Goel and N. Richter-Dyn. *Stochastic Models in Biology*. Academic Press, 1974.

[27] L. Grady. "Random Walks for Image Segmentation". In: *IEEE Transaction on Pattern Analysis and Machine Intelligence* 28.11 (2006), pp. 1768–1783.

[28] L. K. Grover. "A fast quantum mechanical algorithm for database search". In: *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing*. Vol. 212. 1996.

[29] J. Kempe. "Quantum random walks: an introductory overview". In: *Contemporary Phys.* 44.4 (2003), pp. 307–327.

[30] V. Kendon. "A random walk approach to quantum algorithms". In: *Philos. T. Roy. Soc. A* 364.1849 (2006), pp. 3407–3422.

[31] B. P. Kitchens. *Symbolic Dynamics: One-sided, Two-sided, and Countable State Markov Shifts*. Springer, 1998. Chap. 7.

[32] A. N. Kolmogorov. "New Metric Invariant of Transitive Dynamical Systems and Endomorphisms of Lesbesgue Spaces". In: *Doklady of Russian Academy of Sciences* 119.5 (1958), pp. 861–864.

[33] A. N. Kolmogorov. "Three approaches to the quantitative definition of information". In: *Probl. Inf. Transm.* 1.1 (1965), pp. 1–7.

[34] N. Konno. “Quantum random walks in one dimension”. In: *Quantum Inf. Process.* 1.5 (2002), pp. 345–354.

[35] N. Konno. “Quantum Walks and Quantum Cellular Automata”. In: *International Conference on Cellular Automata for Research and Industry Proceedings.* 2008, pp. 12–21.

[36] N. Konno, T. Luczak, and E. Segawa. “Limit measures of inhomogeneous discrete-time quantum walks in one dimension”. In: *Quantum Inf. Process.* 12 (2013), pp. 33–53.

[37] N. Konno et al. “Quantum Walks and Reversible Cellular Automata”. In: *Phys. Lett. A* 330.4 (2004), pp. 408–417.

[38] A. Kossakowski, M. Ohya, and N. Watanabe. “Quantum Dynamical Entropy for Completely Positive Maps”. In: *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* 2.2 (1999), pp. 267–282.

[39] L. G. Kraft. “A device for quantizing, grouping, and coding amplitude-modulated pulses”. MA thesis. Massachusetts Institute of Technology, 1949.

[40] N. B. Lovett et al. “Universal quantum computation using the discrete-time quantum walk”. In: *Phys. Rev. A* 81.4 (2010), p. 042330.

[41] I. Martínez-Martínez and E. Sánchez-Burillo. “Quantum Stochastic Walks on Networks for Decision-Making”. In: *Sci. Rep.* 6.23812 (2016).

[42] B. McMillan. “Two inequalities implied by unique decipherability”. In: *IRE T. Inform. Theory* 2.4 (1956), pp. 115–116.

[43] D. A. Meyer. “From Quantum Cellular Automata to Quantum Lattice Gases”. In: *J. Stat. Phys.* 85.5-6 (1996), pp. 551–574.

[44] N. Muraki and M. Ohya. “Entropy functionals of Kolmogorov-Sinai type and their limit theorems”. In: *Lett. Math. Phys.* 36.3 (1996), pp. 327–335.

[45] N. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2000.

[46] M. Ohya and N. Watanabe. "Quantum entropy and its applications to quantum communication and statistical physics". In: *Entropy* 12.5 (2010), pp. 1194–1245.

[47] Henri Poincaré. *New methods of celestial mechanics*. Vol. 13. Springer Science & Business Media, 1992.

[48] R. Portugal. *Quantum Random Walks and Search Algorithms*. New York: Springer, 2013.

[49] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*. Vol. 10. John Wiley & Sons, 2016.

[50] M. Schuld, I. Sinayskiy, and F. Petruccione. "Quantum Walks on graphs representing the firing patterns of a quantum neural network". In: *Phys. Rev. A* 89.032333 (2014).

[51] M. Schuld, I. Sinayskiy, and F. Petruccione. "The quest for a Quantum Neural Network". In: *Quantum Inf. Process.* 13.11 (2014), pp. 2567–2586.

[52] B. Schumacher. "presentation at the Santa Fe Institute workshop on Complexity, Entropy, and the Physics of Information". In: *(unpublished)* (1994).

[53] B. Schumacher and M. D. Westmoreland. "Indeterminate-length quantum coding". In: *Phys. Rev. A* 64.4 (2001), p. 042304.

[54] C. E. Shannon. "A Mathematical Theory of Communication". In: *Bell System Tech. J.* 27.3 (July 1948), pp. 379–423.

[55] C. E. Shannon. "A Mathematical Theory of Communication". In: *Bell System Tech. J.* 27.4 (Oct. 1948), pp. 623–666.

[56] P. W. Shor. "Algorithms for quantum computation: Discrete logarithms and factoring". In: *Found. Comput. Sci., 1994 Proc., 35th Annu. Symp.* IEEE. 1994, pp. 124–134.

[57] Y. G. Sinai. "On the Notion of Entropy of a Dynamical System". In: *Doklady of Russian Academy of Sciences* 124 (1959), pp. 768–771.

[58] W. Słomczyński and A. Szczepanek. "Quantum Dynamical Entropy, Chaotic Unitaries and Complex Hadamard Matrices". In: *IEEE Trans. Inform. Theory* 63.12 (Dec. 2017), pp. 7821–7831.

[59] W. Słomczyński and K. Życzkowski. "Quantum Chaos: An Entropy Approach". In: *J. Math. Phys.* 35.2 (1994), pp. 5674–5700.

[60] J. M. Steele. *Stochastic calculus and financial applications*. Vol. 45. Springer Science & Business Media, 2012.

[61] P. Tuyls. "Comparing quantum dynamical entropies". In: *Banach Cent.* 43.1 (1998), pp. 411–420.

[62] P. Tuyls. "Towards Quantum Kolmogorov-Sinai Entropy". PhD thesis. KU Leuven Institute Theor. Phys., 1997.

[63] S. E. Venegas-Andraca. "Quantum Walks: A Comprehensive Overview". In: *Quantum Inf. Process.* 11.5 (2012), pp. 1015–1106.

[64] P. Walters. *An Introduction to Ergodic Theory*. New York: Springer-Verlag, 1982.

[65] A. Werhl. "General Properties of Entropy". In: *Rev. Modern Phys.* 50.2 (1978), pp. 221–260.