Measures of dynamical complexity

Andrei N. Soklakov

Technical Report

RHUL–MA–2001–1

31 January 2001

Department of Mathematics

Royal Holloway, University of London

Egham, Surrey TW20 0EX, England

http://www.rhul.ac.uk/mathematics/techreports

MEASURES OF DYNAMICAL COMPLEXITY

Andrei N. Soklakov

Royal Holloway

University of London

A thesis submitted for the degree of Doctor of Philosophy

in the Faculty of Science of the University of London, January 2001

Abstract

The aim of the thesis is to define, develop, and consider applications of different measures of dynamical complexity, i.e. the measures that would quantify complexity of system dynamics. These measures are based on the two fundamental notions of Kolmogorov (or algorithmic) complexity and von Neumann entropy.

Our main applications are in the theory of chaos and of open quantum systems. In such applications the interaction of the system with its environment is crucial. Consider a joint quantum state of a system and its environment. A measurement on the environment induces a decomposition of the system state. Using algorithmic information theory, we define the preparation information of a pure or mixed quantum state in a given decomposition. We demonstrate that the minimal value, $I_{\min}$, of the average preparation information of the system state characterizes the complexity of system-environment correlations which develop as a result of the system dynamics. Comparing the change of $I_{\min}$ with the change of the von Neumann entropy ΔH of the system induced by an optimal measurement we introduce a measure of complexity of the system dynamics ($\chi \equiv I_{\min}/\Delta H$). We discuss this measure of dynamical complexity in the context of the hypersensitivity approach to quantum chaos.

The partial development of a quantum version of symbolic dynamics for the quantum baker's map is one of the achievements presented in this thesis. Although our methods are not yet as powerful and general as the classical symbolic dynamics, we were able to recover the classical symbolic dynamics for the baker's map starting from a purely quantum version of the map and taking the classical limit. We use these results in the framework of the decoherent (consistent) histories approach to introduce a measure of dynamical complexity which is conceptually equivalent to the Kolmogorov-Sinai entropy which quantifies the degree of chaos in classical systems.

Often the mathematical formalism of algorithmic measures of complexity is very difficult to apply in a concrete physical setting. In such cases entropy-like measures of dynamical complexity can become the only practical choice. We consider a general setting of homodyne measurements in cavity QED. As our first objective we use the formalism of stochastic master equations to calculate the system entropy reduction due to the measurements. This quantity provides fundamental limits on the experimental resolution of the conditional system dynamics. We go beyond the limitations of the formalism of stochastic master equations end develop analytical tools for calculating the system dynamics conditional on the discrete photocount record.

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Rüdiger Schack for all his help, encouragement and guidance over the last three years.

I am very grateful to Carlton M. Caves, Christopher A. Fuchs, James B. Hartle, Jens G. Jensen, Alex S. Johnson, Yuri Kalnishkan, H. Jeff Kimble, Hideo Mabuchi and M. Saraceno for their support, encouragement, and numerous valuable discussions.

I would like to acknowledge the hospitality of the Isaac Newton Institute in Cambridge, the Department of Physics and Astronomy at the University of New Mexico and the Laboratory of Quantum Optics at the California Institute of Technology where various parts of this work were completed.

This work was supported by a research studentship from the Engineering and Physical Sciences Research Council, an Overseas Research Student award and by the Valerie Myerscough Prize 1999.

Contents

List of Figures

Introduction and outline

An incredible variety of different complexity arguments can be found in all areas of science. Often such arguments are used to guide our intuition as in biology [68, 35], chemistry [71] and environmental sciences [69, 52]. Sometimes, however, the notion of complexity plays a well-defined role as part of a mathematical formalism as in thermodynamics and the theory of chaos [5]. Our experience suggests that, although very simple in form, the fundamental laws of physics can lead to highly complex phenomena. The complexity often appears as an inevitable loss of predictability of the system evolution. How is this possible? How do we loose predictability in a situation where a perfectly reversible simple dynamical law governs the system dynamics? Several reasons can be mentioned to explain this fact. Some of them can be traced back to the fundamental principles of physics such as the concept of measurement in quantum mechanics and the Heisenberg uncertainty principle, which sets a limit on the accuracy of simultaneous measurements of conjugate observables. However, we do not need to look very deep to see that unpredictability is a very common phenomenon and can become an unavoidable problem, whether we have reached any fundamental limitations or not. Any situation where there is a rapid loss of control with time leads to unpredictability. Given a physical system we can quantify the resources necessary to simulate (or equivalently predict) the system evolution with some arbitrarily fixed precision. Such resources could for instance include the number of elementary operations needed to complete the computation, or the memory needed to assist it. In general, such resources depend on the physical interval of time Δt during which we simulate the evolution of the given system. Consider a situation in which the amount of

resources needed to predict the system evolution during Δt increases very rapidly with Δt, say exponentially or faster. Then no matter how much of the resources we have, at some point the prediction will suddenly become impossible: taking the calculations just a little further would require more resources then all of the previous computation. Physical systems which are difficult to predict in the above sense are called chaotic. One can understand this phenomenon in greater detail using the concept of hypersensitivity. If a system is hypersensitive to perturbations the knowledge of simple dynamical laws is not enough to predict the evolution of the system in the presence of an unknown environment. Every little perturbation should be recorded as it causes dramatic changes of the evolution. In this case, the necessary memory resources can grow exponentially in time which leads to unpredictability or apparent complexity of the system dynamics.

How do we measure unpredictability or, more generally, the complexity of system dynamics? In the information-theoretic approach the first step is to translate the system dynamics into a discrete form. As an example, one can think of a computer simulation of the dynamics. In this case the dynamics is simply a transformation of binary strings representing consecutive states of the system.

Among different information-theoretic measures of complexity at least three can be regarded as fundamental. These are Kolmogorov (or algorithmic) complexity, Shannon entropy and von Neumann entropy. Roughly speaking, the Kolmogorov complexity $K(\alpha)$ of a binary string α is the shortest length of a program which would cause a universal computer to compute α. In other words, $K(\alpha)$ is the smallest number of bits needed to describe α. For this reason Kolmogorov complexity is sometimes called descriptive complexity. Given a statistical ensemble of strings we can introduce an alternative measure of information known as Shannon information

$$S(\{p_k\}) \equiv -\sum_s p_s \log_2 p_s \ , \tag{0.1}$$

where p_s is the probability of the string s in the ensemble. This definition can be derived from a very natural set of requirements for an information measure as a function of the probabilities $\{p_k\}$ (Ref. [7], chapter 3). While Kolmogorov

complexity measures the complexity of an individual string, Shannon information measures the average complexity in a statistical ensemble of strings [39]. Shannon information is easier to deal with than Kolmogorov complexity which is not computable in the general case.

Consider a quantum system described by a density matrix ρ. Such a density matrix can always be decomposed as a mixture

$$\rho = \sum_s p_s \rho_s \ , \tag{0.2}$$

where $\{\rho_s\}$ is a set of pure states and $\{p_s\}$ is a set of probabilities. The von Neumann entropy $H(\rho)$ is defined as the minimum over all such decompositions of the Shannon information

$$H(\rho) \equiv \min_{\text{decompositions}} S(\{p_k\}) \ . \tag{0.3}$$

Like Kolmogorov complexity and the Shannon information the von Neumann entropy can be introduced independently starting from a set of general requirements [67]. Such arguments result in a simpler expression

$$H(\rho) = -\mathrm{Tr}(\rho \log_2 \rho) \ , \tag{0.4}$$

which is equivalent to (0.3).

In this thesis we will use the above definitions as building blocks for developing more specialized measures of dynamical complexity, i.e. the measures that quantify the complexity of system dynamics. Our main applications are in the theory of chaos and of open quantum systems. Interaction of such systems with the environment can result in highly complex dynamics. Given a quantum system interacting with some environment we consider a general situation when information about the system can be obtained via a fixed set of measurements performed on the environment. This situation is typical for experimental setups involving quantum-optical systems, where most of the observations are based on detection of photons which escape from the system into the surrounding environment. Mathematically, a measurement on the environment can be described as a decomposition of the system state into a sum of conditional states: each conditioned on a particular

outcome of the measurement. In Chapter 1 we use algorithmic information theory to define the preparation information of a pure or mixed quantum state in a given decomposition. We then define an optimal decomposition as a decomposition for which the average preparation information is minimal. This minimal value, $I_{\min}$, of the average preparation information of the system state characterizes the complexity of system-environment correlations which develop as a result of the system dynamics. Comparing the change of $I_{\min}$ with the change of the von Neumann entropy ΔH of the system induced by an optimal measurement we introduce a measure of complexity of the system dynamics as

$$\chi \equiv \frac{I_{\min}}{\Delta H} \ . \tag{0.5}$$

This measure of dynamical complexity plays a central role in the hypersensitivity approach to quantum chaos. Formally we can speak of a chaotic dynamics whenever $\chi \gg 1$. This condition formalizes the above explained situation when the memory resources $I_{\min}$ required to acquire ΔH bits of new information about the system are much larger than the amount of this new information. We give several examples revealing some important differences between classical and quantum chaos [60]. Although these examples are considered analytically they explicitely use a conjecture due to Schack and Caves which has only numerical justification [60, 61]. Roughly speaking, this conjecture implies that under the interaction with a random environment the state of a chaotic system becomes a mixture of pure states which are randomly distributed in the Hilbert space.

The above conjecture can be considered as a quantum analogue of the well known fact that the probability distribution describing the state of a classically chaotic system tends to spread over the entire phase space of the system under the influence of random perturbations at however small scale. This fact has a simple proof in terms of classical symbolic dynamics [60]. It would therefore be reasonable to anticipate some further development by looking for quantum generalizations of the classical symbolic dynamics. In the general case, this programme turned out to be rather difficult. There are however models, such as the classical baker's map, which have exceptionally simple symbolic dynamics. Partial development of

a quantum version of symbolic dynamics for the quantum baker's map is one of the achievements presented in Chapter 2 of this thesis. Although our methods are not yet as powerful and general as the classical symbolic dynamics, we were able to recover the classical symbolic dynamics for the baker's map starting from a purely quantum version of the map and taking the classical limit. We use these results in the framework of the decoherent (consistent) histories approach to introduce a measure of dynamical complexity which is conceptually related to the Kolmogorov-Sinai entropy which quantifies a degree of chaos in classical systems.

Often the standard mathematical formalism of algorithmic measures of complexity is very difficult to apply in a concrete physical setting. In such cases entropy-like measures of dynamical complexity can become the only practical choice. A good example is the net change of von Neumann entropy of the system resulting from both the interaction with the environment and from experiments conducted on the environment. In Chapter 3 we consider a general setting of homodyne measurements in cavity QED. Information about the system is obtained by performing a measurement on the output modes. Evolution of the system becomes conditional on the measurement records. As in Chapter 1, the set of possible measurement records defines the set of conditional states each conditioned on a particular measurement record. Physically, each conditional state describes the system given a particular history of the measurements. Dynamics of the conditional density matrix can be analyzed in the framework of stochastic master equations. As our first objective we calculate the system entropy reduction due to the measurements. This quantity is important as it provides fundamental limits on the experimental resolution of the conditional system dynamics. The formalism of stochastic master equations is sound if some assumptions on the capability of detectors are made. These assumptions include infinitesimally small response time without which the limit of continuous observations cannot be justified. We go beyond these limitations and analyze the system dynamics conditional on the discrete photocount record as opposed to the photocurrent. Considering the important case when the initial state of the system is the steady state we find an analytic expression

for the system density matrix during the evolution in the presence of continuous homodyne monitoring of the output modes.

Chapter 1

Preparation information for a quantum state as a measure of dynamical complexity

1.1 Introduction

The concept of dynamical complexity arises naturally in the context of chaos. Consider the time evolution of an open system in the presence of unknown environment. The uncertainty about the system state can be quantified by the Gibbs entropy in the classical case or by the von Neumann entropy in the quantum case [67]. In general, the entropy of the system state grows as a result of interaction with the environment. Measurements on the environment, however, can provide some information about the system state, i.e. can result in the decrease of entropy. How much information from the measurements is needed to keep the entropy below a certain "tolerable" level? More importantly, how is this information related to the corresponding decrease of entropy? Answers to these questions quantify the degree of unpredictability in the behavior of an open system which we summarize by a single quantity – the dynamical complexity.

Let $\Delta I_{\min}$ be the minimal amount of information about the environment which

is needed to decrease the system entropy by ΔH. The complexity of the system dynamics can be quantified by the ratio

$$\chi \sim \frac{\Delta I_{\min}}{\Delta H} \, . \tag{1.1}$$

An important feature of this quantity is that it can be introduced in an absolutely analogous way for both classical and quantum dynamics.

Before we convert (1.1) into a precise mathematical definition and discuss technical aspects of calculating χ it is relevant to mention our main motivation for considering this type of measure of dynamical complexity. Both classical and quantum systems can display complex (unpredictable) behavior. However, even when classical and quantum systems are closely related, it can be rather difficult to find a common reason for the loss of unpredictability. For instance, complexity or unpredictability of classical (Hamiltonian) dynamics is often described as complex or unpredictable in terms of sensitivity to initial conditions [40]. In such cases the accuracy of prediction of the phase-space trajectory will be necessarily getting worse if the accuracy of the initial conditions is finite. In order to maintain the accuracy of prediction at a constant level we would need extra information about the initial conditions supplied at a rate given by the Kolmogorov-Sinai entropy of the system (see section 2.2.3). By contrast, the angle (or Fubini Study distance [26]) between any two states in the Hilbert space of the system is preserved by the unitary evolution of quantum mechanics. This may suggest that dynamical unpredictability of quantum evolution can be avoided and is not fundamental. However, the Liouville equation in classical mechanics likewise preserves the overlap between phase-space distributions yet dynamical unpredictability of classical chaotic systems is an established fact. This argument (due to Berry [9]) suggests that there is a more fundamental signature of unpredictability than sensitivity to initial conditions. After all, the sensitivity to initial conditions is formalized using the essentially classical concept of trajectories and may not have a satisfactory analogue in the quantum case. Similarly, in quantum mechanics there is a source of unpredictability that has no analogue in the classical case. This is the unpredictability of quantum measurements which have little to do with dynamics. So

how do we quantify *dynamical* complexity of both classical and quantum systems without being distracted by differences in their formal descriptions? Clearly we need to find a source of unpredictability which is common in both classical and quantum dynamics. Perturbation of the evolution via interaction with an unknown environment is such a source. In this case of an open system the dynamical complexity can be quantified using the above measure χ. If $\chi \gg 1$ the system dynamics is called complex in the sense of hypersensitivity to perturbation and it is expected to be chaotic [64, 60, 61].

In this thesis we focus our attention on the complexity of quantum dynamics. The classical case was considered previously in full generality and we will use it as a guidance. Looking at the definition of χ we see that the complexity of the system dynamics depends strongly on the type of correlations that develop between the system and the environment as a result of their joint evolution. These are the correlations which allow information about the system to be accessible via the measurements performed on the environment. Such correlations can be studied via the conditional time evolution of the system conditioned on the measurement record. Fix any moment of time and consider the set of possible conditional states together with the probabilities of the corresponding measurement results. This set of states and corresponding probabilities defines a decomposition of the unconditional system density matrix at the given time. In section 1.2 we will see how studying such decompositions can be used to define and calculate χ. It should however be mentioned in advance that calculation of χ can become very complicated for interesting physical systems. At this point observations taken from numerical simulations of chaotic systems can be very helpful.

It has been noticed by Schack, D'Ariano and Caves [64, 61] that interaction of a chaotic system with a random (unknown) environment seems to result typically in ensembles of conditional states which are distributed uniformly over the entire Hilbert space of the system. Although lacking an analytical prove, this observation has been confirmed by extensive numerical simulations where the interaction with the environment was modeled as a stochastic perturbation of the Hamilto-

nian. We use this observation to calculate the dynamical complexity in a number of simple examples. We shall see that, whenever valid, this observation enables quantum systems to exhibit exponential hypersensitivity to perturbation similar to the classical case [58, 60].

Using the method of symbolic dynamics Schack and Caves were able to define χ rigorously and estimate it for a large family of classical chaotic maps [58, 60]. Unlike the quantum case the estimations were performed analytically due to the universality of the classical method of symbolic dynamics. It would therefore be reasonable to anticipate some further insight into the quantum origins of hypersensitivity by looking for quantum generalizations of the classical symbolic dynamics (see section 2.2 for an overview). This links the material of this chapter with Chapter 2 of this thesis where we develop a quantum version of symbolic dynamics using the example of the quantum baker's map. In the general quantum case, however, the symbolic analysis of χ can become rather difficult. This suggests that other measures of dynamical complexity should be introduced depending on the context of a particular application.

In chapters 2 and 3 we give more specialized definitions of dynamical complexity for a system with simple symbolic dynamics and for a very general experimental setup in cavity QED. An essential feature of the method of symbolic dynamics is coarse graining of the state space of the system (see sec 2.2.2). The problem of quantum coarse-graining has been discussed in the framework of the decoherent (consistent) histories approach [28, 46, 24] which links quantum and classical physics. Using this approach we introduce a measure of dynamical complexity which, in the classical limit, is related to the Kolmogorov-Sinai entropy. We use the quantum baker's map to give a detailed analysis. Although being an important example for studying quantum and classical chaos, the baker's map is an artificial system. What are we going to do with real systems, say an atom in a cavity coupled to the environment, for which finding the symbolic dynamics is a difficult open problem? It has been suggested by Zurek and Paz [75] that the rate of von Neumann entropy production can already be a good indication of the degree

of chaos (and therefore dynamical complexity) for an open quantum system. In Chapter 3 we consider such a measure of dynamical complexity in the context of a realistic experimental setup in cavity QED.

1.2 Preparation information and optimal ensembles

Let D and $D_{\mathcal{E}}$ denote the Hilbert-space dimensions of the system $\mathcal{S}$ and the environment $\mathcal{E}$, respectively. We will normally assume that $D_{\mathcal{E}} \gg D$. Now consider a joint state ρ_{total} on $\mathcal{S} \otimes \mathcal{E}$. The state of the system alone, ρ, is then obtained by tracing out the environment,

$$\rho = \text{tr}_{\mathcal{E}}(\rho_{\text{total}}) \; . \tag{1.2}$$

The von Neumann entropy of the system is

$$H = -\text{tr}(\rho \log \rho) \; , \tag{1.3}$$

where here and throughout this chapter, log denotes the base-2 logarithm. We now perform an arbitrary measurement on the environment [36], described by a POVM, $\{E_r\}$, where the E_r are positive environment operators such that

$$\sum_r E_r = \mathbb{1}_{\mathcal{E}} = (\text{environment unit operator}). \tag{1.4}$$

The probability of obtaining result r is given by

$$p_r = \text{tr}(\rho_{\text{total}} E_r) \; , \tag{1.5}$$

and the system state after a measurement that yields result r is

$$\rho_r = \frac{\text{tr}_{\mathcal{E}}(\rho_{\text{total}} E_r)}{p_r} \; . \tag{1.6}$$

By summing over r and using the completeness of the POVM, we obtain

$$\sum_r p_r \rho_r = \text{tr}_{\mathcal{E}}(\rho_{\text{total}}) = \rho \; . \tag{1.7}$$

The ensemble $\{\rho_r, p_r\}$ forms a decomposition of ρ. To characterize the ensemble, we define the system entropy conditional on measurement outcome r,

$$H_r = -\mathrm{tr}(\rho_r \log \rho_r) \ , \tag{1.8}$$

the average conditional entropy,

$$\bar{H} = \sum_r p_r H_r \ , \tag{1.9}$$

and the average entropy decrease due to the measurement, $\Delta\bar{H} = H - \bar{H}$. These quantities obey the double inequality

$$0 \leq \Delta\bar{H} \leq H \ , \tag{1.10}$$

which follows from the concavity of the von Neumann entropy [7, 15]. The content of the first of these inequalities is that a measurement on the environment will not, on the average, increase the system entropy.

Now let $\{\rho_r, p_r\}$ be the ensemble induced by the POVM $\{E_r\}$. We denote by $I(\rho_k|\rho_{\mathrm{total}}, \{E_r\})$ the conditional algorithmic information to specify ρ_k, given the ensemble (see [16] and references in [56]). The quantity $I(\rho_k|\rho_{\mathrm{total}}, \{E_r\})$ defines the *preparation information* of the state ρ_k, given the total state ρ_{total} and the POVM. We also define the *average preparation information*

$$\bar{I}(\rho_{\mathrm{total}}, \{E_r\}) \equiv -\sum_r p_r \log p_r \ . \tag{1.11}$$

This definition is justified, because the average algorithmic information can be bounded above and below as follows: [56]

$$-\sum_r p_r \log p_r \leq \sum_k p_k I(\rho_k|\rho_{\mathrm{total}}, \{E_r\}) \leq -\sum_r p_r \log p_r + 1 \ . \tag{1.12}$$

The average preparation information is never smaller than the average system entropy decrease $\Delta\bar{H}$,

$$\bar{I}(\rho_{\mathrm{total}}, \{E_r\}) \geq \Delta\bar{H} \ . \tag{1.13}$$

This inequality is a consequence of a general theorem about average density operators [7, 15].

The next step is to define a ΔH-*decomposition* of ρ as a decomposition for which $\Delta\bar{H} \geq \Delta H$, and an *optimal* ΔH-*decomposition* of ρ as a ΔH-decomposition with minimal average preparation information $\bar{I}$. The average preparation information for an optimal ΔH-decomposition,

$$\bar{I}_{\min} = \inf_{\Delta H-\text{decompositions}} \bar{I} \,, \tag{1.14}$$

is then a property of ρ_{total}, and characterizes the system-environment correlations. (If there is no ΔH decomposition for which $\bar{I}$ is minimal, we will call any decomposition optimal for which $\bar{I} < \bar{I}_{\min} + \epsilon$ for some given small constant ϵ.) The quantity $\bar{I}_{\min}$ is the information about the environment needed to reduce the system entropy by ΔH. Now we can define χ rigorously as

$$\chi \equiv \frac{\bar{I}_{\min}}{\Delta H} \,. \tag{1.15}$$

A useful generalization results from taking the infinum in Eq. (1.14) over a restricted class of POVMs, as in the quantum-optical example of Ref. [11]. This defines ensembles that are optimal with respect to a given class of environment measurements.

To illustrate the above introduced concepts we conclude this section with a couple of simple examples. In section 1.4 we consider more complicated examples illustrating how the above measure of dynamical complexity distinguishes between regular and chaotic dynamics.

Let the system be a qubit, for which the dimension of Hilbert space is $D = 2$. Let $|0\rangle$ and $|1\rangle$ be orthogonal basis states for the qubit. Define $|\psi_1\rangle = |0\rangle$, $|\psi_2\rangle = |1\rangle$, $|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|\psi_4\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, and let the joint density operator of system and environment be

$$\rho_{\text{total}} = \frac{1}{4}\sum_{k=1}^{4} |\psi_k\rangle\langle\psi_k| \otimes P_k^{\mathcal{E}} \,, \tag{1.16}$$

where $\{P_k^{\mathcal{E}}\}$ are projectors on the Hilbert space of the environment. The state of the system alone is then given by

$$\rho = \text{tr}_{\mathcal{E}}(\rho_{\text{total}}) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \,, \tag{1.17}$$

for which the system entropy in the absence of measurements is given by $H = 1$. Suppose we want to reduce the system entropy by $\Delta H = 1$bit, i.e., we want the conditional system state to be pure. The only environment measurement achieving this is given by $E_r = P_r^{\mathcal{E}}$, which results in the unique and therefore optimal $\Delta H = 1$ ensemble given by $\rho_r = |\psi_r\rangle\langle\psi_r|$, $p_r = 1/4$. The average preparation information for this ensemble is $\bar{I} = 2$bits, and hence $\bar{I}_{\min} = 2$bits.

For a different value of ΔH, consider the ensemble defined by $\rho_1 = \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_3\rangle\langle\psi_3|)$, $\rho_2 = \frac{1}{2}(|\psi_2\rangle\langle\psi_2| + |\psi_4\rangle\langle\psi_4|)$, $p_1 = p_2 = 1/2$, which is induced by the POVM $E_1 = P_1^{\mathcal{E}} + P_3^{\mathcal{E}}$ and $E_2 = P_2^{\mathcal{E}} + P_4^{\mathcal{E}}$. For this ensemble, $H_1 = H_2 = -\mathrm{tr}(\rho_1 \log \rho_1) \simeq 0.81$, and hence $\Delta\bar{H} = H - H_1 \simeq 0.19$. The average preparation information is $\bar{I} = 1$. It is easy to see that, with respect to our restricted class of measurements, this is an optimal $\Delta\bar{H}$-decomposition, and hence $\bar{I}_{\min} = 1$. In this example, to obtain 0.19 bits of information about the system, 1 bit of information about the environment is needed.

1.3 Hypersensitivity criterion

Following [61] we assume that the interaction with the environment can be described as a stochastic perturbation of the Hamiltonian. Let the system be prepared in a pure state and let N be the number of different perturbation histories which can happen during time interval Δt. During this time the perturbed evolution brings the system in one of the states $(|\psi_1\rangle, \ldots, |\psi_N\rangle)$ with corresponding probabilities $(p_1, \ldots, p_N)$. Averaging over the perturbation leads to the system density operator

$$\rho_{\mathcal{S}} = \sum_{k=1}^{N} p_k |\psi_k\rangle\langle\psi_k| \,. \tag{1.18}$$

We will call this density matrix *unconditional* meaning that it is not conditioned on any measurement record. We already mentioned that the measure of dynamical complexity χ introduced above characterizes system-environment correlations which develop during the perturbed time evolution of the system. To study these

correlations we need to know the joint system-environment density matrix. Having restricted ourselves by modeling the system-environment interaction by a stochastic Hamiltonian, we can assume without further loss of generality that the joint state of system and its environment can be described by a density matrix of the form

$$\rho_{\text{total}} = \frac{1}{D_{\mathcal{E}}} \sum_{k=1}^{N} |\psi_k\rangle\langle\psi_k| \otimes P_k^{\mathcal{E}} , \tag{1.19}$$

where $\{P_k^{\mathcal{E}}\}$ is a complete set of mutually orthogonal projectors on the $D_{\mathcal{E}}$-dimensional Hilbert space of the environment. This equation states in an explicit form that there exists a unique history of perturbations of the system by the environment. Each state $|\psi_k\rangle$ in (1.19) is a conditional state of the system conditioned on a particular perturbation history which is realized with probability $p_k = \text{Tr}(P_k^{\mathcal{E}})/D_{\mathcal{E}}$. This state is classically correlated with some state of the environment. This type of correlations naturally arise in periodically perturbed systems such as the quantum kicked top [61].[1]

Schack and Caves conjectured [61, 60] that the distribution of the vectors $\{|\psi_k\rangle\}$ holds the key to the complexity of system dynamics. They noticed that in the chaotic regime the time evolution of the system often results in a random distribution of the vectors $\{|\psi_k\rangle\}$ over the entire Hilbert space of the system. They also argued that regular evolution can only explore a relatively small submanifold of the system Hilbert space. In the next section we consider two examples corresponding to chaotic and regular dynamics.

1.4 Examples

In this section, $\{P_k^{\mathcal{E}}, k = 1, \ldots, D_{\mathcal{E}}\}$ denotes a complete set of orthogonal environment projectors. In the two examples discussed below, we will restrict the class of

[1] Here we do not aim for generality. In the most general case the system would become entangled with the environment and the system-environment interaction could not be described by a unique perturbation history.

environment measurements to orthogonal projections of the form

$$E_r = \sum_{k \in K_r} P_k^{\mathcal{E}} \ , \tag{1.20}$$

where $K_r \subset \{1, \ldots, D_{\mathcal{E}}\}$. For the given examples, it seems intuitively clear that ensembles which are optimal with respect to this class of measurements are also, to a good approximation, optimal with respect to the class of all possible environment measurements. We have not, however, been able to prove this statement rigorously.

1.4.1 Random vectors in Hilbert space

In the trivial example considered at the end of section 1.2, the average preparation information $\bar{I}_{\min}$ for an optimal ensemble is significantly larger than the corresponding entropy reduction $\Delta\bar{H}$. In the present subsection, we show that $\bar{I}_{\min}$ can vastly exceed $\Delta\bar{H}$.

Assume that $\log D_{\mathcal{E}} \gg \log D$ and consider

$$\rho_{\text{total}} = \frac{1}{D_{\mathcal{E}}} \sum_{k=1}^{D_{\mathcal{E}}} |\psi_k\rangle\langle\psi_k| \otimes P_k^{\mathcal{E}} \ , \tag{1.21}$$

where the $|\psi_k\rangle$ are distributed randomly in D-dimensional (projective) Hilbert space [74]. Here, the system entropy in the absence of measurements is $H \simeq \log D$. We will see that in this case the complexity of the system-environment correlations, as quantified by χ, is very large. This is in marked contrast to the second example discussed below.

Environment measurements of the form (1.20) correspond to grouping the vectors $|\psi_k\rangle$ into disjoint groups. We construct an approximation to an optimal measurement by grouping the vectors into Hilbert-space spheres of radius ϕ. (See Ref. [61] for a detailed argument.) We assume that $D_{\mathcal{E}}$ is sufficiently large so that the state vectors in each such sphere fill it randomly. Since all spheres are chosen to be of equal size, the average entropy $\bar{H}$ is equal to the entropy of one sphere, i.e., the entropy of a uniform mixture of states within a Hilbert-space sphere of radius ϕ,

given by [64]

$$\bar{H} = -\left(1 - \frac{D-1}{D}\sin^2\phi\right)\log\left(1 - \frac{D-1}{D}\sin^2\phi\right) - \frac{D-1}{D}\sin^2\phi\,\log\left(\frac{\sin^2\phi}{D}\right) . \tag{1.22}$$

The volume contained within a sphere of radius ϕ in Hilbert space is $(\sin\phi)^{2(D-1)}V_D$, where V_D is the total volume of projective Hilbert space [64]. The number of spheres of radius ϕ in D-dimensional Hilbert space is thus $(\sin\phi)^{-2(D-1)}$, so the information needed to specify a particular sphere is

$$\bar{I}_{\min} \simeq \tilde{I}_{\min} \equiv \log\left((\sin\phi)^{-2(D-1)}\right) = -(D-1)\log(\sin^2\phi) . \tag{1.23}$$

The information $\tilde{I}_{\min}$ slightly underestimates the actual value of $\bar{I}_{\min}$, because the perfect grouping into nonoverlapping spheres of the same size assumed by Eq. (1.23) does not exist.

As an example, let us choose a Hilbert-space dimension $D = 101$ and a radius of $\phi \simeq 1.025$. Equations (1.22,1.23) then give $\Delta\bar{H} = \log D - \bar{H} \simeq 1$ and $\tilde{I}_{\min} \simeq 45.3$, which means that here, to obtain 1 bit of information about the system, more than 45 bits of information about the environment are needed.

Using Eq. (1.23) to eliminate ϕ from Eq. (1.22) gives a complicated expression for $\tilde{I}_{\min}$ as a function of $\Delta\bar{H}$ [62], which is a good approximation to the average preparation information for an optimal $\Delta\bar{H}$ ensemble. Figure 1.1 shows this function for a Hilbert space dimension $D = 101$. To obtain more insight into the properties of this curve, we consider the derivative [62]

$$\frac{d\tilde{I}_{\min}}{d\Delta\bar{H}} = \frac{D}{\sin^2\phi\ln(1 + D\cot^2\phi)} , \tag{1.24}$$

which is the marginal tradeoff between between information and entropy. For ϕ near $\pi/2$, so that $\epsilon = \pi/2 - \phi \ll 1$, the information becomes $\tilde{I}_{\min} = (D-1)\epsilon^2/\ln 2$, and the derivative (1.24) can be written as

$$\frac{d\tilde{I}_{\min}}{d\Delta\bar{H}} \simeq \frac{D}{\ln(1 + D\epsilon^2)} , \tag{1.25}$$

which is proportional to D with a slowly varying logarithmic correction. We have thus identified a situation where the average preparation information is of the same

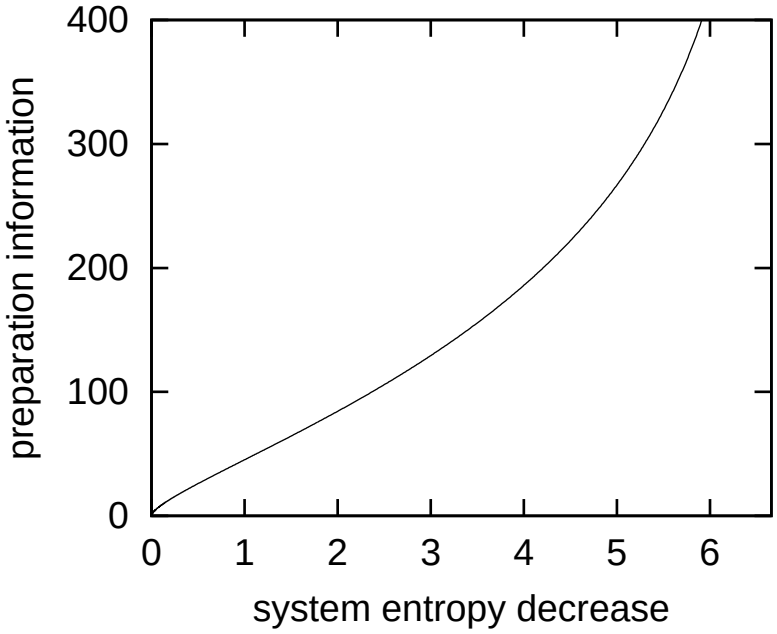


Figure 1.1: Average preparation information $\tilde{I}_{\min}$ in bits, versus average entropy reduction $\Delta\bar{H}$ in bits, for the example of subsection 1.4.1.

order as the dimension of Hilbert space D, despite the fact that the von Neumann entropy of a state cannot exceed $\log D$.

1.4.2 Random coherent states

In this example the system considered is a spin-j particle, for which the dimension of Hilbert space is $D = 2j + 1$. As in the preceding section, we assume that the Hilbert-space dimension of the environment is much larger than D, $D_{\mathcal{E}} \gg D$, and consider

$$\rho_{\text{total}} = \frac{1}{D_{\mathcal{E}}} \sum_{k=1}^{D_{\mathcal{E}}} |\psi_k\rangle\langle\psi_k| \otimes P_k^{\mathcal{E}} \ , \tag{1.26}$$

but now we choose the $|\psi_k\rangle$ to be distributed randomly on the submanifold of angular-momentum coherent states (1.28). We will see that the resulting complexity of the system-environment correlations, as quantified by χ, is small.

The angular momentum coherent state $|\theta, \phi\rangle$ can be defined by rotating the $\hat{J}_z$ eigenstate $|j; j\rangle$ through Euler angles ϕ around the z-axis, and then by θ around the new y-axis. This gives [49]

$$|\theta, \phi\rangle = \sum_{m=-j}^{j} |j; m\rangle \binom{2j}{j+m}^{\frac{1}{2}} \cos^{j+m}(\theta/2) \sin^{j-m}(\theta/2) e^{-im\phi} \ . \tag{1.27}$$

Each coherent states corresponds to a point on the surface of a three-dimensional sphere. Assuming that $D_{\mathcal{E}}$ is sufficiently large the state of the system alone is

$$\hat{\rho} = \text{tr}_{\mathcal{E}}(\rho_{\text{total}}) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} |\theta, \phi\rangle\langle\theta, \phi| \sin\theta \, d\theta \, d\phi \ . \tag{1.28}$$

As in the previous section, environment measurements of the form (1.20) correspond to grouping the vectors $|\psi_k\rangle$ into disjoint groups. Approximately optimal measurements correspond to grouping the vectors into approximately equal, compact areas on the surface of the sphere. We choose the areas to be of the form

$$\Omega_r(\Theta) = \{\theta, \phi : \arccos[\underline{\text{n}}(\theta, \phi)\underline{\text{n}}(\theta_r, \phi_r)] \leq \Theta\} \tag{1.29}$$

centered at points (θ_r, ϕ_r). The corresponding density operators

$$\hat{\rho}_r(\Theta) = \frac{\int_{\Omega_r(\Theta)} |\theta, \phi\rangle\langle\theta, \phi| \sin\theta \, d\theta \, d\phi}{\int_{\Omega_r(\Theta)} \sin\theta \, d\theta \, d\phi} \ , \tag{1.30}$$

can be used to construct a nearly optimal decomposition of $\hat{\rho}$. The preparation information $\bar{I}_{\min}$ is then approximately given by

$$\bar{I}_{\min} \simeq \tilde{I}_{\min} \equiv \log \frac{4\pi}{2\pi(1-\cos\Theta)} , \tag{1.31}$$

where the denominator is the area of $\Omega_r(\Theta)$

In the following section, we show that $\hat{\rho}_r(\Theta)$, in the coordinates where $(\theta_r, \phi_r) = (0,0)$, can be written in the diagonal form

$$\hat{\rho}_r(\Theta) = \sum_{m=-j}^{j} |j;m\rangle \lambda_m^{\Theta} \langle j;m| , \tag{1.32}$$

where

$$\lambda_m^{\Theta} = \frac{(2j)!\ \sin^{2(j-m)}\frac{\Theta}{2}}{(j+m)!\ (j-m+1)!} F(-j-m, j-m+1, j-m+2, \sin^2\frac{\Theta}{2}) , \tag{1.33}$$

and where F is the hypergeometric function ${}_2F_1$. Since all density operators $\rho_r(\Theta)$ in the decomposition of $\hat{\rho}$ have the same entropy, the average entropy $\bar{H}$ can be written as

$$\bar{H} = -\sum_{m=-j}^{j} \lambda_m^{\Theta} \log \lambda_m^{\Theta} . \tag{1.34}$$

For the entropy of the system, Eq. (1.3), in the absence of measurements, $\hat{\rho} = \hat{\rho}_r(\pi)$, we have (see Eq. 1.48)

$$H = -\sum_{m=-j}^{j} \lambda_m^{\pi} \log \lambda_m^{\pi} = \log(2j+1) = \log D . \tag{1.35}$$

We can analyze $\bar{H}$ in detail for the important special value $\Theta = \pi/2$, for which the measurement corresponds to a grouping into two disjoint hemispheres, and is therefore strictly optimal. For such a grouping, $\bar{I}_{\min} = 1$. In the next section, we show that the eigenvalues $\lambda_m^{\pi/2}$ obey the bounds

$$\begin{aligned} 0 \le \lambda_m^{\pi/2} \le \frac{1}{j} e^{-\sqrt[3]{j}/3} , \qquad & m < -1 - j^{2/3} , \\ \frac{2}{2j+1}(1 - e^{-\sqrt[3]{j}/3} - 4^{-j}) \le \lambda_m^{\pi/2} \le \frac{1}{j} , \qquad & m > 1 + j^{2/3} . \end{aligned} \tag{1.36}$$

Using these bounds we have derived the following asymptotic expression for the average entropy:

$$\bar{H} = -\sum_{m=-j}^{j} \lambda_m^{\pi/2} \log \lambda_m^{\pi/2} = \log j + O(\sqrt[3]{j} e^{-\sqrt[3]{j}/3}) , \tag{1.37}$$

and hence, in the limit $j \to \infty$,

$$\frac{\bar{I}_{\min}}{\Delta\bar{H}} = \frac{1}{\log(2j+1) - \bar{H}} \longrightarrow 1 \ . \tag{1.38}$$

Figure 1.2 shows a parametric plot of the average preparation information $\tilde{I}_{\min}$ versus the average entropy reduction $\Delta\bar{H} = H - \bar{H}$ for $j = 50$, i.e., D=101. It can be seen that for moderate values of $\Delta\bar{H}$, $\tilde{I}_{\min} \simeq \Delta\bar{H}$. To reduce the system entropy by 1 bit, not more than approximately 1 bit of information about the environment is needed. This should be compared to the previous example, where the required environment information is of the same order as D. In the limit of $\Delta\bar{H}$ approaching its maximum value $H = \log D$, the information $\tilde{I}_{\min}$ diverges. This is due to the fact that an infinite amount of information is needed to specify a general state exactly. The complexity of the system-environment correlations is characterized by the slope of the curve for small values of $\Delta\bar{H}$ rather than its asymptotic behaviour for $\Delta\bar{H} \to \log D$.

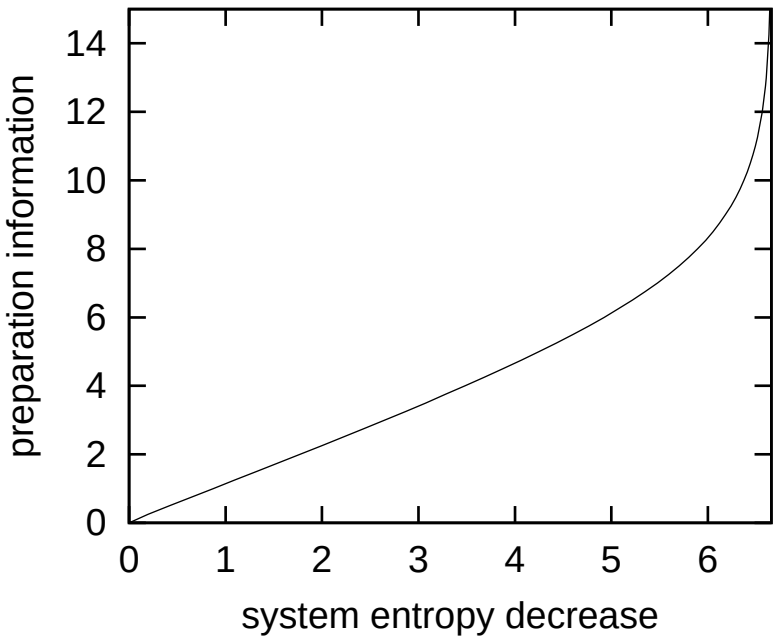


Figure 1.2: Average preparation information $\tilde{I}_{\min}$ in bits, versus average entropy reduction $\Delta\bar{H}$ in bits, for the example of subsection 1.4.2.

1.4.3 Mathematical details

Here we provide rigorous proofs for the results used in section 1.4.2. Namely, we calculate the eigenvalues of $\hat{\rho}_r(\Theta)$ given by Eq. (1.33) and derive the expression (1.36) for the case $\Theta = \pi/2$. Choosing the coordinate system such that $(\theta_r, \phi_r) = (0, 0)$, we have

$$\hat{\rho}_r(\Theta) = \frac{\int_0^{2\pi}\int_0^{\Theta} |\theta,\phi\rangle\langle\theta,\phi| \sin\theta \, d\theta \, d\phi}{2\pi(1-\cos\Theta)} \, . \tag{1.39}$$

Substitution of (1.27) and subsequent integration over ϕ gives

$$\hat{\rho}_r(\Theta) = \frac{4}{1-\cos\Theta} \sum_{m=-j}^{j} |j;m\rangle\langle j;m| \binom{2j}{j+m} \Lambda^{\Theta}(j+m, j-m) \, , \tag{1.40}$$

where

$$\Lambda^{\Theta}(p,q) = \int_0^{\Theta/2} \cos^{2p+1}\vartheta \sin^{2q+1}\vartheta \, d\vartheta \, . \tag{1.41}$$

Since $\hat{\rho}_r(\Theta)$ is diagonal in the $|j;m\rangle$ basis, the task of finding the eigenvalues is equivalent to the problem of evaluating the integral $\Lambda^{\Theta}(p,q)$. Consider the integrand

$$\cos^{2p+1}\vartheta \sin^{2q+1}\vartheta = -\frac{1}{2}\cos^{2p}\vartheta(1-\cos^2\vartheta)^q \frac{d\cos^2\vartheta}{d\vartheta} \, . \tag{1.42}$$

Using the formula [27]

$$(1+z)^n = F(-n,b;b;-z) \, , \tag{1.43}$$

where F is the hypergeometric function ${}_2F_1$, we obtain

$$\cos^{2p+1}\vartheta \sin^{2q+1}\vartheta = -\frac{1}{2}(\cos^2\vartheta)^p F(-q,b;b;\cos^2\vartheta) \frac{d\cos^2\vartheta}{d\vartheta} \, . \tag{1.44}$$

We now use [1]

$$\frac{d^n}{dz^n}[z^{c-1}F(a,b;c;z)] = \frac{\Gamma(c)}{\Gamma(c-n)} z^{c-n-1}F(a,b;c-n;z), \tag{1.45}$$

with $a = -q$, $c = b+1$ and $z = \cos^2\vartheta$ to get

$$b(\cos^2\vartheta)^{b-1}F(-q,b;b;\cos^2\vartheta) = \frac{d}{d\cos^2\vartheta}[(\cos^2\vartheta)^b F(-q,b;b+1;\cos^2\vartheta)]. \tag{1.46}$$

Comparing (1.44) and (1.46) we see that, choosing $b = p+1$, we find

$$\cos^{2p+1}\vartheta \sin^{2q+1}\vartheta = -\frac{1}{2(p+1)}\frac{d}{d\vartheta}[\cos^{2p+2}\vartheta F(-q,p+1;p+2;\cos^2\vartheta)] \, , \tag{1.47}$$

and hence, using [27]

$$F(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \quad , \qquad \mathrm{Re}(c) > \mathrm{Re}(a+b)\,, \tag{1.48}$$

one obtains

$$\Lambda^{\Theta}(p,q) = \frac{p!\,q!}{2\cdot(p+q+1)!} - \frac{\cos^{2p+2}\frac{\Theta}{2}}{2(p+1)} F(-q,p+1;p+2;\cos^2\frac{\Theta}{2}) \;, \quad q > -1 \;. \tag{1.49}$$

To simplify the above equation we return to the definition (1.41) and split the integration to get

$$\Lambda^{\Theta}(p,q) = \int_0^{\frac{\pi}{2}} \cos^{2p+1}\vartheta \sin^{2q+1}\vartheta \, d\vartheta + \int_{\frac{\pi}{2}}^{\frac{\Theta}{2}} \cos^{2p+1}\vartheta \sin^{2q+1}\vartheta \, d\vartheta \;. \tag{1.50}$$

The first integral is proportional to the beta function and the second integral can be transformed into $\Lambda^{\pi-\Theta}(q,p)$ by substitution $\vartheta \to \pi/2 - \vartheta$ so that

$$\Lambda^{\Theta}(p,q) = \frac{p!\,q!}{2\cdot(p+q+1)!} - \Lambda^{\pi-\Theta}(q,p) \;. \tag{1.51}$$

Using this formula Eq. (1.49) can be transformed into

$$\Lambda^{\Theta}(p,q) = \frac{\sin^{2q+2}\frac{\Theta}{2}}{2(q+1)} F(-p,q+1;q+2;\sin^2\frac{\Theta}{2}) \;. \tag{1.52}$$

Expression (1.40) for $\hat{\rho}_r(\Theta)$ can now be rewritten in the compact form of Eq. (1.32). To investigate the case $\Theta = \pi/2$, we calculate directly

$$\Lambda^{\frac{\pi}{2}}(p,q) = \int_0^{\frac{\pi}{4}} \cos^{2p+1}\vartheta \sin^{2q+1}\vartheta \, d\vartheta \;. \tag{1.53}$$

Substitution of $t = \tan^2\vartheta$ gives

$$\Lambda^{\frac{\pi}{2}}(p,q) = \frac{1}{2}\int_0^1 t^q (1+t)^{-(p+q+2)}\, dt. \tag{1.54}$$

Using the integral representation of the hypergeometric function [1],

$$F(a,b;c;z) = \frac{1}{B(b,c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\, dt \quad , \quad \mathrm{Re}(c) > \mathrm{Re}(b) > 0 \;, \tag{1.55}$$

we find

$$\Lambda^{\frac{\pi}{2}}(p,q) = \frac{F(p+q+2,q+1;q+2;-1)}{2(q+1)} \;, \; q > -1 \;, \tag{1.56}$$

which is a rather compact expression. We can obtain additional insight in the following way. Consider the Gauss formula for the so-called contiguous functions $F(a, b-1; c; z)$ and $F(a, b; c+1; z)$ [1]

$$c(1-z)F(a,b;c;z) - cF(a,b-1;c;z) + (c-a)zF(a,b;c+1;z) = 0\ . \qquad (1.57)$$

For the case $b = c$ and $z = -1$ we get using (1.43)

$$F(a,b;b+1;-1) = \alpha(b)F(a,b-1;b;-1) + \beta(b), \qquad (1.58)$$

where

$$\alpha(b) = \frac{b}{a-b}\ , \qquad \beta(b) = -2^{1-a}\alpha(b)\ . \qquad (1.59)$$

Iteration of (1.58) $b-1$ times gives

$$F(a,b;b+1;-1) = F(a,1;2;-1)\prod_{k=0}^{b-2}\alpha(b-k) + \sum_{s=0}^{b-2}\alpha(b-s)\prod_{k=0}^{s-1}\beta(b-k)\ . \qquad (1.60)$$

Noticing that

$$\prod_{k=0}^{s}\alpha(b-k) = \frac{b!\,(a-b-1)!}{(b-s-1)!\,(a-b+s)!} \qquad (1.61)$$

and substituting $x = b-1-s$, we obtain

$$F(a,b;b+1;-1) = \frac{b!\,(a-b-1)!}{(a-1)!}[(a-1)F(a,1;2;-1) - 2^{1-a}\sum_{x=1}^{b-1}\binom{a-1}{x}]\ . \qquad (1.62)$$

The value of $F(a, 1; 2; -1)$ can be calculated using the integral representation (1.55)

$$F(a,1;2;-1) = \frac{1-2^{1-a}}{a-1}\ , \qquad (1.63)$$

and hence we find

$$F(a,b;b+1;-1) = \frac{b!\,(a-b-1)!}{(a-1)!}[1 - 2^{1-a}\sum_{x=0}^{b-1}\binom{a-1}{x}]\ . \qquad (1.64)$$

Using this formula, Eq. (1.56) can be rewritten as

$$\Lambda^{\frac{\pi}{2}}(p,q) = \frac{1}{2(p+q+1)}\binom{p+q}{p}^{-1}[1 - 2^{-(p+q+1)}\sum_{x=0}^{q}\binom{p+q+1}{x}]\ . \qquad (1.65)$$

Using (1.51) we have

$$\Lambda^{\frac{\pi}{2}}(p,q) = \frac{2^{-(p+q+2)}}{p+q+1} \sum_{x=0}^{p} \binom{p+q+1}{x} , \tag{1.66}$$

and therefore

$$\lambda_m^{\frac{\pi}{2}} = \frac{2}{2j+1} \sum_{x=0}^{j+m} \binom{2j+1}{x} 2^{-(2j+1)} . \tag{1.67}$$

The sum on the right hand side is the univariate cumulative distribution function [1]

$$G(y;n,p) = \sum_{x=0}^{y} P(x;n,p) \tag{1.68}$$

for the binomial distribution

$$P(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} , \tag{1.69}$$

where $n = 2j+1$, $p = 1/2$ and $y = j+m$. In the limit $j \to \infty$, $G(y;n,p)$ approaches the step function

$$\lim_{j\to\infty} G(j+m;2j+1,1/2) = \begin{cases} 0 , & \lim_{j\to\infty} \frac{m}{j} < 0 , \\ \frac{1}{2} , & m = 0 , \\ 1 , & \lim_{j\to\infty} \frac{m}{j} > 0 . \end{cases} \tag{1.70}$$

To obtain the behaviour for large but finite values of j, we use the Chernoff bounds [17]

$$\begin{aligned} \sum_{x<np(1-\epsilon)} P(x;n,p) &\leq e^{-\epsilon^2 np/3} , \\ \sum_{x>np(1+\epsilon)} P(x;n,p) &\leq e^{-\epsilon^2 np/3} , \end{aligned} \tag{1.71}$$

which are valid for $0 \leq \epsilon \leq 1$. Choosing $\epsilon = j^{-1/3}$, we obtain Eq. (1.36), in agreement with Eq. (1.70).

Chapter 2

Complexity of dynamics in the context of chaos: case study of the quantum baker's map

2.1 Introduction

In the previous chapter we advocated the information-theoretic approach to the problem of quantifying the complexity of dynamics for both classical and quantum systems. In this approach the system dynamics can be viewed as the minimal knowledge which allows us to predict the results of experiments conducted upon the system given the initial conditions with some fixed accuracy. The amount of this minimal knowledge (in bits) quantifies the complexity of the system dynamics. We have shown that our ability to predict the system behaviour can be strongly affected by unknown perturbations of the system. This is because the information about the history of such perturbations as well as additional information about the initial conditions beyond the originally given accuracy can be crucial if we want to predict the evolution of a system. This contributes to the complexity of the system dynamics, and, clearly, the amount of this contribution can grow with time making the prediction of evolution practically impossible. In such cases we can

talk about chaos and use the above contributions to the complexity of dynamics as a measure of chaoticity.

It has become a normal practice to use continuum spaces to model the state space of the system. This is convenient in the sense that we can use methods of differential and integral calculus. In contrast, information-theoretic methods often require a discrete description of the system in question. It is highly debatable which description, discrete or continuous, is more fundamental. Fortunately, for a large number of interesting physical systems the gap between continuous and discrete descriptions is not as big as it may seem at the first glance. In fact, a discrete representation of continuous nonlinear systems can often be introduced without loss of relevant information. This can be achieved using a general method of *symbolic dynamics* which we review in the next section. The basic idea of the method is to use symbolic sequences, i.e. string of letters from a finite alphabet $\mathbb{A}$, to represent every possible state of the system. In order to do this in the most general case we need to discretize time and partition the phase space of the system into a finite number of cells which can be labeled by the letters from the alphabet $\mathbb{A}$. Then, we can list a sequence of symbols corresponding to the cells visited by the exact trajectory of the system during its evolution from some initial state. Such a sequence of finite-precision points, coarse grained both on phase space and in time, can be viewed as a *coarse-grained trajectory*. For a large class of interesting systems, e.g. chaotic systems, symbolic sequences introduced above, retain all the vital information about the system dynamics. At the same time, the discretized system dynamics becomes just a shift on the symbolic sequences.

Methods of symbolic dynamics have proved extremely useful in the field of classical chaos. It is possible, for instance, to calculate the average information, $\langle I \rangle$, necessary to specify a particular coarse-grained trajectory. For classical chaotic systems this information has been shown [60] to grow linearly with time

$$\langle I \rangle \sim H_{\mathrm{KS}}\, t\ , \tag{2.1}$$

where "$\sim$" means the leading asymptotic behaviour, and the linear rate of infor-

mation increase, $H_{\rm KS}$, called the Kolmogorov-Sinai or metric entropy, quantifies the degree of classical chaos. The information $\langle I \rangle$ can be viewed as the average information missing from the coarse-grained history of the system evolution: we need more and more information about the initial condition to predict the coarse-grained trajectory. As we explained earlier, this information contributes to the complexity of the system dynamics and, as we see from Eq. (2.1), the rate at which this contribution grows with time is a signature of chaos.

The generalization of the classical theory of chaos to the purely quantum case is a nontrivial procedure. One problem is that quantum systems are not sensitive to the initial conditions due to the unitary nature of quantum evolution. The information-theoretic approach, however, can be used to define chaos in such a way that classical and quantum systems are treated on an equal footing. Open system, classical or quantum, can be unpredictable due to interaction with a partially unknown environment. Given, with some finite accuracy, the initial state of the system we ask how much additional information $I_{\rm min}$ we need in order to predict the system evolution within the required precision up to time t. This quantity is a generalization of I introduced above.

In the previous chapter we argued that $I_{\rm min}$ is a good measure of unpredictability in the general case of an open system with a partially unknown environment. We pointed out that the analytical results supporting this statement were obtained for classical systems only using the methods of classical symbolic dynamics. It is therefore reasonable to anticipate that a quantum generalization of classical symbolic dynamics can provide a useful tool for the investigation of quantum unpredictability (or chaos).

The classical baker's map is a particularly simple example of a genuine classical chaotic system: its symbolic dynamics is just a shift on bi-infinite strings of bits. In section 2.3 we review the definition of the classical baker's map in terms of its symbolic dynamics. Together with section 2.2 on classical symbolic dynamics and related measures of dynamical complexity, this would constitute a classical background on which we would develop a quantum analogue of classical symbolic

dynamics and generalize the related measures of dynamical complexity to the quantum case. We introduce the notion of the qubit representation where the basis states are chosen as direct products of qubits (states of a two-level system). Any state in a 2^N-dimensional Hilbert space can therefore be represented as a superposition of qubit strings. We use this representation to define a family of quantum baker's maps and show that in the classical limit such maps can be viewed as shifts on bi-infinite strings of bits as required by the classical symbolic representation of the baker's map.

We use the framework of the decoherent histories approach to introduce a coarse-graining of the state space of the quantum baker's map and to define $I_{\min}$ relative to the coarse-graining. We then compute $I_{\min}$ in the classical limit for several types of coarse-graining. To our knowledge, this is the first time that the framework of decoherent histories was successfully applied to a quantized version of a nonlinear classical system.

2.2 Method of symbolic dynamics

Before we discuss complexity of dynamics in the context of chaos we must introduce a convenient description of the dynamics. In our case this means that the dynamics should be represented by a discrete model suitable for information-theoretic study. The method of symbolic dynamics was specifically built for this purpose. Within this method all possible states of the system are encoded in the form of symbolic sequences and the evolution of the system state is translated into a discrete transformation of such symbolic sequences. In order to emphasize the generality of the method we consider a physical system whose dynamical laws can be stated as a system of differential equations

$$\dot{x} = F(x) , \qquad x \in \mathbb{S} \subset \mathbb{R}^d , \tag{2.2}$$

where x represent the state of the system, and the state space $\mathbb{S}$ is a compact subset of $\mathbb{R}^d$. Within this chapter we shall identify $\mathbb{S}$ with the phase space of

the physical system, giving time a special role as is customary in nonrelativistic mechanics[1]. The (smooth) vector field F generates a flow $f_t : \mathbb{S} \to \mathbb{S}$, such that $f_t(x_0)$ is the image x_t of the initial condition $x_0 \in \mathbb{S}$ at time t. The pair $(F, \mathbb{S})$ is usually referred to as a dynamical system.

2.2.1 Discretization of time

Discretization of time is the first necessary step in an information-theoretic study of chaotic dynamical systems. This can be introduced without losing the relevant information about the structure of the motion by using the appropriate Poincaré surface $\Xi \in \mathbb{R}^d$. The surface Ξ is chosen to be $d-1$ dimensional and satisfying the requirement that every trajectory in $\mathbb{S}$ must intersect Ξ. To be more precise, for any initial condition x_0 there must exist a moment of time t such that $f_t(x_0) \in \Xi$. It is further required that, for any $x \in \Xi$, $F(x)$ is not tangent to Ξ. Whenever motion takes place in a compact set $\mathbb{S}$, as assumed in (2.2), it is possible to find a Poincaré surface. This follows from the fact that an arbitrary component $x^{(i)}$ of x is a bounded function and therefore $\dot{x}_t^{(i)} = F^{(i)}(x_t) = 0$ at some instant t, which defines a possible $d-1$ dimensional surface of section that satisfies all the mentioned requirements. We define $\{x_1, x_2, \dots\}$ to be the sequence of successive returns onto the surface Ξ. In the new discrete time k they can be connected by a Poincaré map

$$x_k = M(x_{k-1}) \,, \tag{2.3}$$

where $M(x_{k-1}) = f_{\Delta t_k}(x_{k-1})$ and Δt_k is the time elapsed between the events x_{k-1} and x_k.

If the discrete time was used to define the dynamical laws of our physical system there would be no reason to introduce a $d-1$ dimensional surface of section and we would not lose any information about the dynamics of the system. However, even in the continuous case considered above the discretization of time still preserves a

[1]Alternatively, we could combine the time, coordinates and momenta into one collective variable and use this variable to address the states of the system. This can become important only in cases when the time is considered on equal footing with other parameters of the system state, as in relativistic mechanics.

lot of interesting features of the motion. In particular, the fixed points of F and M coincide since the surface Ξ passes through them; and any periodic orbit of F is transformed into a periodic orbit[2] of M.

2.2.2 Coarse-graining

We divide the state space $\mathbb{S}$ of the system into a finite number b of disjoint cells by introducing a partition

$$\mathcal{E} \equiv \{\mathbb{E}_k | \bigcup_{k=1}^{b} \mathbb{E}_k = \mathbb{S} \ , \ \mathbb{E}_j \cap \mathbb{E}_k = 0 \ \text{ for } j \neq k\}. \tag{2.4}$$

Starting from some initial condition x_0 the action of M defines a trajectory $f(x_0) \equiv \{x_0, x_1, \dots\}$ which visits various elements of $\mathcal{E}$. Let the symbol s_k denote the index of the cell $\mathbb{E}_{s_k}$ visited by the trajectory at time k. One can think of s_k as a coarse-grained value of the exact x_k and define a coarse-grained trajectory as a symbolic sequence ${}^J f(x_0) \equiv \{s_0, \dots, s_J\}$. Knowledge of ${}^J f(x_0)$ can be used to recover the history of the system to a better precision than given by the coarse-graining. For instance, the observation of the first and the second symbols s_0 and s_1 of ${}^J f(x_0)$ implies that the initial state x_0 was not only in $\mathbb{E}_{s_0}$ but also in $M^{-1}\mathbb{E}_{s_1}$. The sequence ${}^J f(x_0)$ can be produced only if the initial state x_0 belongs to the intersection

$$\mathbb{E}_{s_0 \dots s_J} \equiv \mathbb{E}_{s_0} \cap M^{-1}\mathbb{E}_{s_1} \cap \dots \cap M^{-J}\mathbb{E}_{s_J} \ . \tag{2.5}$$

The size of $\mathbb{E}_{s_0 \dots s_J}$ determines the accuracy to which we must know x_0 to be able to reconstruct the coarse-grained trajectory ${}^J f(x_0)$. An interesting situation arises when repeating this argument for increasing J we arrive at smaller and smaller intersections $\mathbb{E}_{s_0 \dots s_J}$. In this situation the amount of information to specify a specific intersection $\mathbb{E}_{s_0 \dots s_J}$ as a part of $\mathbb{S}$ increases with J. This means that in order to predict the system evolution for a longer period of time we need more and more information about the initial condition x_0, or, equivalently, we need to know the initial condition with a better and better precision. If the available precision is

[2] i.e. a finite sets of points which are mapped cyclically by M

limited, as is usually the case in practical situations, we lose the ability to predict the system behaviour. How fast this may happen is one of the key questions in the theory of chaos.

In this subsection, we are not introducing any perturbations. More generally, we would calculate the minimal information which is necessary to specify a coarse-grained trajectory given the initial condition up to some fixed accuracy. This information quantifies the complexity of the system dynamics and it can be defined in both classical and quantum cases. The complexity of system dynamics can grow with time for different reasons – such as perturbations or sensitivity to initial conditions. In any case, the rate at which it grows will quantify the resources needed to predict the system evolution, and thus it is also an indicator for chaotic behaviour.

All nonempty sets $\mathbb{E}_{s_0 \ldots s_J}$ for given J define a new partition ${}^J\mathcal{E}$ of the phase space $\mathbb{S}$ as

$$ {}^J\mathcal{E} \equiv \{\, \mathbb{E}_{s_0 \ldots s_J} \,|\, (s_0, \ldots, s_J) \in \mathbb{A}^{J+1} \,:\, \mathbb{E}_{s_0 \ldots s_J} \neq \emptyset \,\}. \tag{2.6} $$

We say that ${}^J\mathcal{E}$ is a *refinement* of $\mathcal{E}$ under M.[3] It is important to emphasize that the refinements $\{{}^J\mathcal{E} \,|\, J = 1, 2 \ldots\}$ are produced by the dynamics itself (dynamical refinements). The study of the infinite symbolic sequences $\{{}^\infty f(x_0) \,|\, x_0 \in \mathbb{S}\}$ is equivalent to that of real trajectories of the original system (2.2) if for every such sequence ${}^\infty f(x_0) \equiv \{s_0, s_1, \ldots\}$ there exists only one point (initial condition) $x_0 \in \mathbb{S}$ that generates ${}^\infty f(x_0)$ under the action of M. This is achieved if the original partition $\mathcal{E}$ is *generating*, i.e. if $\mathcal{E}$ refines itself indefinitely under the dynamics. It is easy to show that if one generating partition exists, there are infinitely many of them. For example, if $\mathcal{E}$ is generating then any refinement $\mathcal{E}'$ of $\mathcal{E}$ is also generating. However, the problem of constructing a single generating partition is highly nontrivial and has not been entirely solved (see Ref. [5], p73). The ultimate yet unsolved problem is to find a generating partition with the minimum number of elements.

[3] More generally, a partition $\mathcal{E}'$ is called a refinement of $\mathcal{E}$ if any element of $\mathcal{E}$ is a union of some elements of $\mathcal{E}'$.

If the map M is invertible, it is convenient to extend each infinite sequence ${}^{\infty}f(x_0)$ to a doubly-infinite sequence

$$\ldots s_{-2}s_{-1}.\, s_0 s_1 \ldots , \tag{2.7}$$

where the symbols to the left of the dot represent the backward iterates of x_0. In this representation the system dynamics is equivalent to a shift of the dot to the right for the forward iterations of M and to the left for the backward iterations. The study of the system dynamics is thereby reduced to the study of the *symbolic dynamics* of a map, i.e. the study of the doubly-infinite symbolic sequences generated by the map.

2.2.3 Information-theoretic measures of dynamical complexity

The behaviour of the cardinality $|\,{}^{J}\mathcal{E}\,|$ in the limit $J \to \infty$ can be used to quantify the complexity achievable by the system dynamics. We define[4]

$$h(M|\,\mathcal{E}) = \lim_{J\to\infty} \frac{\log|\,{}^{J}\mathcal{E}\,|}{J}\,. \tag{2.8}$$

This quantity still depends on the choice of partitioning $\mathcal{E}$ which can be highly non-unique. To remove this dependence we define

$$h(M) \equiv \sup_{\mathcal{E}} h(M|\,\mathcal{E})\,. \tag{2.9}$$

This quantity is known as the topological entropy of the map M (see e.g. ref. [2]). It quantifies the maximum complexity of dynamics generated by the map M.

A more detailed measure of dynamical complexity can be constructed using the notion of Kolmogorov (or algorithmic) complexity. A rigorous mathematical definition of Kolmogorov complexity can be found in [39]. Here it is sufficient to know that the Kolmogorov complexity $K(\boldsymbol{\alpha})$ of a symbolic string $\boldsymbol{\alpha}$ is the smallest number of bits needed to encode $\boldsymbol{\alpha}$ using a universal computer. The algorithmic

[4] This limit has been shown to exist (see lemma 1.18 in Ref. [10]).

complexity of dynamics generated by the map M relative to the partition $\mathcal{E}$ given the initial condition x_0 is defined as [5]

$$K(M|\mathcal{E}, x_0) \equiv \limsup_{J\to\infty} \frac{K[{}^J f(x_0)]}{J} . \tag{2.10}$$

This quantity is also known as the complexity of the trajectories of a dynamical system [12].

In many physical cases the state space of the system $\mathbb{S}$ is measurable and the measure can be introduced in a very natural way. For instance, integration over the phase space of a Hamiltonian system is routinely performed with respect to the measure $dp\,dq$. Using the measure we can define a different measure of dynamical complexity known as metric or Kolmogorov-Sinai entropy. Fix a normalized Borel measure μ on $\mathbb{S}$ and consider a symbolic sequence $(s_0, s_1, \ldots, s_J)$. Such a sequence represents a coarse-grained trajectory ${}^J f(x_0)$ where the initial point x_0 belongs to a set of measure

$$\mu(s_0 \ldots s_J) \equiv \mu(\mathbb{E}_{s_0} \cap M^{-1}\mathbb{E}_{s_1} \cap \cdots \cap M^{-J}\mathbb{E}_{s_J}) . \tag{2.11}$$

This quantifies the "probability" of the symbolic sequence $(s_0, s_1, \ldots, s_J)$. The average uncertainty about the symbolic sequence (or, which is equivalent, the coarse-grained trajectory) can be quantified by the entropy

$$H(\mu, \mathcal{E}, J) \equiv -\sum_{s_0, s_1, \ldots, s_J} \mu(s_0 \ldots s_J) \log \mu(s_0 \ldots s_J) . \tag{2.12}$$

The KS-entropy is defined as the asymptotic rate at which the uncertainty about the dynamics accumulates with time:

$$H_{\mathrm{KS}}(\mu) \equiv \sup_{\mathcal{E}} \lim_{J\to\infty} \frac{H(\mu, \mathcal{E}, J)}{J} , \tag{2.13}$$

where as in Eq. (2.9) the supremum is taken to remove any dependence on the partitioning.

The three measures of dynamical complexity defined above in (2.9), (2.10) and (2.13) are closely related. For instance, the KS-entropy and the topological entropy

[5] $\left[A = \limsup_{J\to\infty} a(J)\right] \Leftrightarrow \left[\exists N \text{ such that } \forall n > N \; A = \sup\{a(n), a(n+1), \ldots\}\right]$.

satisfy the equality

$$h(M) = \sup_{\mu} H_{\rm KS}(\mu)\,, \tag{2.14}$$

where the supremum is taken over the set of all M-invariant normalized Borel measures. Moreover, if μ is ergodic[6] then for μ-almost all x_0 the algorithmic complexity of dynamics and the KS-entropy are related via the equation

$$\sup_{\mathcal{E}} K(M|\mathcal{E}, x_0) = H_{\rm KS}(\mu)\,, \tag{2.15}$$

where the supremum is taken over all measurable finite partitions of $\mathbb{S}$. These measures of dynamical complexity can also be related to a more familiar characterization of chaos in terms of Lyapunov exponents. For a large class of dynamical systems the following simple identity holds [50]

$$H_{\rm KS} = \sum_j \lambda_j^+\,, \tag{2.16}$$

where the right hand side is the sum of all positive Lyapunov exponents of the system.

In conclusion of this section we note that all three measures considered above are defined as asymptotic quantities. In other words they all quantify the complexity of system dynamics "in the long run". However, very often we are interested in the behaviour of the system during a particular period of time. In such cases we may drop the limit $J \to \infty$ in the above equations leaving the complexities as functions of J.

2.3 Classical baker's map

The classical baker's transformation [4] maps the unit square $\mathbb{S} = \{\,(q,p)\,|\,0 \le q,p \le 1\}$ onto itself according to

$$(q,p) \longmapsto \begin{cases} \left(2q, \frac{1}{2}p\right)\,, & \text{if } 0 \le q \le \frac{1}{2}, \\ \left(2q-1, \frac{1}{2}(p+1)\right)\,, & \text{if } \frac{1}{2} < q \le 1. \end{cases} \tag{2.17}$$

[6]The measure μ is called ergodic if all M-invariant subsets of $\mathbb{S}$ are either of measure 0 or 1.

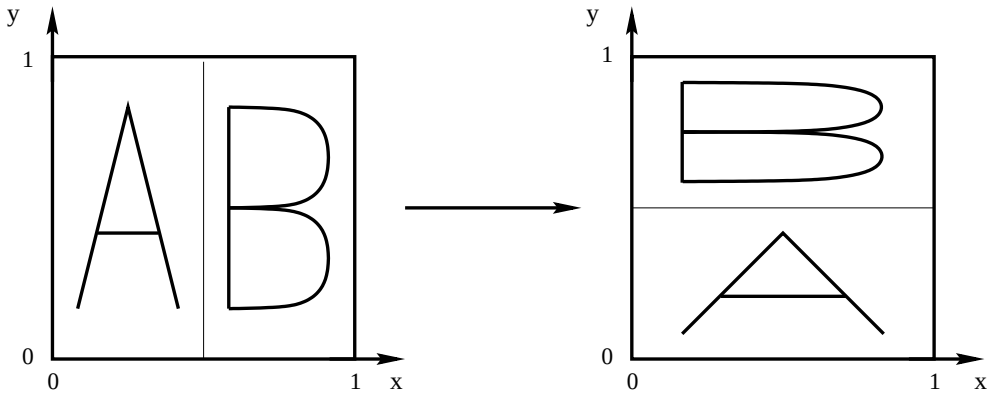


Figure 2.1: The classical baker's map on the unit square.

This corresponds to compressing the unit square in the p direction and stretching it in the q direction, while preserving the area, then cutting it vertically, and finally stacking the right part on top of the left part as shown in Fig. 2.1. This is a kind of transformation that a piece of dough would normally undergo under the action of a baker.

Choose a partition $\mathcal{E} = \{\mathbb{E}_0, \mathbb{E}_1\}$ where $\mathbb{E}_0$ and $\mathbb{E}_1$ denote the left and the right halves of the unit square[7]

$$\begin{aligned} \mathbb{E}_0 &\equiv \{\,(q,p)\,|\,0 \le q \le 1/2;\, 0 \le p \le 1\} \\ \mathbb{E}_1 &\equiv \{\,(q,p)\,|\,1/2 < q \le 1;\, 0 \le p \le 1\}\,. \end{aligned} \tag{2.18}$$

We see that this partition is generating and moreover it has the minimal number of elements. As shown in the previous section, we can now use this partition to obtain an alternative (but equivalent) representation of the map in terms of its symbolic dynamics. In this representation the action of the map is given by the (Bernoulli) shift U defined on a symbolic string of zeros and ones

$$s = \cdots s_{-2}s_{-1}\,.\,s_0 s_1 \cdots \qquad s_k \in \{0,1\}\ \forall k \in \mathbb{Z} \tag{2.19}$$

[7]Here we have chosen disjoint sets in agreement with the definition (2.4). In this case there is one-to-one correspondence between the bi-infinite sequences and the points in the state space unlike in the usual case where we would some times assign two different strings to one point (e.g.: $\ldots 000.0111\ldots$ and $\ldots 000.1000\ldots$). As a consequence, in our case not all the binary strings correspond to a point in $\mathbb{S}$. This subtlety in "grammar" affects a set of points of measure zero and is therefore trivial.

as $Us = s'$, where $s'_k = s_{k+1}$. In other words, at each time step, the entire string is shifted one place to the left while the dot remains fixed. As we explained in the previous section each symbolic string (2.19) corresponds to a point (q, p) in $\mathbb{S}$. In this particular case this correspondence is given by the following relations:

$$q = \sum_{k=0}^{\infty} s_k 2^{-k-1} , \quad p = \sum_{k=1}^{\infty} s_{-k} 2^{-k} . \tag{2.20}$$

These relations between points (q, p) and symbolic strings are particular to the baker's transformation, however, as we have seen in the previous section, the method of symbolic dynamics is very general and can be applied to a large class of chaotic maps.

2.4 Quantum baker's map

2.4.1 Notation

Most results of this chapter are phrased in terms of finite binary strings. It will be convenient to adopt a slightly different and more flexible notation than the one used in Eq. (2.19). Here, a binary string

$$\xi_{s:f} \equiv \begin{cases} \xi_s \xi_{s+1} \cdots \xi_f & (s \leq f) \\ \xi_s \xi_{s-1} \cdots \xi_f & (s > f) , \end{cases} \tag{2.21}$$

where $\xi_i \in \{0, 1\}$ is a bit, can have increasing $(s < f)$ or decreasing $(s > f)$ indices. We will use bold Greek and Latin letters to denote binary strings, e.g.,

$$\boldsymbol{\alpha} = \xi_{s:f} \quad \text{or} \quad \boldsymbol{x} = x_{h:t} . \tag{2.22}$$

The length of a string $\boldsymbol{\alpha}$ will be denoted by $|\boldsymbol{\alpha}|$; e.g., in the above example, $|\boldsymbol{\alpha}| = |f - s| + 1$. Concatenation of strings is defined in the usual way. Again considering the above example, $\boldsymbol{\alpha x}$ is the string $\boldsymbol{\alpha x} = \xi_s ... \xi_f x_h ... x_t$. Any string $\boldsymbol{\alpha}$ represents a natural number through its binary expansion

$$\boldsymbol{\alpha} = \sum_{k=1}^{|\boldsymbol{\alpha}|} 2^{|\boldsymbol{\alpha}|-k} \boldsymbol{\alpha}_{(k)} , \tag{2.23}$$

where $\boldsymbol{\alpha}_{(k)}$ denotes the k-th bit of $\boldsymbol{\alpha}$, $1 \leq k \leq |\boldsymbol{\alpha}|$, such that

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_{(1)}\boldsymbol{\alpha}_{(2)}...\boldsymbol{\alpha}_{(|\boldsymbol{\alpha}|)} \ . \tag{2.24}$$

Thus our notation does not distinguish between a binary string and the corresponding natural number. Similarly, two strings $\boldsymbol{\alpha}$ and $\boldsymbol{x}$ can be combined to represent a rational number

$$\boldsymbol{\alpha}.\boldsymbol{x} \equiv \sum_{k=1}^{|\boldsymbol{\alpha}|} 2^{|\boldsymbol{\alpha}|-k}\boldsymbol{\alpha}_{(k)} + \sum_{k=1}^{|\boldsymbol{x}|} 2^{-k}\boldsymbol{x}_{(k)} \ . \tag{2.25}$$

2.4.2 Quantum baker's map on a finite qubit string

A general procedure of quantization can be broken into two distinct steps [70]. The first step is to define the kinematics which includes the state space of the system and the operators that describe the system. In our case that would be a Hilbert space and a pair of conjugate variables representing position q and momentum p. Given the kinematics we should be able to described the system at any moment of time. After the kinematics is defined, the second step is to introduce the dynamics or a time evolution of the system. Both of these steps can be highly non-unique. A rough guide can be given by the correspondence principle i.e. a requirement that the quantized version of the classical system should in some sense resemble the classical system in the "classical limit".

Numerous quantization schemes have been proposed and a lot of attention has been devoted to the study of the resulting quantum baker's maps. The original definition of the map [6, 54] is based on Weyl's quantization [70] of the unit square. Essentially the same map has been derived by algebraic methods [53, 38] as well as by considering the transition from ray to wave optics [29]. Path integral quantization [18], and a definition on a sphere [48] have been proposed.

During the last decade, the semiclassical properties of the quantum baker's map have been studied extensively [47, 19, 55, 37, 34, 18], its long-time behavior has been investigated [19, 34], and it has been shown to display hypersensitivity to perturbation [59, 61]. The vast majority of the results are based on numerical

simulations and not on rigorous proofs. Experimentally feasible realizations of the map were proposed [29, 57, 14].

In the previous sections we have seen that a large class of classically chaotic maps can be viewed as Bernoulli shifts on bi-infinite symbolic strings. In particular, we demonstrated that the classical baker's map is equivalent to a shift on bi-infinite strings of bits. Recently, a whole class of quantum baker's maps has been defined on the finite strings of quantum bits [63]. This class of quantum baker's maps, which can also be derived from the semi-quantum maps introduced in Ref. [55], is the subject of our research. In the rest of this section we review the definitions and the necessary background.

To define the kinematics of the quantum baker's map we need a state space $\mathbb{S}$ and a couple of conjugate variables $(\hat{q}, \hat{p})$. We choose $\mathbb{S} = \mathcal{H}_D$ where $\mathcal{H}_D$ is a D-dimensional Hilbert space. We further assume that $D = 2^N$, which is the dimension of the Hilbert space of N qubits, i.e., N two-state systems.

For each qubit we use $|0\rangle$ and $|1\rangle$ to denote two orthonormal states, and define

$$|q_j\rangle \equiv |\xi_1\rangle \otimes |\xi_2\rangle \otimes \cdots \otimes |\xi_N\rangle \,, \tag{2.26}$$

where $\xi_l \in \{0, 1\}$ and for each $j \equiv \sum_{l=1}^{N} \xi_l 2^{N-l}$ we define[8]

$$q_j \equiv 0.\xi_1\xi_2 \ldots \xi_N 1 = 0.\xi_{1:N}1 \,. \tag{2.27}$$

Using the basis $\{|q_j\rangle\}$ we define the position operator as $\hat{q} \equiv \mathrm{diag}(q_1, \ldots, q_{2^N})$. We see that the eigenvalues of $\hat{q}$ each satisfy the inequality $0 < q_j < 1$ in analogy with the values of "position" variable x in the classical baker's map. We introduce the symbolic notation

$$|.\xi_{1:N}\rangle = |.\xi_1\xi_2 \ldots \xi_N\rangle = e^{i\pi/2}|q_j\rangle \,, \tag{2.28}$$

which is closely analogous to Eq. (2.19), where the bits to the right of the dot specify the position variable; see Ref. [63] for the reason for the phase shift $e^{i\pi/2}$. To construct the "momentum" operator $\hat{p}$ we use the fact that momentum and

[8] Here we use the string notation introduced in section 2.4.1.

position eigenstates are related through a generalized quantum Fourier transform $\hat{F}$ [54], i.e.,

$$|p_k\rangle \equiv \hat{F}|q_k\rangle\,, \tag{2.29}$$

where

$$\begin{aligned} e^{-i\pi(0.\xi_{1:N}1)}\hat{F}|.\xi_{1:N}\rangle \;\equiv\; & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_N 1)]|1\rangle\}\otimes \\ & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_{N-1}\xi_N 1)]|1\rangle\}\otimes\cdots\otimes \\ & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_{1:N} 1)]|1\rangle\}\,, \end{aligned} \tag{2.30}$$

We use this as a definition of $|p_k\rangle$ and again, in analogy to Eq. (2.19), we define the notation

$$|\xi_{1:N}.\rangle = |p_k\rangle\,, \tag{2.31}$$

where $p_k = 0.\xi_{N:1}1$. Since $\hat{F}$ is a unitary operator, each number p_k is an eigenvalue of $\hat{p}$ corresponding to the eigenvector $|p_k\rangle$. By applying a *partial* quantum Fourier transform [63] to the position eigenstates, one obtains the family of states

$$\begin{aligned} |\xi_{1:n}.\xi_{n+1:N}\rangle \;\equiv\; & |\xi_{n+1}\rangle\otimes\cdots\otimes|\xi_N\rangle e^{i\pi(0.\xi_{n:1}1)}\otimes \\ & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_1 1)]|1\rangle\}\otimes \\ & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_2\xi_1 1)]|1\rangle\}\otimes\cdots\otimes \\ & \sqrt{1/2}\{|0\rangle + \exp[2\pi i(0.\xi_{n:1} 1)]|1\rangle\}\,, \end{aligned} \tag{2.32}$$

where $1 \le n \le N-1$. More precisely, the state $|\xi_{1:n}.\xi_{n+1:N}\rangle$ is obtained by applying the Fourier transform operator to the n rightmost bits of the position eigenstate $|.\xi_{n+1:N}\xi_{n:1}\rangle$. For given n, these states form an orthonormal basis. The state $|\xi_{1:n}.\xi_{n+1:N}\rangle$ is localized in both position and momentum: it is strictly localized within a position region of width $1/2^{N-n}$, centered at position $q = 0.\xi_{n+1:N}1$, and it is crudely localized within a momentum region of width $1/2^n$, centered at momentum $p = 0.\xi_{n:1}1$.

Now as we have the quantum kinematics on the unit square we are in the position to introduce the dynamics of the baker's map. For each n, $1 \le n \le N-2$, a quantum baker's map can be defined as the transformation

$$\hat{B}_n|\xi_{1:n}.\xi_{n+1:N}\rangle \equiv |\xi_{1:n+1}.\xi_{n+2:N}\rangle\,, \tag{2.33}$$

where the dot is shifted by one position. In phase-space language, the map $\hat{B}_n$ takes a state localized at $(q,p) = (0.\xi_{n+1:N}1, 0.\xi_{n:1}1)$ to a state localized at $(q',p') = (0.\xi_{n+2:N}1, 0.\xi_{n+1:1}1)$, while it stretches the state by a factor of two in the q direction and squeezes it by a factor of two in the p direction. This analogy with the classical baker's map (see Fig. 2.1) motivates calling the maps (2.33) "quantum baker's maps." For $n = N-1$, the map is the original quantum baker's map as defined in Ref. [54], which in our notation becomes

$$\hat{B}_{N-1}|\xi_{1:N-1}.\xi_N\rangle \equiv |\xi_{1:N}.\rangle \ , \tag{2.34}$$

and for $n = 0$, the map is

$$\hat{B}_0|.\xi_{1:N}\rangle \equiv |\xi_1.\xi_{2:N}\rangle \ . \tag{2.35}$$

Below we show that all the maps (2.33,2.34,2.35) reduce to the classical baker's map in the limit $\hbar \to 0$. We then use the framework of the decoherent histories approach to generalize the notion of dynamical complexity for the case of quantum dynamics. Finally, we consider this quantity using the quantum baker's map as an example.

2.5 Classical limit in terms of symbolic dynamics for the quantum baker's map

We derive a simple closed form for the matrix elements of the quantum baker's map that shows that the map is an approximate shift in a symbolic representation based on discrete phase space. We use this result to give a formal proof that the quantum baker's map approaches a classical Bernoulli shift in the limit of a small effective Planck's constant. In subsection 2.5.1, we state the results of this section and discuss their significance. Subsection 2.5.2 contains the derivations and proofs.

2.5.1 Results

Equation (2.33) is a mixed representation of the quantum baker's map, using different bases on both sides of the equation. To go beyond the heuristic phase-space interpretation of the map given at the end of the last section, we need to express the matrix elements of $\hat{B}_n$ with respect to a single basis, i.e., we need to find

$$C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) \equiv \begin{cases} \langle .\xi^1_{1:N}|\hat{B}_0|.\xi^0_{1:N}\rangle & \text{if } n=0 \\ \langle \xi^1_{1:n}.\xi^1_{n+1:N}|\hat{B}_n|\xi^0_{1:n}.\xi^0_{n+1:N}\rangle & \text{if } 1\le n\le N-1\ . \end{cases} \tag{2.36}$$

A main result of this chapter is the following simple formula, which will be proved in Sec. 2.5.2:

$$C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) = \Phi(\xi^0_1,\xi^1_N)\frac{\delta(\xi^0_{n+2:N}-\xi^1_{n+1:N-1})}{2^{n+1}\sin[\pi(0.\xi^0_{n+1:1}1-0.\xi^1_{n:1}1)]}\ , \tag{2.37}$$

where $1\le n\le N-2$ and Φ is a phase factor given by

$$\Phi(\xi^0_1,\xi^1_N) = \frac{1}{\sqrt{2}}[i(-1)^{\xi^1_N}-(-1)^{\xi^0_1}]\ . \tag{2.38}$$

For the case $n=0$, one obtains

$$C_0^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) = \frac{1-i}{2}\delta(\xi^0_{2:N}-\xi^1_{1:N-1})e^{i\frac{\pi}{2}|\xi^0_1-\xi^1_N|}\ , \tag{2.39}$$

and for $n=N-1$,

$$C_{N-1}^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) = \frac{\Phi(\xi^0_1,\xi^1_N)}{2^N\sin[\pi(0.\xi^0_{N:1}1-0.\xi^1_{N-1:1}1)]}\ . \tag{2.40}$$

The coefficients $C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N})$ given in (2.37) are zero unless $\xi^0_{n+2:N}=\xi^1_{n+1:N-1}$, i.e., unless the position bits $\xi^1_{n+1:N-1}$ of the final state are obtained by shifting the corresponding position bits of the initial state. Furthermore, the sin term in the denominator ensures that $C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N})$ is strongly peaked for $\xi^0_{n+1:2}=\xi^1_{n:1}$, i.e., if the momentum bits $\xi^1_{n:1}$ of the final state are obtained by shifting the corresponding momentum bits of the initial state. The formula (2.37) therefore establishes rigorously that the maps (2.33) are approximate shift maps, a result which had been obtained numerically in Ref. [55].

To formulate the question of the classical limit of the baker's map, we use the concept of coarse-graining in the spirit of the consistent (or decoherent) histories approach [28, 46, 24]. For this, we introduce projectors on subspaces corresponding to symbolic strings $\boldsymbol{y}$ of length l. We fix in advance an upper limit, $k_{\max}$, on the number of iterations, k, considered; this is necessary because in computing the classical limit of a chaotic map, the limit $\hbar = 2^{-(N+1)}/\pi \to 0$ has to be taken before the limit $k \to \infty$ [8]. We will show that, for given l and $k_{\max}$, it is always possible to choose $\hbar$ in such a way that the coarse-grained quantum dynamics is arbitrarily close to a shift of the string $\boldsymbol{y}$. In contrast to the approach of Refs. [53, 38], in taking the limit $\hbar \to 0$, we always remain in the finite-dimensional Hilbert space on which our maps are defined.

As before, we are considering basis states of the form $|\xi_{1:n}.\xi_{n+1:N}\rangle$. As we let N increase, the number of position bits to the right of the dot, $m \equiv N - n$, remains fixed. We define $r = N - l$ as the number of bits ignored in the coarse graining. In the following, we always assume that $k \leq k_{\max} < r$. We also simplify the notation by writing $\hat{B}$ instead of $\hat{B}_n$ assuming implicitly the dependence on n.

We are now in a position to introduce a family of projectors

$$P_{\boldsymbol{y}}^{r,k} \equiv \begin{cases} \displaystyle\sum_{|\boldsymbol{x}|=r-k,\,|\boldsymbol{g}|=k} |\boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2\boldsymbol{g}\rangle\langle\boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2\boldsymbol{g}| & \text{if } k < m \\ \displaystyle\sum_{\substack{|\boldsymbol{x}|=r-k,\,|\boldsymbol{g}^2|=m \\ |\boldsymbol{g}^1|=k-m}} |\boldsymbol{x}\boldsymbol{y}\boldsymbol{g}^1.\boldsymbol{g}^2\rangle\langle\boldsymbol{x}\boldsymbol{y}\boldsymbol{g}^1.\boldsymbol{g}^2| & \text{if } k \geq m\ , \end{cases} \tag{2.41}$$

where $\boldsymbol{y}^1\boldsymbol{y}^2 = \boldsymbol{y}$ and $|\boldsymbol{y}^2| = m - k$. A measurement of the projectors $P_{\boldsymbol{y}}^{r,k}$ can be interpreted as a measurement in which the $r-k$ leftmost bits and the k rightmost bits of the symbolic string are not resolved. By normalizing these projectors, we obtain a family of uniform density matrices,

$$\rho_k \equiv 2^{-r} P_{\boldsymbol{y}}^{r,k}\ . \tag{2.42}$$

For convenience we will write the action of the classical shift $B_{\rm cl}$ on state ρ as $B_{\rm cl}\rho B_{\rm cl}^\dagger$. In this notation we have for the k-th iteration of the classical shift

$$B_{\rm cl}^k\, \rho_0 (B_{\rm cl}^\dagger)^k = \rho_k\ , \tag{2.43}$$

and therefore

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} B_{\mathrm{cl}}^k\, \rho_0 (B_{\mathrm{cl}}^\dagger)^k] = 1\ . \tag{2.44}$$

By normalization condition $\sum_{\boldsymbol{z}} \mathrm{Tr}\,[P_{\boldsymbol{z}}^{r,k} B_{\mathrm{cl}}^k\, \rho_0 (B_{\mathrm{cl}}^\dagger)^k] = 1$ this implies that the shifted state $B_{\mathrm{cl}}^k\, \rho_0 (B_{\mathrm{cl}}^\dagger)^k$ lies entirely in the shifted space defined by $P_{\boldsymbol{y}}^{r,k}$. We will prove that in the quantum case we have a similar expression

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} \hat{B}^k \rho_0 (\hat{B}^\dagger)^k] = 1 - O(\frac{r}{2^{r-k}}) \tag{2.45}$$

or, since k is bounded from above by $k_{\max}$, and $r = N - l$, where l is fixed,

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} \hat{B}^k \rho_0 (\hat{B}^\dagger)^k] = 1 - O(\frac{N}{2^N}) = 1 - O(\hbar \log \hbar)\ . \tag{2.46}$$

Comparing Eqs. (2.44) and (2.46), one sees that the coarse-grained quantum evolution approaches the shift-map behavior to any required accuracy as $\hbar \to 0$.

Equation (2.45) can be rewritten as

$$2^{-r} \sum_{|\boldsymbol{x}|=r} \mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} \hat{B}^k |\boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle \boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2| (\hat{B}^\dagger)^k] = 1 - O(\frac{r}{2^{r-k}})\ , \tag{2.47}$$

which is a sum of 2^r terms each of which is bounded from above as

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} \hat{B}^k |\boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle \boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2| (\hat{B}^\dagger)^k] \le 1\ . \tag{2.48}$$

Here, $\boldsymbol{y}^1\boldsymbol{y}^2 = \boldsymbol{y}$ and $|\boldsymbol{y}^2| = m$ as before. Equations (2.47) and (2.48) can be satisfied simultaneously only if the condition

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k} \hat{B}^k |\boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle \boldsymbol{x}\boldsymbol{y}^1.\boldsymbol{y}^2| (\hat{B}^\dagger)^k] = 1 - O(\frac{r}{2^{r-k}}) \tag{2.49}$$

holds for all $\boldsymbol{x}$ except for a fraction of order $O(r/2^{r-k})$, i.e., for all basis states $|\xi_{1:n}.\xi_{n+1:N}\rangle$ except for an exponentially small fraction. In other words, the property (2.49) holds for *typical* basis states.

An interesting feature of the quantum baker's map is that there are atypical basis states for which Eq. (2.49) does not hold. At the end of the next section we give an example of an atypical state $|\boldsymbol{x}^{\mathrm{atyp}}\boldsymbol{y}^1.\boldsymbol{y}^2\rangle$ for which

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,1} \hat{B} |\boldsymbol{x}^{\mathrm{atyp}}\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle \boldsymbol{x}^{\mathrm{atyp}}\boldsymbol{y}^1.\boldsymbol{y}^2| \hat{B}^\dagger] = \frac{\pi^2 + 8G}{2\pi^2} + O(4^{r-n}) + O(2^{-r})\ , \tag{2.50}$$

where $G \simeq 0.915965$ is Catalan's constant [27]. For sufficiently large $n-r$ and r, this expression is less than 0.872. This is an example where the quantum evolution in the limit $\hbar \to 0$ differs substantially from the classical evolution, already after the first iteration of the map. If, however, the initial state is a mixture in which atypical states have an exponentially small weight, such as ρ_0, the correspondence principle is obeyed.

2.5.2 Derivations and proofs

First iteration

In this section we prove the formula (2.37) for the matrix elements $C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N})$. Equation (2.40) for the case $n = N-1$ follows from almost identical arguments, and Eq. (2.39) for $n = 0$ is essentially trivial. A direct calculation yields

$$\begin{aligned} C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) &= \delta(\xi^0_{n+2}-\xi^1_{n+1})\delta(\xi^0_{n+3}-\xi^1_{n+2})\cdots\delta(\xi^0_N-\xi^1_{N-1}) \\ &\quad\times\sqrt{1/2}\{\delta(\xi^0_N)+\delta(\xi^0_N-1)\exp[2\pi i(0.\xi^0_1 1)]\} \\ &\quad\times\exp\{i\pi(0.\xi^0_{n+1:1}1-0.\xi^1_{n:1}1)\} \\ &\quad\times 1/2[1+\exp\{2\pi i(0.\xi^0_2\xi^0_1 1-0.\xi^1_1 1)\}]\times\cdots \\ &\quad\times 1/2[1+\exp\{2\pi i(0.\xi^0_{n+1:1}1-0.\xi^1_{n:1}1)\}]\;. \end{aligned} \tag{2.51}$$

Using the identity $1+e^{i\phi} = 2e^{i\phi/2}\cos(\phi/2)$ and noticing that $\delta(\xi^1_N)+\delta(\xi^1_N-1)\exp[2\pi i(0.\xi^0_1 1)]$
$=\exp[i\pi\xi^1_N(\xi^0_1+1/2)]$, we have

$$\begin{aligned} C_n^{\rm 1st}(\xi^0_{1:N},\xi^1_{1:N}) &= \sqrt{1/2}\exp[i\pi\xi^1_N(\xi^0_1+1/2)]\delta(\xi^0_{n+2:N}-\xi^1_{n+1:N-1}) \\ &\quad\times\exp[i\pi(0.\xi^0_{n+1:1}1-0.\xi^1_{n:1}1)] \\ &\quad\times\prod_{k=2}^{n+1}\cos[\pi(0.\xi^0_{k:1}1-0.\xi^1_{k-1:1}1)] \\ &\quad\times\prod_{k=2}^{n+1}\exp[i\pi(0.\xi^0_{k:1}1-0.\xi^1_{k-1:1}1)] \\ &= \sqrt{1/2}\exp[i\pi\xi^1_N(\xi^0_1+1/2)]\delta(\xi^0_{n+2:N}-\xi^1_{n+1:N-1}) \\ &\quad\times e^{i\phi_n}(\prod_{k=1}^{n}\cos\phi_k)(\prod_{k=1}^{n}e^{i\phi_k})\;, \end{aligned} \tag{2.52}$$

where

$$\phi_k \equiv \pi(0.\xi^0_{k+1:1}1 - 0.\xi^1_{k:1}1) \; . \tag{2.53}$$

To simplify Eq. (2.52), we first consider the products of cosines and exponents separately and then combine them to formulate the final result for the first iteration of the baker's map. Note that

$$\begin{aligned} 2\phi_k &= \pi(0.\xi^0_{k:1}1 - 0.\xi^1_{k-1:1}1) + \pi(\xi^0_{k+1} - \xi^1_k) \\ &= \phi_{k-1} + \pi(\xi^0_{k+1} - \xi^1_k) \; , \end{aligned} \tag{2.54}$$

so

$$\begin{aligned} \cos\phi_{k-1} &= \cos[2\phi_k + \pi(\xi^1_k - \xi^0_{k+1})] \\ &= (-1)^{\xi^1_k - \xi^0_{k+1}} \cos(2\phi_k) \; , \quad k \leq n \; . \end{aligned} \tag{2.55}$$

From Eq. (2.54), we have $2\phi_k = 4\phi_{k+1} \pmod{2\pi}$ and thus $2\phi_k = 2^{n+1-k}\phi_n \pmod{2\pi}$, so the previous formula can be rewritten as

$$\cos\phi_k = (-1)^{\xi^1_{k+1} - \xi^0_{k+2}} \cos(2^{n-k}\phi_n) \; , \quad k \leq n-1 \; . \tag{2.56}$$

Using this formula the product of cosines can be expressed as

$$\begin{aligned} \prod_{k=1}^{n} \cos\phi_k &= \cos\phi_n \prod_{k=1}^{n-1} \cos\phi_k \\ &= (-1)^{\sigma(\xi^1_{2:n}) - \sigma(\xi^0_{3:n+1})} \prod_{k=1}^{n} \cos[2^{-k}(2^n\phi_n)] \; , \end{aligned} \tag{2.57}$$

where $\sigma(\xi_{k:n}) \equiv \sum_{s=k}^{n} \xi_s$. It is easy to check by induction that

$$\prod_{k=0}^{n-1} \cos 2^k x = \frac{\sin 2^n x}{2^n \sin x} \; , \quad x \neq \pi j \; , \quad j = 0, \pm 1, \pm 2, \ldots \tag{2.58}$$

In our case

$$\prod_{k=1}^{n} \cos[2^{-k}(2^n\phi_n)] = \frac{\sin(2^n\phi_n)}{2^n \sin\phi_n} \; . \tag{2.59}$$

Putting everything together, the product of cosines becomes

$$\prod_{k=1}^{n} \cos\phi_k = (-1)^{\sigma(\xi^1_{2:n}) - \sigma(\xi^0_{3:n+1})} \frac{\sin(2^n\phi_n)}{2^n \sin\phi_n} \; , \tag{2.60}$$

where $\phi_n = \pi(0.\xi^0_{n+1:1}1 - 0.\xi^1_{n:1}1)$. Now we simplify the product of exponents in (2.52). Equation (2.54) implies

$$\phi_{n-k} = 2^k\phi_n + \sum_{s=1}^{k} 2^{k-s}\pi(\xi^1_{n+1-s} - \xi^0_{n+2-s}),\ k \geq 1\ , \tag{2.61}$$

so

$$\begin{aligned}
\sum_{k=1}^{n}\phi_k &= \phi_n + \sum_{k=1}^{n-1}\phi_{n-k} \\
&= \phi_n\sum_{k=0}^{n-1}2^k + \sum_{k=1}^{n-1}\sum_{s=1}^{k}2^{k-s}\pi(\xi^1_{n+1-s} - \xi^0_{n+2-s}) \\
&= [\phi_n(2^n-1) + \sum_{k=1}^{n-1}\pi(\xi^1_{n+1-k} - \xi^0_{n+2-k})](\mathrm{mod}\ 2\pi) \\
&= \{\phi_n(2^n-1) + \pi[\sigma(\xi^1_{2:n}) - \sigma(\xi^0_{3:n+1})]\}(\mathrm{mod}\ 2\pi)\ .
\end{aligned} \tag{2.62}$$

The product of exponents is thus given by

$$\prod_{k=1}^{n} e^{i\phi_k} = \exp(i\sum_{k=1}^{n}\phi_k) = (-1)^{\sigma(\xi^1_{2:n})-\sigma(\xi^0_{3:n+1})}\frac{\exp(i2^n\phi_n)}{\exp(i\phi_n)}\ . \tag{2.63}$$

Using (2.60) and (2.63) one can rewrite (2.52) as

$$C_n^{\mathrm{1st}}(\xi^0_{1:N},\xi^1_{1:N}) = \frac{(-1)^{\xi^1_N(\xi^0_1+1/2)}}{2^{n+1/2}}\delta(\xi^0_{n+2:N}-\xi^1_{n+1:N-1})\frac{\sin(2^n\phi_n)}{\sin\phi_n}\exp(i2^n\phi_n)\ . \tag{2.64}$$

Further simplification is possible due to the fact that $2^n\phi_n = 2\pi(0.\xi^0_2\xi^0_1 1 - 0.\xi^1_1 1)$ (mod 2π). The final result is

$$C_n^{\mathrm{1st}}(\xi^0_{1:N},\xi^1_{1:N}) = \Phi(\xi^0_1,\xi^1_N)\frac{\delta(\xi^0_{n+2:N}-\xi^1_{n+1:N-1})}{2^{n+1}\sin[\pi(0.\xi^0_{n+1:1}1 - 0.\xi^1_{n:1}1)]}\ , \tag{2.65}$$

where the phase factor Φ is given by

$$\begin{aligned}
\Phi(\xi^0_1,\xi^1_N) &= \sqrt{2}e^{i\pi\xi^1_N(\xi^0_1+1/2)}\sin(2^n\phi_n)\exp(i2^n\phi_n) \\
&= \frac{1}{\sqrt{2}}[i(-1)^{\xi^1_N} - (-1)^{\xi^0_1}]\ .
\end{aligned} \tag{2.66}$$

This formula is an exact expression for the matrix elements (2.36) of the quantum baker's map.

kth iteration

In this section we prove that for all $k \leq k_{\max}$,

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k}\hat{B}^k\rho_0(\hat{B}^\dagger)^k] = 1 - O(\frac{r}{2^{r-k}})\,, \tag{2.67}$$

where the projectors $P_{\boldsymbol{y}}^{r,k}$ and the density operators ρ_j are defined in Eqs. (2.41) and (2.42). The first step is to prove that

$$\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k}\hat{B}\rho_{k-1}\hat{B}^\dagger] = 1 - O(\frac{r}{2^{r-k}})\,. \tag{2.68}$$

By a direct calculation, we obtain

$$\begin{aligned}\mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,k}\hat{B}\rho_{k-1}\hat{B}^\dagger] &= 2^{k-r}\sum_{\substack{|\boldsymbol{\alpha}|=r-k+1\\|\boldsymbol{\beta}|=r-k}} |2^{n+1}\sin[2^{-(n-r+k)}\pi(0.\boldsymbol{\alpha}1 - 0.\boldsymbol{\beta}1)]|^{-2} \\ &\geq \frac{4}{\pi^2 2^{r-k}}\sum_{u=0}^{2^{r-k+1}-1}\sum_{v=0}^{2^{r-k}-1}(2u-4v-1)^{-2}\,, \end{aligned}\tag{2.69}$$

where the inequality $\sin^2 x \leq x^2$ was used. Let $L = r-k$, and let $Q(s)$ be the number of different pairs (u,v), $0 \leq u < 2^{L+1}$, $0 \leq v < 2^L$, for which $u - 2v = s$. It follows that

$$\begin{aligned}\sum_{u=0}^{2^{r-k+1}-1}\sum_{v=0}^{2^{r-k}-1}(2u-4v-1)^{-2} &= \sum_{s=-2(2^L-1)}^{2^{L+1}-1} Q(s)(2s-1)^{-2} \\ &= \sum_{s=1}^{2^{L+1}-1}\frac{Q(s)+Q(1-s)}{(2s-1)^2}\,. \end{aligned}\tag{2.70}$$

Using a simple counting argument based on the register principle [45], one can show that

$$Q(s) + Q(1-s) = 2^{L+1} - s + \frac{1}{2}[1-(-1)^s]\,, \tag{2.71}$$

from which one obtains

$$\begin{aligned}\sum_{s=1}^{2^{L+1}-1}\frac{Q(s)+Q(1-s)}{(2s-1)^2} &= \sum_{s=1}^{2^L-1}\frac{2^{L+1}-s}{(2s-1)^2} - \sum_{t=1}^{2^L-1}(4t-1)^{-2} \\ &= 2^L[\frac{\pi^2}{4} - O(\frac{L}{2^L})]\,, \end{aligned}\tag{2.72}$$

where we have used the relations

$$\sum_{s=1}^{2^L}(2s-1)^{-2} = \frac{\pi^2}{8} + O(2^{-L})\,, \qquad 2^{-L}\sum_{s=1}^{2^L}\frac{1}{2s-1} = O(\frac{L}{2^L})\,. \tag{2.73}$$

Combining Eqs. (2.69), (2.70) and (2.72), we obtain Eq. (2.68) as required. We now rewrite Eq. (2.68) in the symmetric form

$$\text{Tr } [\rho_k \hat{B} \rho_{k-1} \hat{B}^\dagger] = 2^{-r}[1 - O(\frac{r}{2^{r-k}})] \tag{2.74}$$

and introduce the distance measure between density matrices induced by the Euclidian norm [33],

$$d(\rho, \rho') \equiv \sqrt{\text{Tr } (\rho - \rho')^2} \; . \tag{2.75}$$

This distance measure is unitarily invariant and obeys the triangle inequality. We will now prove that (2.74) and (2.75) imply

$$d(\rho_k, \hat{B}^k \rho_0 [\hat{B}^\dagger]^k) = O(2^{\frac{k}{2}-r}\sqrt{r}) \; . \tag{2.76}$$

Using the cyclic property of the trace, we have

$$d(\rho_k, \hat{B} \rho_{k-1} \hat{B}^\dagger) = \sqrt{\text{Tr } \rho_k^2 + \text{Tr } \rho_{k-1}^2 - 2\text{Tr } (\rho_k \hat{B} \rho_{k-1} \hat{B}^\dagger)} \; . \tag{2.77}$$

Since $\text{Tr } \rho_k^2 = 2^r/2^{2r} = 2^{-r}$ for any k,

$$[d(\rho_k, \hat{B} \rho_{k-1} \hat{B}^\dagger)]^2 = 2^{-r+1} - 2\text{Tr } (\rho_k \hat{B} \rho_{k-1} \hat{B}^\dagger) \; , \tag{2.78}$$

which, together with Eq. (2.74), implies

$$d(\rho_k, \hat{B} \rho_{k-1} \hat{B}^\dagger) = O(2^{\frac{k}{2}-r}\sqrt{r}) \; . \tag{2.79}$$

The case $k = 1$ of (2.76) follows directly from (2.79). Assuming that (2.76) is true for a given value of k and using the unitary invariance of the distance (2.75), we have

$$d(\hat{B} \rho_k \hat{B}^\dagger, \hat{B}^{k+1} \rho_0 [\hat{B}^\dagger]^{k+1}) = O(2^{\frac{k}{2}-r}\sqrt{r}) \; . \tag{2.80}$$

Substituting $k + 1$ for k in Eq. (2.79), we get

$$d(\rho_{k+1}, \hat{B} \rho_k \hat{B}^\dagger) = O(2^{\frac{1}{2}(k+1)-r}\sqrt{r}) \; . \tag{2.81}$$

Using the triangle inequality for the distance measure (2.75), it follows from (2.80) and (2.81) that

$$d(\rho_{k+1}, \hat{B}^{k+1} \rho_0 [\hat{B}^\dagger]^{k+1}) \;\; = \;\; O(2^{\frac{k}{2}-r}\sqrt{r}) + O(2^{\frac{1}{2}(k+1)-r}\sqrt{r})$$

$$= \quad O(2^{\frac{1}{2}(k+1)-r}\sqrt{r}) \ . \tag{2.82}$$

By induction, this completes the proof of (2.76) for any $k \leq k_{\max}$. On the other hand

$$d(\rho_k, \hat{B}^k\rho_0[\hat{B}^\dagger]^k) = \sqrt{\text{Tr } \rho_k^2 + \text{Tr } \rho_0^2 - 2\text{Tr } (\rho_k\hat{B}^k\rho_0[\hat{B}^\dagger]^k)} \ , \tag{2.83}$$

hence using Eq. (2.76) it follows that

$$\sqrt{2^{1-r} - 2\text{Tr } (\rho_k\hat{B}^k\rho_0[\hat{B}^\dagger]^k)} = O(2^{\frac{k}{2}-r}\sqrt{r}) \ , \tag{2.84}$$

and finally

$$\text{Tr } (\rho_k\hat{B}^k\rho_0[\hat{B}^\dagger]^k) = 2^{-r}[1 - O(\frac{r}{2^{r-k}})] \ , \tag{2.85}$$

which is equivalent to (2.67) as required.

Atypical initial states

In this section, we show that the state $|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle$, where 0^r is a string of r zeros, is an atypical state in the sense of the discussion at the end of Sec. 2.5.1, i.e., we show that the state $|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle$ satisfies Eq. (2.50). A direct calculation gives

$$\begin{aligned} \text{Tr } [P_{\boldsymbol{y}}^{r,1}\hat{B}|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle 0^r\boldsymbol{y}^1.\boldsymbol{y}^2|\hat{B}^\dagger] \ &= \ \sum_{|\boldsymbol{x}|=r-1}\sum_{g=0}^{1}|C_n^{\text{1st}}(0^r\boldsymbol{y}^1\boldsymbol{y}^2, \boldsymbol{x}\boldsymbol{y}^1\boldsymbol{y}^2 g)|^2 \\ &= \ \frac{8}{\pi^2}\sum_{v=0}^{2^{r-1}-1}\frac{1+O(4^{r-n})}{(4v+1)^2} \ . \end{aligned} \tag{2.86}$$

Substituting $t = 2v$, we have

$$\begin{aligned} \text{Tr } [P_{\boldsymbol{y}}^{r,1}\hat{B}|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle 0^r\boldsymbol{y}^1.\boldsymbol{y}^2|\hat{B}^\dagger] &= \frac{8+O(4^{r-n})}{\pi^2}\sum_{t=0}^{2^r}\frac{1+(-1)^t}{2(2t+1)^2} \\ &= \frac{4+O(4^{r-n})}{\pi^2}\left(\sum_{s=1}^{2^r+1}(2s-1)^{-2} + \sum_{t=0}^{2^r}\frac{(-1)^t}{(2t+1)^2}\right) \ . \end{aligned} \tag{2.87}$$

Using Eq. (2.73) and the series representation of Catalan's constant $G \simeq 0.915965$ [27],

$$G = \sum_{t=0}^{\infty}\frac{(-1)^t}{(2t+1)^2} \ , \tag{2.88}$$

it follows that

$$\begin{aligned} \mathrm{Tr}\,[P_{\boldsymbol{y}}^{r,1}\hat{B}|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle\langle 0^r\boldsymbol{y}^1.\boldsymbol{y}^2|\hat{B}^\dagger] &= \frac{\pi^2+8G}{2\pi^2}+O(4^{r-n})+O(2^{-r}) \\ &\simeq 0.871+O(4^{r-n})+O(2^{-r})\,. \end{aligned} \tag{2.89}$$

Since one can treat $n-r$ and r as independent variables, this expression can be made smaller than 0.872 by choosing $n-r$ and r large enough. For the initial state $|0^r\boldsymbol{y}^1.\boldsymbol{y}^2\rangle$, the asymptotic relation (2.49) is thus violated.

2.6 Decoherent histories for the quantum baker's map

2.6.1 Background

Decoherent histories approach

In this subsection we review the terminology of the decoherent histories approach [28, 46, 24]. Although the definitions given below are usually presented in a more general form, we do not aim for full generality here. Rather we introduce a notation which is specific for our applications, i.e. the quantum baker's map.

Let $\mathbb{A}^\gamma$ be a set of binary strings of length γ

$$\mathbb{A}^\gamma = \left\{\boldsymbol{\alpha}\middle|\ \boldsymbol{\alpha}\equiv\alpha_1\ldots\alpha_\gamma : \alpha_i\in\{0,1\}\right\}\,. \tag{2.90}$$

At each moment (of time) t we choose a set of projectors $\mathbb{P}_\gamma^t \equiv \{\pi_{\boldsymbol{\alpha}}^t|\,\boldsymbol{\alpha}\in\mathbb{A}^\gamma\}$. Each such set represents a set of events labeled by strings from $\mathbb{A}^\gamma$. We use a superscript t to indicate that at each moment the set of events $\mathbb{P}_\gamma^t$ can be chosen independently in contrast to the set of labels $\mathbb{A}^\gamma$ which is fixed.

In the Schrödinger picture, a coarse-grained[9] history $h_{\vec{\boldsymbol{\alpha}}(k)}$ of length k is defined as an ordered (in time) sequence of k projectors

$$h_{\vec{\boldsymbol{\alpha}}(k)} \equiv (\pi_{\boldsymbol{\alpha}^1}^1,\ldots,\pi_{\boldsymbol{\alpha}^t}^t,\ldots,\pi_{\boldsymbol{\alpha}^k}^k)\,, \tag{2.91}$$

[9] A fine-grained history is a specific case of the definition (2.91), where all the projectors $\{\pi_{\boldsymbol{\alpha}^k}^k\}$ are one-dimensional.

where $\vec{\boldsymbol{\alpha}}(k)$ is the ordered sequence of labels $(\boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^k)$, where $\boldsymbol{\alpha}^t \in \mathbb{X}$ for $1 \leq t \leq k$, each of which specifies a particular projector $\pi^t_{\boldsymbol{\alpha}^t} \in \mathbb{P}^t_\gamma$. Two histories $h_{\vec{\boldsymbol{\alpha}}(k)}$ and $h_{\vec{\boldsymbol{\beta}}(k)}$ are said to be mutually exclusive (or disjoint) if at one of the times $t \in \{1, \ldots, k\}$ we have $\pi^t_{\boldsymbol{\alpha}^t}\pi^t_{\boldsymbol{\beta}^t} = 0$. A set of mutually exclusive (i.e. pairwise disjoint) histories $\{h_{\vec{\boldsymbol{\alpha}}(k)} |\, k \text{ fixed}\}$ is called exhaustive if at each moment t the set of projectors $\mathbb{P}^t_\gamma$ is complete: $\sum_{\boldsymbol{\alpha}^t \in \mathbb{X}} \pi^t_{\boldsymbol{\alpha}^t} = \mathbb{1}$ for all $t \in \{1, \ldots, k\}$. The decoherence functional is then defined as

$$\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] \equiv \mathrm{Tr}\,[\pi^k_{\boldsymbol{y}^k}\hat{B}\pi^{k-1}_{\boldsymbol{y}^{k-1}}\hat{B}\cdots\pi^1_{\boldsymbol{y}^1}\hat{B}\rho_{\boldsymbol{x}}\hat{B}^\dagger\pi^1_{\boldsymbol{z}^1}\cdots\hat{B}^\dagger\pi^{k-1}_{\boldsymbol{z}^{k-1}}\hat{B}^\dagger\pi^k_{\boldsymbol{z}^k}]\ , \quad (2.92)$$

where $\rho_{\boldsymbol{x}}$ denotes the initial state and $\hat{B}$ is a unitary dynamics (evolution operator) of the system. An exhaustive set of mutually exclusive histories is said to decohere if

$$\mathrm{Re}\,\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] = 0, \qquad h_{\vec{\boldsymbol{y}}(k)} \neq h_{\vec{\boldsymbol{z}}(k)}\ . \quad (2.93)$$

This is the so-called "weak decoherence" condition [25]. For every history $h_{\vec{\boldsymbol{y}}(k)}$ we define a nonnegative number

$$p[\,h_{\vec{\boldsymbol{y}}(k)}\,] \equiv \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{y}}(k)}\,]\ . \quad (2.94)$$

These numbers can be regarded as probabilities if in addition to the above definition we have

$$\begin{aligned} p[(\pi^1_{\boldsymbol{\alpha}^1}, \ldots, \pi^t_{\boldsymbol{\alpha}^t}, \ldots, \pi^k_{\boldsymbol{\alpha}^k})] = {} & p[(\pi^1_{\boldsymbol{\alpha}^1}, \ldots, \pi^t_{\boldsymbol{\alpha}^t_1}, \ldots, \pi^k_{\boldsymbol{\alpha}^k})] \\ & + p[(\pi^1_{\boldsymbol{\alpha}^1}, \ldots, \pi^t_{\boldsymbol{\alpha}^t_2}, \ldots, \pi^k_{\boldsymbol{\alpha}^k})] \end{aligned} \quad (2.95)$$

if $\pi^t_{\boldsymbol{\alpha}^t} = \pi^t_{\boldsymbol{\alpha}^t_1} + \pi^t_{\boldsymbol{\alpha}^t_2}$ and $\pi^t_{\boldsymbol{\alpha}^t_1}\pi^t_{\boldsymbol{\alpha}^t_2} = 0$ for any $t \in \{1, \ldots, k\}$. This is the ordinary sum rule for the probabilities of mutually exclusive events. It turns out that the weak decoherence condition introduced above is necessary and sufficient for this sum rule to hold. This can be understood using an illustrative example of a pure initial state $\rho_{\boldsymbol{x}} = |\Psi\rangle\langle\Psi|$ as follows. We have in this case

$$p[(\pi^1_{\boldsymbol{y}^1}, \pi^2_{\boldsymbol{y}^2}, \pi^3_{\boldsymbol{y}^3})] = \langle\Psi|\hat{B}^\dagger\pi^1_{\boldsymbol{y}^1}\hat{B}^\dagger\pi^2_{\boldsymbol{y}^2}\hat{B}^\dagger\pi^3_{\boldsymbol{y}^3}\hat{B}\pi^2_{\boldsymbol{y}^2}\hat{B}\pi^1_{\boldsymbol{y}^1}\hat{B}|\Psi\rangle\ , \quad (2.96)$$

and

$$p[(\pi^1_{\boldsymbol{y}^1}, \mathbb{1}, \pi^3_{\boldsymbol{y}^3})] = \langle\Psi|\hat{B}^\dagger\pi^1_{\boldsymbol{y}^1}\hat{B}^\dagger\mathbb{1}\hat{B}^\dagger\pi^3_{\boldsymbol{y}^3}\hat{B}\mathbb{1}\hat{B}\pi^1_{\boldsymbol{y}^1}\hat{B}|\Psi\rangle\ . \quad (2.97)$$

Using the completeness of the projectors $\pi^2_{\boldsymbol{y}^2}$: $\mathbb{1} = \sum_{\boldsymbol{y}^2} \pi^2_{\boldsymbol{y}^2}$, we have

$$\begin{aligned} p[(\pi^1_{\boldsymbol{y}^1}, \mathbb{1}, \pi^3_{\boldsymbol{y}^3})] &= \sum_{\boldsymbol{y}^2, \boldsymbol{z}^2} \langle\Psi|\hat{B}^\dagger \pi^1_{\boldsymbol{y}^1} \hat{B}^\dagger \pi^2_{\boldsymbol{y}^2} \hat{B}^\dagger \pi^3_{\boldsymbol{y}^3} \hat{B} \pi^2_{\boldsymbol{z}^2} \hat{B} \pi^1_{\boldsymbol{y}^1} \hat{B}|\Psi\rangle \\ &= \sum_{\boldsymbol{y}^2} p[(\pi^1_{\boldsymbol{y}^1}, \pi^2_{\boldsymbol{y}^2}, \pi^3_{\boldsymbol{y}^3})] + \\ &\quad \mathrm{Re} \sum_{\boldsymbol{y}^2 \neq \boldsymbol{z}^2} \mathcal{D}[|\Psi\rangle\langle\Psi|, (\pi^1_{\boldsymbol{y}^1}, \pi^2_{\boldsymbol{y}^2}, \pi^3_{\boldsymbol{y}^3}), (\pi^1_{\boldsymbol{y}^1}, \pi^2_{\boldsymbol{z}^2}, \pi^3_{\boldsymbol{y}^3})\,], \end{aligned} \tag{2.98}$$

where we used hermicity of the decoherence functional. This expression gives the classical sum rule for probabilities provided that the second sum (over the off-diagonal elements) vanishes. This is equivalent to the weak decoherence requirement. Realistic mechanisms of decoherence often lead to the stronger condition of "medium decoherence":

$$\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] = 0, \qquad h_{\vec{\boldsymbol{y}}(k)} \neq h_{\vec{\boldsymbol{z}}(k)}\ . \tag{2.99}$$

We will also need the following inequalities

$$\left|\, \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] \,\right|^2 \leq \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{y}}(k)}\,] \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{z}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] \tag{2.100}$$

and

$$\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] \leq \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k-1)}, h_{\vec{\boldsymbol{z}}(k-1)}\,]\ , \tag{2.101}$$

where we assume that the first $k-1$ projectors in the history $h_{\vec{\boldsymbol{y}}(k)}$ coincide with those in $h_{\vec{\boldsymbol{y}}(k-1)}$:

$$h_{\vec{\boldsymbol{y}}(k)} = (\pi^1_{\boldsymbol{y}^1}, \ldots, \pi^k_{\boldsymbol{y}^k}) \quad \Rightarrow \quad h_{\vec{\boldsymbol{y}}(k-1)} \equiv (\pi^1_{\boldsymbol{y}^1}, \ldots, \pi^{k-1}_{\boldsymbol{y}^{k-1}})\ , \tag{2.102}$$

and similarly for $h_{\vec{\boldsymbol{z}}(k)}$. To prove the first of these inequalities we fix any decomposition of $\rho_{\boldsymbol{x}}$ in terms of a complete set of mutually orthogonal pure states $\{|\psi_k\rangle\}$

$$\rho_{\boldsymbol{x}} = \sum_s p_s |\psi_s\rangle\langle\psi_s|\ . \tag{2.103}$$

Then we can write

$$\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] = \sum_s p_s \mathrm{Tr}\,[C^\dagger_{\vec{\boldsymbol{z}}(k)} C_{\vec{\boldsymbol{y}}(k)} |\psi_s\rangle\langle\psi_s|]\,, \tag{2.104}$$

where we used the notation $C_{\vec{\boldsymbol{y}}(k)} \equiv \pi^k_{\boldsymbol{y}^k} \cdots \hat{B} \pi^1_{\boldsymbol{y}^1} \hat{B}$ and the cyclic property of the trace. Writing explicitly the trace in terms of $\{|\psi_j\rangle\}$ we have

$$\begin{aligned} \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] &= \sum_s p_s \sum_j \langle\psi_j|C^\dagger_{\vec{\boldsymbol{z}}(k)} C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\langle\psi_s|\psi_j\rangle \\ &= \sum_s p_s \langle\psi_s|C^\dagger_{\vec{\boldsymbol{z}}(k)} C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\,. \end{aligned} \tag{2.105}$$

This implies that

$$\begin{aligned} \left|\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,]\right| &\leq \sum_s p_s \left|\langle\psi_s|C^\dagger_{\vec{\boldsymbol{z}}(k)} C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\right| \\ &\leq \sum_s p_s \left|C_{\vec{\boldsymbol{z}}(k)}|\psi_s\rangle\right| \cdot \left|C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\right|\,, \end{aligned} \tag{2.106}$$

where we used the Cauchy-Schwarz inequality. According to (2.105) we furthermore have

$$\mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{\alpha}}(k)}, h_{\vec{\boldsymbol{\alpha}}(k)}\,] = \sum_s p_s |C_{\vec{\boldsymbol{\alpha}}(k)}|\psi_s\rangle|^2\,, \qquad \vec{\boldsymbol{\alpha}}(k) = \vec{\boldsymbol{y}}(k) \;\text{ or }\; \vec{\boldsymbol{z}}(k)\,. \tag{2.107}$$

Rewrite the inequality

$$\sum_{s,t} (Y_s Z_t - Y_t Z_s)^2 \geq 0 \tag{2.108}$$

as

$$\sum_s Y_s^2 \cdot \sum_t Z_t^2 \geq \sum_s Y_s Z_s \cdot \sum_t Y_t Z_t \tag{2.109}$$

and set $Y_s = \sqrt{p_s}\left|C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\right|$, $Z_t = \sqrt{p_t}\left|C_{\vec{\boldsymbol{z}}(k)}|\psi_t\rangle\right|$. The resulting inequality reads

$$\sum_s p_s \left|C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\right|^2 \cdot \sum_t p_t \left|C_{\vec{\boldsymbol{z}}(k)}|\psi_t\rangle\right|^2 \geq \left(\sum_s p_s \left|C_{\vec{\boldsymbol{y}}(k)}|\psi_s\rangle\right| \cdot \left|C_{\vec{\boldsymbol{z}}(k)}|\psi_s\rangle\right|\right)^2 \tag{2.110}$$

Using (2.106) and (2.107) we therefore have Eq. (2.100) as required.

The second of the inequalities (2.101) can be proven directly from definition (2.92) by rewriting it as

$$\begin{aligned} \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k)}, h_{\vec{\boldsymbol{z}}(k)}\,] &= \mathrm{Tr}\,[(\pi^k_{\boldsymbol{z}^k}\pi^k_{\boldsymbol{y}^k})\hat{B}\pi^{k-1}_{\boldsymbol{y}^{k-1}}\hat{B}\cdots\pi^1_{\boldsymbol{y}^1}\hat{B}\rho_{\boldsymbol{x}}\hat{B}^\dagger\pi^1_{\boldsymbol{z}^1}\cdots\hat{B}^\dagger\pi^{k-1}_{\boldsymbol{z}^{k-1}}\hat{B}^\dagger] \\ &\leq \mathrm{Tr}\,[\hat{B}\pi^{k-1}_{\boldsymbol{y}^{k-1}}\hat{B}\cdots\pi^1_{\boldsymbol{y}^1}\hat{B}\rho_{\boldsymbol{x}}\hat{B}^\dagger\pi^1_{\boldsymbol{z}^1}\cdots\hat{B}^\dagger\pi^{k-1}_{\boldsymbol{z}^{k-1}}\hat{B}^\dagger] \\ &= \mathrm{Tr}\,[\pi^{k-1}_{\boldsymbol{y}^{k-1}}\hat{B}\cdots\pi^1_{\boldsymbol{y}^1}\hat{B}\rho_{\boldsymbol{x}}\hat{B}^\dagger\pi^1_{\boldsymbol{z}^1}\cdots\hat{B}^\dagger\pi^{k-1}_{\boldsymbol{z}^{k-1}}] \\ &= \mathcal{D}[\rho_{\boldsymbol{x}},\, h_{\vec{\boldsymbol{y}}(k-1)}, h_{\vec{\boldsymbol{z}}(k-1)}\,]\,, \end{aligned} \tag{2.111}$$

where we used the cyclic property of the trace and the fact that projectors decrease the trace of hermitian matrices.

Simplified notation for the quantum baker's map

In order to simplify further arguments we introduce some new notation. First, we omit the dot:

$$|\xi_{1:N}\rangle \equiv |\xi_{1:n}.\xi_{n+1:N}\rangle\,, \tag{2.112}$$

assuming that n and N are given constants throughout the rest of this chapter. Secondly, we introduce a more general notation for the projectors of the type (2.41). We define

$$P_{\boldsymbol{y}}^{(l,r)} \equiv \sum_{a_{1:l}} \sum_{b_{1:r}} |a_1 a_2 \ldots a_l \;\boldsymbol{y}\; b_1 \ldots b_r\rangle\langle a_1 a_2 \ldots a_l \;\boldsymbol{y}\; b_1 \ldots b_r|\,, \tag{2.113}$$

where, for clarity, we used bold string notation for $\boldsymbol{y} = y_{1:N-l-r}$, which enumerates the projectors, to separate it from the dummy variables. In order to make the above definition even more transparent we can write each $P_{\boldsymbol{y}}^{(l,r)}$ as a diagram

$$P_{\boldsymbol{y}}^{(l,r)} \equiv (\underbrace{\square\square\ldots\square}_{l}\;\boldsymbol{y}\;\underbrace{\square\square\ldots\square}_{r})\,, \tag{2.114}$$

where the empty boxes indicate l leftmost and r rightmost bits which are coarse grained over. In what follows we would also assume for simplicity that $l < n$ and $r < N - n$. In this case l and r acquire a more specific meaning as the number of "momentum" and "position" bits that are ignored in the coarse-graining. Using the diagrammatic notation, the projectors $P_{\boldsymbol{y}}^{l,k}$ introduced in (2.41) can now be rewritten as

$$P_{\boldsymbol{y}}^{l,k} = P_{\boldsymbol{y}}^{(l-k,k)} = (\underbrace{\square\square\ldots\square}_{l-k}\;\boldsymbol{y}\;\underbrace{\square\square\ldots\square}_{k})\,. \tag{2.115}$$

Finally, we introduce the characteristic function

$$D_{\boldsymbol{y};\rho}^{(l-k,k)} \equiv \mathrm{Tr}\,[P_{\boldsymbol{y}}^{(l-k,k)} \hat{B}^k \rho\, (\hat{B}^\dagger)^k]\,. \tag{2.116}$$

The main result of section 2.5.1 then takes the form

$$D_{\boldsymbol{y};\rho_{\boldsymbol{x}}^{(l,0)}}^{(l-k,k)} = \delta_{\boldsymbol{x}}^{\boldsymbol{y}} - O(\frac{l}{2^{l-k}})\,, \tag{2.117}$$

where $\rho_{\boldsymbol{x}}^{(l,0)} \equiv 2^{-l} P_{\boldsymbol{x}}^{(l,0)}$. This result was used to show that the coarse-grained quantum evolution approaches the shift-like symbolic behaviour of the classical baker's map to any required accuracy. We will show that this result also implies the existence of a set of decoherent histories.

2.6.2 Types of histories

Depending on the type of projectors constituting each set $\mathbb{P}_\gamma^t$ $(t = 1, \ldots, k)$, we can have different types of histories of length k. Here we list four different types of histories that will be considered later in the chapter. In what follows we shall always assume the initial state

$$\begin{aligned} \rho_{\boldsymbol{x}}^{(l,r)} &= 2^{-(l+r)} P_{\boldsymbol{x}}^{(l,r)} \\ &= 2^{-(l+r)} \, (\underbrace{\square\square\ldots\square}_{l} \; \boldsymbol{x} \; \underbrace{\square\square\ldots\square}_{r}) \, . \end{aligned} \tag{2.118}$$

Definition 1

Type-1 history

$$h^1_{\vec{\boldsymbol{y}}(k)} \equiv (P_{\boldsymbol{y}^1}^{(l-1,r+1)}, P_{\boldsymbol{y}^2}^{(l-2,r+2)}, \ldots, P_{\boldsymbol{y}^k}^{(l-k,r+k)}) \, . \tag{2.119}$$

Histories of this type are motivated by the symbolic dynamics of the quantum baker's map considered earlier in this thesis. We see that such histories consist of projectors that are "shifted" one bit per step relative to the initial state (2.118). Using our diagrammatic notation we have

$$\begin{aligned} h^1_{\vec{\boldsymbol{y}}(k)} \equiv \; (& \underbrace{\square\ldots\square\square\square}_{l-1} \; \boldsymbol{y}_1 \; \underbrace{\square\ldots\square}_{r+1} \, , \\ & \underbrace{\square\ldots\square\square}_{l-2} \; \boldsymbol{y}_2 \; \underbrace{\square\square\ldots\square}_{r+2} \, , \; \ldots \, , \\ & \underbrace{\square\ldots\square}_{l-k} \; \boldsymbol{y}_k \; \underbrace{\square\square\square\ldots\square}_{r+k} \;) \, . \end{aligned} \tag{2.120}$$

As an auxiliary tool we will also use a similar, but more coarse-grained type of histories:

Definition 2

Type-2 history

$$h^2_{\boldsymbol{y}} \equiv (\underbrace{\mathbb{1}, \ldots, \mathbb{1}}_{k-1 \text{ times}}, P_{\boldsymbol{y}}^{(l-k,r+k)}) \, . \tag{2.121}$$

This type of histories can be obtained from the type-1 histories by summing over the first $k-1$ projectors. Physically, this would correspond to resolving alternative histories only after the k-th iteration of the map.

Type-1 and type-2 histories are particularly useful for seeing how close the map is to a classical shift. Histories of this type, however, are somewhat artificial because the level of coarse-graining over the "momentum" and "position" changes in time: after the k-th step only $l-k$ momentum bits are coarse-grained over compared to l in the initial state $\rho_{\boldsymbol{x}}^{(l,r)}$; as for the position, $r+k$ bits are coarse-grained over after the k-th step compared to r in the initial state. We therefore introduce a more realistic type of histories with a constant level of coarse-graining over both "momentum" and "position" bits:

Definition 3

Type-3 history

$$\begin{aligned} h^3_{\tilde{\boldsymbol{y}}(k)} &\equiv (P^{(l,r)}_{\boldsymbol{y}^1}, P^{(l,r)}_{\boldsymbol{y}^2}, \ldots, P^{(l,r)}_{\boldsymbol{y}^k}) \\ &= \Big(\underbrace{\square\square\ldots\square}_{l}\, \boldsymbol{y}_1\, \underbrace{\square\square\ldots\square}_{r}, \\ &\quad \underbrace{\square\square\ldots\square}_{l}\, \boldsymbol{y}_2\, \underbrace{\square\square\ldots\square}_{r}, \ \ldots\ , \\ &\quad \underbrace{\square\square\ldots\square}_{l}\, \boldsymbol{y}_k\, \underbrace{\square\square\ldots\square}_{r} \Big) \, . \end{aligned} \tag{2.122}$$

In this definition we use the same set of projectors $\{P^{(l,r)}_{\boldsymbol{y}}\}$ at all times $t = 1, \ldots, k$ with the same coarse-graining parameters l and r as in the initial state $\rho_{\boldsymbol{x}}^{(l,r)}$. Analogously to the case of type-2 histories we introduce more coarse-grained histories by summing over the first $k-1$ projectors in (2.122):

Definition 4

Type-4 history

$$h^4_{\boldsymbol{y}} \equiv (\underbrace{\mathbb{1}, \ldots, \mathbb{1}}_{k-1 \text{ times}}, P^{(l,r)}_{\boldsymbol{y}}) \, . \tag{2.123}$$

Within each type the histories are mutually exclusive and form an exhaustive set. In the case of type-1 histories, for instance, we have $\mathbb{A}^{\gamma} = \mathbb{A}^{N-l-r} = \big\{\boldsymbol{\alpha} \big|\ |\boldsymbol{\alpha}| = N-l-r\big\}$ and each projector $\pi^t_{\boldsymbol{\alpha}^t}$ in (2.91) is identified with $P^{(l-t,r+t)}_{\boldsymbol{\alpha}^t}$, so that for any $t = 1, \ldots, k$ the set $\mathbb{P}^t_{\gamma} = \big\{P^{(l-t,r+t)}_{\boldsymbol{\alpha}^t} \big|\ \boldsymbol{\alpha}^t \in \mathbb{A}^{\gamma}\big\}$ is a complete set of mutually

orthogonal projectors:

$$\sum_{\boldsymbol{\alpha}\in\mathbb{X}} P_{\boldsymbol{\alpha}}^{(l-t,r+t)} = \mathbb{1} \quad \text{and} \quad P_{\boldsymbol{\alpha}}^{(l-t,r+t)} P_{\boldsymbol{\beta}}^{(l-t,r+t)} = 0, \quad \boldsymbol{\alpha} \neq \boldsymbol{\beta} \,. \tag{2.124}$$

2.6.3 Decoherence results

In this section we estimate the elements of the decoherence functional for the initial state $\rho_{\boldsymbol{x}}^{(l,r)}$. In the limit of large l we establish the medium decoherence condition and calculate the diagonal entries of the decoherence functional for different types of histories. The results are summarized in the form of theorems each followed by a short discussion. The proofs are deferred to section 2.7. All conventions and assumptions listed in the last part of section 2.6.1 are implicitly used throughout the arguments.

Theorem 1

For a fixed number of iterations k, in the limit of large l any two type-2 histories $h_{\boldsymbol{y}}^2$ and $h_{\boldsymbol{z}}^2$ satisfy the asymptotic relation

$$\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)},\, h_{\boldsymbol{y}}^2, h_{\boldsymbol{z}}^2\,] = \mathrm{Tr}\,[P_{\boldsymbol{y}}^{(l-k,r+k)} \hat{B}^k \rho_{\boldsymbol{x}}^{(l,r)} (\hat{B}^\dagger)^k P_{\boldsymbol{z}}^{(l-k,r+k)}] = \delta_{\boldsymbol{y}}^{\boldsymbol{z}}\,[\,\delta_{\boldsymbol{x}}^{\boldsymbol{y}} - O(\frac{l+r}{2^{l-k}})] \,. \tag{2.125}$$

This theorem summarizes two important properties of the decoherence functional for type-2 histories. Firstly, we have that the off-diagonal elements of the decoherence functional are zero

$$\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)},\, h_{\boldsymbol{y}}^2, h_{\boldsymbol{z}}^2\,] = 0\,, \quad h_{\boldsymbol{y}}^2 \neq h_{\boldsymbol{z}}^2\,, \tag{2.126}$$

which immediately follows from the mutual orthogonality of the projectors $\{P_{\boldsymbol{y}}^{(l-k,r+k)}\}$ and the cyclic property of the trace. Secondly, we see that there is only one diagonal element which is close to one:

$$\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)},\, h_{\boldsymbol{x}}^2, h_{\boldsymbol{x}}^2\,] = 1 - O(\frac{l+r}{2^{l-k}})\,. \tag{2.127}$$

Relation (2.126) implies the medium decoherence condition, and therefore the diagonal elements of the decoherence functional can be interpreted as probabilities

of individual histories. Equation (2.127) identifies the "dominant history" – the history which in the limit of large l can be assigned unit probability. The "error" term $O(\frac{l+r}{2^{l-k}})$ arises in the proof of the theorem as a consequence of the estimations performed in the derivations of Eq. (2.117). We therefore acknowledge that the bound on the absolute value of this error term can probably be improved.

Theorem 2

For a fixed number of iterations k the off-diagonal elements of the decoherence functional for type-1 histories can be made arbitrarily small by choosing sufficiently large l. More precisely

$$\left|\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{z}}(k)}]\right| = O(\frac{l+r}{2^{l-2k}}) \qquad h^1_{\vec{\boldsymbol{y}}} \neq h^1_{\vec{\boldsymbol{z}}} \; . \tag{2.128}$$

It follows from this theorem that in the limit of large l type-1 histories satisfy the "medium decoherence" condition, and therefore, within this limit the diagonal elements of the decoherence functional define consistent probabilities. The next theorem estimates these probabilities:

Theorem 3

For a fixed number of iterations k and for sufficiently large l the diagonal elements of the decoherence functional for type-1 histories approach either close to one or to zero. More precisely

$$\begin{aligned}
&\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)},\, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{y}}(k)}\,] \\
&\qquad \equiv \mathrm{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k} \hat{B} \cdots P^{(l-1,r+1)}_{\boldsymbol{y}^1} \hat{B} \rho_{\boldsymbol{x}}^{(l,r)} \hat{B}^\dagger P^{(l-1,r+1)}_{\boldsymbol{y}^1} \cdots \hat{B}^\dagger P^{(l-k,r+k)}_{\boldsymbol{y}^k}] \\
&\qquad = \delta^{\boldsymbol{x}}_{\boldsymbol{y}^k} \delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}} \cdots \delta^{\boldsymbol{x}}_{\boldsymbol{y}^1} + O(\frac{l+r}{2^{l-2k}}) \; .
\end{aligned} \tag{2.129}$$

Similar to the case of type-2 histories, we have that, except one dominant history, all type-1 histories have nearly zero probabilities. Further comparison of the results for type-1 and type-2 histories reveals a noticeable difference in the order of the error terms: $O(\frac{l+r}{2^{l-2k}})$ for type-1 histories and $O(\frac{l+r}{2^{l-k}})$ for type-2 histories. We do not have any evidence of its importance: it may well be just a consequence of the particular choice of the methods used in the proofs of the theorems.

Let us now move on to the second group of results, those for type-3 and type-4 histories. For the same reasons as in the case of type-2 histories, the medium decoherence condition for type-4 histories is satisfied exactly. We can therefore proceed with the estimation of the probabilities associated with the individual histories:

Theorem 4

Fix two strings $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *of the same length:* $|\boldsymbol{x}| = |\boldsymbol{y}| = c$. *For any two strings* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\beta}$ *such that* $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = k$, *where* k *is a fixed number of iterations, we have in the limit of large* l*:*

$$\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{\alpha x}}, h^4_{\boldsymbol{y\beta}}, h^4_{\boldsymbol{y\beta}}] = \mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y\beta}} \hat{B}^k \rho^{(l,r)}_{\boldsymbol{\alpha x}} (\hat{B}^\dagger)^k] = 2^{-k} \delta^{\,\boldsymbol{y}}_{\boldsymbol{x}} - O(\frac{l+r}{2^{l-k}})\,, \tag{2.130}$$

where it is assumed that $r > k$.

In contrast to the previous cases we see that here there is no single dominant history. Instead, after the k-th step we have 2^k different histories each having asymptotically the same probability: $2^{-k} + O(\frac{l+r}{2^{l-k}})$. These histories are defined by the condition $\boldsymbol{x} = \boldsymbol{y}$ and correspond to 2^k different values of $\boldsymbol{\beta}$ in (2.130). Because $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = k$, the condition $\boldsymbol{x} = \boldsymbol{y}$ requires that a shift k binary positions to the left maps the tail of the fixed string $\boldsymbol{\alpha x}$ (which specifies the initial condition) onto the head of the string $\boldsymbol{y\beta}$ (which specifies a history):

$$\begin{array}{c} \underbrace{\square\square\ldots\square}_{l}\ \ \alpha_1 \ldots \alpha_k\ \underline{\boldsymbol{x}_1 \ldots \boldsymbol{x}_c}\ \underbrace{\square\square\ldots\square}_{r}\,, \\ \swarrow \\ \underbrace{\square\square\ldots\square}_{l}\ \ \overline{\boldsymbol{y}_1 \ldots \boldsymbol{y}_c}\ \beta_1 \ldots \beta_k\ \underbrace{\square\square\ldots\square}_{r}\,. \end{array} \tag{2.131}$$

During this transformation all k bits of $\boldsymbol{\alpha}$ are lost as they reach the scale at which the momentum becomes coarse-grained. At the same time k unspecified position bits $\beta_1 \ldots \beta_k$ come out: these unspecified bits are random as they come out at the scale at which position is coarse-grained. Different strings of k bits $\beta_1 \ldots \beta_k$ correspond to 2^k different histories. At each step the number of histories with significant probability doubles, as each history branches with probability $1/2$ into two different histories. Because any other histories (for which $\boldsymbol{x} \neq \boldsymbol{y}$) have

negligible probabilities we see the loss of one bit of information per time step. We shall return to this point in section 2.6.4 where we consider the entropy of the set of decoherent histories.

An interesting technical question is whether we can relax the requirement $r > k$ which is assumed in this theorem. This would imply (like in all the previous theorems) that we could use $\rho^{(l,0)}_{\boldsymbol{\alpha x}}$ as a particular case of $\rho^{(l,r)}_{\boldsymbol{\alpha x}}$. This is interesting because it is equivalent to an intuitive expectation that the coarse-graining over the least significant position bits may be optional. For the cases of the first and the second iteration we find a positive answer to this question (see Appendix I for details).

Theorem 5

Fix any integer $\gamma \geq 1$, any string $\boldsymbol{x}$ of length $|\boldsymbol{x}| = \gamma$, and any two ordered sequences of strings $\vec{\boldsymbol{y}}(k) = (\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots, \boldsymbol{y}^k)$ and $\vec{\boldsymbol{z}}(k) = (\boldsymbol{z}^1, \boldsymbol{z}^2, \ldots, \boldsymbol{z}^k)$ such that $|\boldsymbol{y}^j| = |\boldsymbol{z}^j| = \gamma$, $j = 1, \ldots, k$, where k is the number of iterations. For sufficiently large l we have then:

$$\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}},\, h^3_{\vec{\boldsymbol{y}}(k)}, h^3_{\vec{\boldsymbol{z}}(k)}] \equiv \mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} P^{(l,r)}_{\boldsymbol{y}^{k-1}} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r)}_{\boldsymbol{x}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^{k-1}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^k}] \tag{2.132}$$

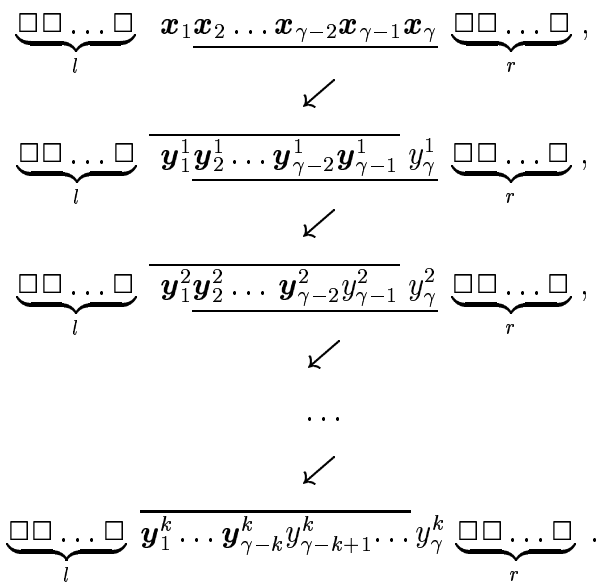
$$= 2^{-k} \left(\prod_{j=1}^{k} \delta^{\boldsymbol{z}^j}_{\boldsymbol{y}^j} \right) \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}_{j+1}}_{\boldsymbol{y}^j_1} \right) \delta^{\boldsymbol{x}_{k+1:\gamma}}_{\boldsymbol{y}^k_{1:\gamma-k}} + O(\frac{l+r-k}{2^{l-2(k^2+k)}}) \tag{2.133}$$

$$= 2^{-k} \underbrace{\left(\prod_{j=1}^{k} \delta^{\boldsymbol{z}^j}_{\boldsymbol{y}^j} \right)}_{\text{diagonal}} \cdot \underbrace{\left(\delta^{\boldsymbol{x}_{2:\gamma}}_{\boldsymbol{y}^1_{1:\gamma-1}} \prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \right)}_{\text{step-by-step shift}} \cdot \underbrace{\left(\delta^{\boldsymbol{x}_{k+1:\gamma}}_{\boldsymbol{y}^k_{1:\gamma-k}} \right)}_{k\text{th shift}}$$
$$+ O(\frac{l+r-k}{2^{l-2(k^2+k)}})\,, \tag{2.134}$$

where it is assumed that $r > k$; and the last equality provides a somewhat redundant but more transparent formulation of the theorem.

We see that the expression in the first parentheses is zero for all off-diagonal elements of the decoherence functional. This implies that in the limit of big l all

off-diagonal elements of the decoherence functional vanish satisfying the medium decoherence condition. The diagonal elements of the decoherence functional can therefore be interpreted as probabilities of the corresponding type-3 histories. Only 2^k diagonal elements (where k is the number of iterations) are nonzero (up to the exponentially small error term). This is similar to the case of type-4 histories considered above: once again, we have only 2^k histories with asymptotically equal probabilities. And again, we see that the number of such histories doubles after each step resulting in a loss of information at the rate of 1 bit per step. The conditions that the histories with nonzero probabilities should satisfy are also very similar to the case of type-4 histories: the only difference is that now each of these histories is a sequence of k projectors and each of these projectors is related to the initial state via a shift according to the position of the projector in the history:



On this diagram the first line represents the initial condition $\rho_{\boldsymbol{x}}^{(l,r)}$; the subsequent lines correspond to the projectors $P_{\boldsymbol{y}^1}^{(l,r)} \dots P_{\boldsymbol{y}^k}^{(l,r)}$ in the history. The bold face is used to indicate the bits which are completely determined by the initial condition for any history with essentially nonzero probability. Such histories should satisfy the step-by-step shift condition denoted on the diagram by the arrow and lines: $\boldsymbol{x}_{2:\gamma}$, for example, is shifted onto $\boldsymbol{y}^1_{1:\gamma-1}$. For the entire history, therefore, there are only k independent bits which can be chosen arbitrarily and still satisfy the step-by-step shift condition. If we choose $y^k_{\gamma-k+1:\gamma}$ as independent and record only the very last projector, ignoring the rest of the trajectory, we get the case of type-4 history considered above.

2.6.4 Entropy of decoherent histories

Once we have established the decoherence condition and calculated the probabilities associated to each of the decoherent histories we can define the entropy of the set of decoherent histories $\{h_{\vec{\boldsymbol{y}}(k)}\}$ as [13, 30, 23]:

$$H(\{h_{\vec{\boldsymbol{y}}(k)}\}) \equiv -\sum_{\boldsymbol{y}} p[h_{\vec{\boldsymbol{y}}(k)}] \log_2 p[h_{\vec{\boldsymbol{y}}(k)}] \,, \tag{2.136}$$

where $p[h_{\vec{\boldsymbol{y}}(k)}]$ is determined by Eq. (2.94). The value of $H(\{h_{\vec{\boldsymbol{y}}(k)}\})$ determines the amount of information which is lost on average if we do not follow a particular history and know only the set of possible alternative histories $\{h_{\vec{\boldsymbol{y}}(k)}\}$. Using the results summarized in the previous chapter, it is now easy to compute the entropy for each of the considered types of histories:

$$H(\{h^1_{\vec{\boldsymbol{y}}(k)}\}) = O(\frac{(l+r)\log_2(l+r)}{2^{l-2k}})\,, \tag{2.137}$$

$$H(\{h^2_{\boldsymbol{y}}\}) = O(\frac{(l+r)\log_2(l+r)}{2^{l-k}})\,, \tag{2.138}$$

$$H(\{h^3_{\vec{\boldsymbol{y}}(k)}\}) = k + O(\frac{(l+r-k)\log_2(l+r-k)}{2^{l-2(k^2+k)}})\,, \tag{2.139}$$

$$H(\{h^4_{\boldsymbol{y}}\}) = k + O(\frac{(l+r)\log_2(l+r)}{2^{l-k}})\,. \tag{2.140}$$

We see that in the realistic cases of type-3 and type-4 histories, for which the level of coarse-graining is fixed at all times, there is a linear growth of entropy at the rate of one bit per step. This is exactly what one would expect from the purely classical consideration of the baker's map. In the classical case the rate of entropy increase is quantified by the Kolmogorov-Sinai entropy which we reviewed in section 2.2.3. The KS-entropy is an established tool for the study of classical chaos and is widely used as a measure of chaoticity of classical chaotic systems. The results presented in this chapter provide a conceptual link between the KS-entropy and the rate of system entropy increase as defined in the decoherent histories approach for quantum systems. Providing a more formal connection is a project for future research.

2.7 Proofs of the theorems

Theorem 1

For a fixed number of iterations k, in the limit of large l any two type-2 histories $h^2_{\boldsymbol{y}}$ and $h^2_{\boldsymbol{z}}$ satisfy the asymptotic relation

$$\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}},\, h^2_{\boldsymbol{y}}, h^2_{\boldsymbol{z}}\,] = \mathrm{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}} \hat{B}^k \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^k P^{(l-k,r+k)}_{\boldsymbol{z}}] = \delta^{\boldsymbol{z}}_{\boldsymbol{y}}\,[\,\delta^{\boldsymbol{y}}_{\boldsymbol{x}} - O(\frac{l+r}{2^{l-k}})]\,. \tag{2.141}$$

Proof

This equation can be considered as a generalization of Eq. (2.117). The proof closely follows the arguments in [65] (section IV.B.). We are happy if we prove that

$$\text{Tr } [P_{\boldsymbol{y}}^{(l-k,r+k)} \hat{B}^k \rho_{\boldsymbol{y}}^{(l,r)} (\hat{B}^\dagger)^k] = 1 - O(\frac{l+r}{2^{l-k}}) \,, \tag{2.142}$$

the rest of the theorem follows from mutual orthogonality of the projectors $\{P_{\boldsymbol{y}}^{(l-k,r+k)}\}$ and from normalization. Eq. (2.68) becomes in the notation of this section

$$\text{Tr } [P_{\boldsymbol{y}}^{(\lambda-\kappa,\kappa)} \hat{B} \rho_{\boldsymbol{y}}^{(\lambda-\kappa+1,\kappa-1)} \hat{B}^\dagger] = 1 - O(\frac{\lambda}{2^{\lambda-\kappa}}) \,. \tag{2.143}$$

We perform the change the variables

$$k + r = \kappa \,, \quad l - k = \lambda - \kappa \,, \tag{2.144}$$

and obtain

$$\text{Tr } [P_{\boldsymbol{y}}^{(l-k,k+r)} \hat{B} \rho_{\boldsymbol{y}}^{(l-k+1,r+k-1)} \hat{B}^\dagger] = 1 - O(\frac{l+r}{2^{l-k}}) \,, \tag{2.145}$$

which is equivalent to

$$\text{Tr } [\varrho_k \hat{B} \varrho_{k-1} \hat{B}^\dagger] = 2^{-(l+r)} [1 - O(\frac{l+r}{2^{l-k}})] \,, \tag{2.146}$$

where we introduced auxiliary matrices $\varrho_k \equiv 2^{-(l+r)} P_{\boldsymbol{y}}^{(l-k,r+k)}$. The above equation can be rewritten in terms of the distance measure given by Eq. (2.75). We have

$$\begin{aligned} d(\varrho_{k+1}, \hat{B} \varrho_k \hat{B}^\dagger) &= \sqrt{\text{Tr } [\varrho_{k+1} - (B \varrho_k B^\dagger)]^2} \\ &= \sqrt{\text{Tr } \varrho_{k+1}^2 + \text{Tr } \varrho_k^2 - 2\text{Tr } (\varrho_{k+1} \hat{B} \varrho_k \hat{B}^\dagger)} \\ &= O(2^{-l-(r-k)/2} \sqrt{l+r}) \,, \end{aligned} \tag{2.147}$$

where we used the equality $\text{Tr } \varrho_k^2 = 2^{l+r}/2^{2(l+r)} = 2^{l+r}$ for any k.

We shall prove (2.142) by induction. The case $k = 1$ of (2.142) follows directly from (2.145). Assuming that (2.142) is true for some value of k we have, as in the previous equation,

$$\begin{aligned} d(\varrho_k, \hat{B}^k \varrho_0 [\hat{B}^\dagger]^k) &= \sqrt{\text{Tr } \varrho_k^2 + \text{Tr } \varrho_0^2 - 2\text{Tr } (\varrho_k \hat{B}^k \varrho_0 [\hat{B}^\dagger]^k)} \\ &= O(2^{-l-(r-k)/2} \sqrt{l+r}) \,. \end{aligned} \tag{2.148}$$

We now use the unitary invariance of the distance measure (2.75) to get

$$d(\hat{B}\varrho_k\hat{B}^\dagger, \hat{B}^{k+1}\varrho_0[\hat{B}^\dagger]^{k+1}) = O(2^{-l-(r-k)/2}\sqrt{l+r}) \ . \tag{2.149}$$

Using the triangle inequality for the distance measure (2.75) we have from (2.147) and (2.149)

$$d(\varrho_{k+1}, \hat{B}^{k+1}\varrho_0[\hat{B}^\dagger]^{k+1}) = O(2^{-l-(r-k)/2}\sqrt{l+r}) \ , \tag{2.150}$$

which implies

$$\text{Tr } [P_{\boldsymbol{y}}^{(l-k-1,k+1+r)}\hat{B}^{k+1}\rho_{\boldsymbol{y}}^{(l,r)}(\hat{B}^\dagger)^{k+1}] = 1 - O(\frac{l+r}{2^{l-k}}) \ . \tag{2.151}$$

By induction this completes the prove of Eq. (2.142) for any $k \leq k_{\max}$. □

Theorem 2

For a fixed number of iterations k the off-diagonal elements of the decoherence functional for type-1 histories can be made arbitrarily small by choosing sufficiently large l. More precisely

$$\left|\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{z}}(k)}]\right| = O(\frac{l+r}{2^{l-2k}}) \qquad h^1_{\vec{\boldsymbol{y}}} \neq h^1_{\vec{\boldsymbol{z}}} \ . \tag{2.152}$$

Proof

Let us consider a sequence of histories $h^1_{\vec{\boldsymbol{y}}(1)}, \ldots, h^1_{\vec{\boldsymbol{y}}(k)}$, such that the strings $\vec{\boldsymbol{y}}(k-1) = \boldsymbol{y}^1, \ldots, \boldsymbol{y}^{k-1}$ coincide with the first $k-1$ strings from $\vec{\boldsymbol{y}}(k) = \boldsymbol{y}^1, \ldots, \boldsymbol{y}^k$. We estimate the difference

$$\begin{aligned}
\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{y}}(k)}] - D^{(l-k,r+k)}_{\boldsymbol{y}\,;\,\rho_{\boldsymbol{x}}^{(l,r)}} &\leq \mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k-1)}, h^1_{\vec{\boldsymbol{y}}(k-1)}] - D^{(l-k,r+k)}_{\boldsymbol{y}\,;\,\rho_{\boldsymbol{x}}^{(l,r)}} \\
&= \mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k-1)}, h^1_{\vec{\boldsymbol{y}}(k-1)}] - D^{(l-k+1,r+k-1)}_{\boldsymbol{y}\,;\,\rho_{\boldsymbol{x}}^{(l,r)}} \\
&\quad + O(\frac{l+r}{2^{l-k+1}}) \ ,
\end{aligned} \tag{2.153}$$

where we first used inequality (2.101) and then Eq. (2.117). For the case $k = 1$ we have $\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(1)}, h^1_{\vec{\boldsymbol{y}}(1)}] = D^{(l-1,1)}_{\boldsymbol{y}\,;\,\rho_{\boldsymbol{x}}^{(l,r)}}$ and therefore by induction we have

$$\mathcal{D}[\rho_{\boldsymbol{x}}^{(l,r)}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{y}}(k)}] \leq D^{(l-k,r+k)}_{\boldsymbol{y}\,;\,\rho_{\boldsymbol{x}}^{(l,r)}} + O(\frac{l+r}{2^{l-2k}}) \ . \tag{2.154}$$

Off-diagonal elements can be estimated using Eq. (2.100)

$$\left|\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{z}}(k)}]\right|^2 \leq \mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}}, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{y}}(k)}]\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}}, h^1_{\vec{\boldsymbol{z}}(k)}, h^1_{\vec{\boldsymbol{z}}(k)}] \;, \tag{2.155}$$

and therefore using (2.154) and (2.117) we have (2.152) as required. □

Theorem 3

For a fixed number of iterations k and for sufficiently large l the diagonal elements of the decoherence functional for type-1 histories approach either close to one or to zero. More precisely

$$\begin{aligned}
&\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}},\, h^1_{\vec{\boldsymbol{y}}(k)}, h^1_{\vec{\boldsymbol{y}}(k)}\,] \\
&\qquad \equiv \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k}\hat{B}\cdots P^{(l-1,r+1)}_{\boldsymbol{y}^1}\hat{B}\rho^{(l,r)}_{\boldsymbol{x}}\hat{B}^\dagger P^{(l-1,r+1)}_{\boldsymbol{y}^1}\cdots\hat{B}^\dagger P^{(l-k,r+k)}_{\boldsymbol{y}^k}] \\
&\qquad = \delta^{\boldsymbol{x}}_{\boldsymbol{y}^k}\delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}}\cdots\delta^{\boldsymbol{x}}_{\boldsymbol{y}^1} + O(\frac{l+r}{2^{l-2k}}) \;.
\end{aligned} \tag{2.156}$$

Proof

For any κ we can write a decomposition of unity $\mathbb{1} = \sum_{\boldsymbol{y}^\kappa} P^{(l-\kappa,r+\kappa)}_{\boldsymbol{y}^\kappa}$ and therefore directly by definition (2.117) we have that

$$\begin{aligned}
D^{(l-k,r+k)}_{\boldsymbol{y}^k\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} &= \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k}\hat{B}^k\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^k] \\
&= \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k}\hat{B}\mathbb{1}\hat{B}^{k-1}\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^{k-1}\mathbb{1}\hat{B}^\dagger] \\
&= \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k}\hat{B}P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}}\hat{B}^{k-1}\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^{k-1}P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}}\hat{B}^\dagger] \\
&\quad + F(\boldsymbol{y}^k, \boldsymbol{y}^{k-1}) \;,
\end{aligned} \tag{2.157}$$

where

$$\begin{aligned}
&F(\boldsymbol{y}^k, \boldsymbol{y}^{k-1}) \\
&\equiv \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k}(\overbrace{\sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}}\sum_{\boldsymbol{\beta}\neq\boldsymbol{y}^{k-1}}\hat{B}P^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}}\hat{B}^{k-1}\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^{k-1}P^{(l-k+1,r+k-1)}_{\boldsymbol{\beta}}\hat{B}^\dagger}^{\text{hermitian}})] \\
&\qquad \leq \sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}}\sum_{\boldsymbol{\beta}\neq\boldsymbol{y}^{k-1}}\text{Tr}\,[P^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}}\hat{B}^{k-1}\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^{k-1}P^{(l-k+1,r+k-1)}_{\boldsymbol{\beta}}] \\
&\qquad\qquad = \sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}}\text{Tr}\,[P^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}}\hat{B}^{k-1}\rho^{(l,r)}_{\boldsymbol{x}}(\hat{B}^\dagger)^{k-1}P^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}}] \\
&\qquad\qquad\qquad = \sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}} D^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} \;.
\end{aligned} \tag{2.158}$$

Because

$$\sum_{\boldsymbol{\alpha}} D^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} = 1 \tag{2.159}$$

we have

$$D^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} + \sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}} D^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} = 1 \tag{2.160}$$

which together with equation (2.141) gives

$$\sum_{\boldsymbol{\alpha}\neq\boldsymbol{y}^{k-1}} D^{(l-k+1,r+k-1)}_{\boldsymbol{\alpha}\,;\,\rho^{(l,r)}_{\boldsymbol{x}}} = 1 - \delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}} + O(\frac{l+r}{2^{l-k}}) \; . \tag{2.161}$$

The function F defined in (2.158) is nonnegative. We therefore have

$$0 \leq F(\boldsymbol{y}^k, \boldsymbol{y}^{k-1}) \leq 1 - \delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}} + O(\frac{l+r}{2^{l-k}}) \; . \tag{2.162}$$

This means that

$$F(\boldsymbol{y}^k, \boldsymbol{y}^{k-1}) = O(\frac{l+r}{2^{l-k}}) \; , \quad \text{for} \;\; \boldsymbol{x} = \boldsymbol{y}^{k-1} \; . \tag{2.163}$$

Applying this knowledge to (2.157) and using (2.141) we have

$$\begin{aligned} \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k} \hat{B} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^{k-1} \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^{k-1} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^\dagger] \\ = \delta^{\boldsymbol{x}}_{\boldsymbol{y}^k} + O(\frac{l+r}{2^{l-k-1}}) , \quad \text{for} \;\; \boldsymbol{x} = \boldsymbol{y}^{k-1} \; . \end{aligned} \tag{2.164}$$

On the other hand

$$\begin{aligned} &\text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k} (\overbrace{\hat{B} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^{k-1} \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^{k-1} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^\dagger}^{\text{hermitian}})] \\ &\leq \text{Tr}\,[\hat{B} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^{k-1} \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^{k-1} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^\dagger] \\ &= \text{Tr}\,[P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^{k-1} \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^{k-1}] \\ &= \delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}} + O(\frac{l+r}{2^{l-k+1}}) \; , \end{aligned} \tag{2.165}$$

which together with (2.164) gives

$$\begin{aligned} \text{Tr}\,[P^{(l-k,r+k)}_{\boldsymbol{y}^k} \hat{B} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^{k-1} \rho^{(l,r)}_{\boldsymbol{x}} (\hat{B}^\dagger)^{k-1} P^{(l-k+1,r+k-1)}_{\boldsymbol{y}^{k-1}} \hat{B}^\dagger] \\ = \delta^{\boldsymbol{x}}_{\boldsymbol{y}^k} \delta^{\boldsymbol{x}}_{\boldsymbol{y}^{k-1}} + O(\frac{l+r}{2^{l-k-1}}) \; . \end{aligned} \tag{2.166}$$

Repeating the same arguments by induction we arrive at (2.156). □

Theorem 4

Fix two strings $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *of the same length:* $|\boldsymbol{x}| = |\boldsymbol{y}|$. *For any two strings* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\beta}$ *such that* $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = k$, *where* k *is a fixed number of iterations, we have in the limit of large* l:

$$\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{\alpha x}}, h^4_{\boldsymbol{y\beta}}, h^4_{\boldsymbol{y\beta}}] = \mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y\beta}} \hat{B}^k \rho^{(l,r)}_{\boldsymbol{\alpha x}} (\hat{B}^\dagger)^k] = 2^{-k} \delta^{\,\boldsymbol{y}}_{\boldsymbol{x}} - O(\frac{l+r}{2^{l-k}})\,, \tag{2.167}$$

where it is assumed that $r > k$.

Proof

Introducing a pair of auxiliary strings $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ such that $|\bar{\boldsymbol{x}}| = |\bar{\boldsymbol{y}}| = k$ we have from (2.141)

$$\mathrm{Tr}\,[P^{(l-k,r+k)}_{\bar{\boldsymbol{y}}\boldsymbol{y}\boldsymbol{\beta}} \hat{B}^k \rho^{(l,r)}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} (\hat{B}^\dagger)^k] = \delta^{\bar{\boldsymbol{y}}\boldsymbol{y}\boldsymbol{\beta}}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} + O(\frac{l+r}{2^{l-k}})\,. \tag{2.168}$$

We redefine the variables by substituting r for $r+k$

$$\mathrm{Tr}\,[P^{(l-k,r)}_{\bar{\boldsymbol{y}}\boldsymbol{y}\boldsymbol{\beta}} \hat{B}^k \rho^{(l,r-k)}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} (\hat{B}^\dagger)^k] = \delta^{\bar{\boldsymbol{y}}\boldsymbol{y}\boldsymbol{\beta}}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} + O(\frac{l+r}{2^{l-k}})\,. \tag{2.169}$$

We now perform the summation over $\bar{y}$ as explained in Appendix II to get

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y\beta}} \hat{B}^k \rho^{(l,r-k)}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} (\hat{B}^\dagger)^k] = \delta^{\boldsymbol{y\beta}}_{\boldsymbol{x}\bar{\boldsymbol{x}}} + O(\frac{l+r}{2^{l-k}})\,. \tag{2.170}$$

Performing further summation over $\bar{\boldsymbol{x}}$ and using the equality

$$\rho^{(l,r)}_{\boldsymbol{\alpha x}} = 2^{-k} \sum_{\bar{\boldsymbol{x}}} \rho^{(l,r-k)}_{\boldsymbol{\alpha x}\bar{\boldsymbol{x}}} \tag{2.171}$$

we derive (2.167) as required. □

Theorem 5

Fix any integer $\gamma \geq 1$, *any string* $\boldsymbol{x}$ *of length* $|\boldsymbol{x}| = \gamma$, *and any two ordered sequences of strings* $\vec{\boldsymbol{y}}(k) = (\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots, \boldsymbol{y}^k)$ *and* $\vec{\boldsymbol{z}}(k) = (\boldsymbol{z}^1, \boldsymbol{z}^2, \ldots, \boldsymbol{z}^k)$ *such that* $|\boldsymbol{y}^j| = |\boldsymbol{z}^j| = \gamma$, $j = 1, \ldots, k$, *where* k *is the number of iterations. For sufficiently large* l *we have then:*

$$\begin{aligned}
\mathcal{D}[\rho^{(l,r)}_{\boldsymbol{x}},\, h^3_{\vec{\boldsymbol{y}}(k)}, h^3_{\vec{\boldsymbol{z}}(k)}] &\equiv \mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} P^{(l,r)}_{\boldsymbol{y}^{k-1}} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r)}_{\boldsymbol{x}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^{k-1}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^k}] \\
&= 2^{-k} \left(\prod_{j=1}^{k} \delta^{\boldsymbol{z}^j}_{\boldsymbol{y}^j} \right) \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}_{j+1}}_{\boldsymbol{y}^j_1} \right) \delta^{\boldsymbol{x}_{k+1:\gamma}}_{\boldsymbol{y}^k_{1:\gamma-k}}
\end{aligned}$$

$$+O(\frac{l+r-k}{2^{l-2(k^2+k)}})\,, \tag{2.172}$$

where it is assumed that $r > k$.

Proof

Before we present a formal proof of the theorem it is helpful to illustrate the idea behind the proof using the diagrams defined in (2.114). Introduce an auxiliary variable

$$r' \equiv r - k\,. \tag{2.173}$$

The idea is to represent every projector $P^{(l,r)}_{\boldsymbol{y}^j}$ in (2.172) as a sum of shifting projectors $P^{(l-j,r'+j)}_{\tilde{\boldsymbol{y}}_{1:j}\boldsymbol{y}^j\tilde{\boldsymbol{y}}_{k-j}}$. Any history $h^3_{\vec{\boldsymbol{y}}(k)}$ therefore becomes

$$\begin{aligned}
h^3_{\vec{\boldsymbol{y}}(k)} \;=\; & \Big(\underbrace{\square\ldots\square\square\square\square}_{l}\,\boldsymbol{y}_1\,\underbrace{\square\square\square\square\square\ldots\square}_{r}\,, \\
& \underbrace{\square\ldots\square\square\square\square}_{l}\,\boldsymbol{y}_2\,\underbrace{\square\square\square\square\square\ldots\square}_{r}\,,\ \ldots\,, \\
& \underbrace{\square\ldots\square\square\square\square}_{l}\,\boldsymbol{y}_k\,\underbrace{\square\square\square\square\square\ldots\square}_{r}\,\Big) \\
= \; & \Big(\underbrace{\square\ldots\square\square\square\square}_{l}\,\boldsymbol{y}_1\,\blacksquare\blacksquare\blacksquare..\blacksquare\,\underbrace{\square\ldots\square}_{r'}\,, \\
& \underbrace{\square\ldots\square\square\square\square}_{l-1}\blacksquare\,\boldsymbol{y}_2\,\blacksquare\blacksquare..\blacksquare\,\underbrace{\square\square\ldots\square}_{r'+1}\,, \\
& \underbrace{\square\ldots\square\square\square}_{l-2}\blacksquare\blacksquare\,\boldsymbol{y}_3\,\blacksquare..\blacksquare\,\underbrace{\square\square\square\ldots\square}_{r'+2}\,,\ \ldots\,, \\
& \underbrace{\square\ldots\square}_{l-k}\blacksquare\blacksquare..\blacksquare\,\boldsymbol{y}_k\,\underbrace{\square\square\square\square\square\ldots\square}_{r'+k=r}\,\Big)\,,
\end{aligned} \tag{2.174}$$

where we used black boxes to indicate the bits which were summed over to make the fixed projectors out of the sliding ones. Then we can use the results on sliding histories given by equations (2.156) and (2.152).

Now we proceed with a formal proof of the theorem. Let $\{\boldsymbol{u}^j\}_{j=1}^k$ be a set of k strings such that for any $j = 1,\ldots,k$ the length $|\boldsymbol{u}^j| = \gamma + k$. We have from Eqs. (2.156) and (2.152)

$$\mathrm{Tr}\,[P^{(l-k,\,r'+k)}_{\boldsymbol{u}^k}\hat{B}\cdots P^{(l-1,\,r'+1)}_{\boldsymbol{u}^1}\hat{B}\rho^{(l,r)}_{\boldsymbol{v}}\hat{B}^\dagger P^{(l-1,\,r'+1)}_{\boldsymbol{w}^1}\cdots\hat{B}^\dagger P^{(l-k,\,r'+k)}_{\boldsymbol{w}^k}] = (\delta^{\boldsymbol{v}}_{\boldsymbol{u}^k}\delta^{\boldsymbol{v}}_{\boldsymbol{u}^{k-1}}\cdots\delta^{\boldsymbol{v}}_{\boldsymbol{u}^1})\,(\delta^{\boldsymbol{v}}_{\boldsymbol{w}^k}\delta^{\boldsymbol{v}}_{\boldsymbol{w}^{k-1}} \tag{2.175}$$

where $\boldsymbol{v}$ is a string of length $|\boldsymbol{v}| = \gamma + k$. We write each $\boldsymbol{u}^j$ and each $\boldsymbol{w}^j$ as a concatenation of three strings:

$$\boldsymbol{u}^j = \bar{\boldsymbol{y}}^j \boldsymbol{y}^j \tilde{\boldsymbol{y}}^j \,, \quad \boldsymbol{w}^j = \bar{\boldsymbol{z}}^j \boldsymbol{z}^j \tilde{\boldsymbol{z}}^j \,, \tag{2.176}$$

where the lengths $|\bar{\boldsymbol{y}}^j| \equiv |\bar{\boldsymbol{z}}^j| \equiv j$ and $|\tilde{\boldsymbol{y}}^j| \equiv |\tilde{\boldsymbol{z}}^j| \equiv k - j$ for $j = 1, \ldots, k$, so that $|\boldsymbol{y}^j| = |\boldsymbol{z}^j| = \gamma$. We also define k different representations of $\boldsymbol{v}$

$$\boldsymbol{v} = \bar{\boldsymbol{x}}^j \boldsymbol{x}^j \tilde{\boldsymbol{x}} \tag{2.177}$$

where $|\tilde{\boldsymbol{x}}| \equiv k$, and $|\bar{\boldsymbol{x}}^j| \equiv j$ for $j = 1, \ldots, k$. Summation over $\{\bar{\boldsymbol{y}}^j\}_{j=1}^k$ and over $\{\bar{\boldsymbol{z}}^j\}_{j=1}^k$ contains 2^{k^2+k} terms and therefore we have

$$\mathrm{Tr}\,[P^{(l,\,r'+k)}_{\boldsymbol{y}^k\tilde{\boldsymbol{y}}^k} \hat{B} \cdots P^{(l,\,r'+1)}_{\boldsymbol{y}^1\tilde{\boldsymbol{y}}^1} \hat{B} \rho^{(l,\,r')}_{\boldsymbol{v}} \hat{B}^\dagger P^{(l,\,r'+1)}_{\boldsymbol{z}^1\tilde{\boldsymbol{z}}^1} \cdots \hat{B}^\dagger P^{(l,\,r'+k)}_{\boldsymbol{z}^k\tilde{\boldsymbol{z}}^k}] = \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^k\tilde{\boldsymbol{y}}^k} \delta^{\boldsymbol{x}^{k-1}\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^{k-1}\tilde{\boldsymbol{y}}^{k-1}} \cdots \delta^{\boldsymbol{x}^1\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^1\tilde{\boldsymbol{y}}^1}\right) \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{z}^k\tilde{\boldsymbol{z}}^k} \delta^{\boldsymbol{x}^k}_{\boldsymbol{z}^k} \tag{2.178}$$

Changing the variables to $r \equiv r' + k$ we have

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k\tilde{\boldsymbol{y}}^k} \hat{B} \cdots P^{(l,r-k+1)}_{\boldsymbol{y}^1\tilde{\boldsymbol{y}}^1} \hat{B} \rho^{(l,r-k)}_{\boldsymbol{v}} \hat{B}^\dagger P^{(l,r-k+1)}_{\boldsymbol{y}^1\tilde{\boldsymbol{y}}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{y}^k\tilde{\boldsymbol{y}}^k}] = \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^k\tilde{\boldsymbol{y}}^k} \delta^{\boldsymbol{x}^{k-1}\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^{k-1}\tilde{\boldsymbol{y}}^{k-1}} \cdots \delta^{\boldsymbol{x}^1\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^1\tilde{\boldsymbol{y}}^1}\right) \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{z}^k\tilde{\boldsymbol{z}}^k} \delta^{\boldsymbol{x}^{k-} \tag{2.179}$$

We now perform a summation over $\{\tilde{\boldsymbol{y}}^j\}_{j=1}^k$ and $\{\tilde{\boldsymbol{z}}^j\}_{j=1}^k$ (the total of 2^{k^2-k} terms) to get

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r-k)}_{\boldsymbol{v}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^k}] = \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}_{1:k}}_{\boldsymbol{y}^k} \delta^{\boldsymbol{x}^{k-1}\tilde{\boldsymbol{x}}_{1:k-1}}_{\boldsymbol{y}^{k-1}} \cdots \delta^{\boldsymbol{x}^1\tilde{\boldsymbol{x}}_1}_{\boldsymbol{y}^1}\right) \left(\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}_{1:k}}_{\boldsymbol{z}^k} \delta^{\boldsymbol{x}^{k-1}\tilde{\boldsymbol{x}}_{1:k-1}}_{\boldsymbol{z}^{k-1}} \tag{2.180}$$

By construction for any $j = 2, \ldots, k$, strings $(\boldsymbol{x}^j\tilde{\boldsymbol{x}}_{1:j})$ and $(\boldsymbol{x}^{j-1}\tilde{\boldsymbol{x}}_{1:j-1})$ have the same length γ, and the first $\gamma - 1$ bits of the string $(\boldsymbol{x}^j\tilde{\boldsymbol{x}}_{1:j})$ coincide with the last $\gamma - 1$ bits of the string $(\boldsymbol{x}^{j-1}\tilde{\boldsymbol{x}}_{1:j-1})$. Formally we write

$$(\boldsymbol{x}^j\tilde{\boldsymbol{x}}_{1:j})_{1:\gamma-1} = (\boldsymbol{x}^{j-1}\tilde{\boldsymbol{x}}_{1:j-1})_{2:\gamma} \,. \tag{2.181}$$

Using this fact and noticing that $|\boldsymbol{y}^j| = |\boldsymbol{u}^j| - |\bar{\boldsymbol{y}}^j| - |\tilde{\boldsymbol{y}}^j| = \gamma$, we have

$$\delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}_{1:k}}_{\boldsymbol{y}^k} \delta^{\boldsymbol{x}^{k-1}\tilde{\boldsymbol{x}}_{1:k-1}}_{\boldsymbol{y}^{k-1}} \cdots \delta^{\boldsymbol{x}^1\tilde{\boldsymbol{x}}_1}_{\boldsymbol{y}^1} = \delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^k} \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}^j_1}_{\boldsymbol{y}^j_1}\right) . \tag{2.182}$$

Equation (2.180) therefore becomes

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r-k)}_{\boldsymbol{v}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{y}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{y}^k}] = \delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{y}^k} \delta^{\boldsymbol{x}^k\tilde{\boldsymbol{x}}}_{\boldsymbol{z}^k} \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}^j_1}_{\boldsymbol{y}^j_1}\right) \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{z}^j_{2:\gamma}}_{\boldsymbol{z}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}^j_1}_{\boldsymbol{z}^j_1}\right) + C \tag{2.183}$$

We see that the product of delta-functions in the rhs of this equation is nonzero only if $\boldsymbol{y}^j = \boldsymbol{z}^j$, $j = 1, \ldots, k$. Using this fact and the identity $\delta^{\boldsymbol{x}^k \tilde{\boldsymbol{x}}}_{\boldsymbol{y}^k} = \delta^{\boldsymbol{x}^k}_{\boldsymbol{y}^k_{1:\gamma-k}} \delta^{\tilde{\boldsymbol{x}}_{1:k}}_{\boldsymbol{y}^k_{\gamma-k+1:\gamma}}$ we have

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r-k)}_{\boldsymbol{v}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{y}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{y}^k}] = \delta^{\boldsymbol{x}^k}_{\boldsymbol{y}^k_{1:\gamma-k}} \delta^{\tilde{\boldsymbol{x}}_{1:k}}_{\boldsymbol{y}^k_{\gamma-k+1:\gamma}} \left(\prod_{j=1}^{k} \delta^{\boldsymbol{z}^j}_{\boldsymbol{y}^j} \right) \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}^j_1}_{\boldsymbol{y}^j_1} \right) + O \tag{2.184}$$

After summing over $\tilde{\boldsymbol{x}}_{1:k}$ and noticing that $\sum_{\tilde{\boldsymbol{x}}} \rho^{(l,r-k)}_{\boldsymbol{v}} = 2^k \rho^{(l,r)}_{\boldsymbol{x}}$, where $\boldsymbol{x} \equiv \bar{\boldsymbol{x}}^j \boldsymbol{x}^j$, we finally obtain

$$\mathrm{Tr}\,[P^{(l,r)}_{\boldsymbol{y}^k} \hat{B} \cdots P^{(l,r)}_{\boldsymbol{y}^1} \hat{B} \rho^{(l,r)}_{\boldsymbol{x}} \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^1} \cdots \hat{B}^\dagger P^{(l,r)}_{\boldsymbol{z}^k}] = 2^{-k} \left(\prod_{j=1}^{k} \delta^{\boldsymbol{z}^j}_{\boldsymbol{y}^j} \right) \left(\prod_{j=1}^{k-1} \delta^{\boldsymbol{y}^j_{2:\gamma}}_{\boldsymbol{y}^{j+1}_{1:\gamma-1}} \delta^{\boldsymbol{x}^j_1}_{\boldsymbol{y}^j_1} \right) \delta^{\boldsymbol{x}^k}_{\boldsymbol{y}^k_{1:\gamma-k}} + O(\frac{l+r}{2^{l-2(k}} \tag{2.185}$$

which is equivalent to (2.172). □

2.8 Appendix 2.A

Here we prove that for the first and the second iterations the coarse-graining over the least significant position bits in Eq.(2.167) is optional. Because

$$\sum_{\boldsymbol{g}} \mathrm{Tr}\,[P^{l,k}_{\boldsymbol{y}\boldsymbol{g}} \hat{B}^k \rho_{\boldsymbol{x}_{1:k}\boldsymbol{y}} (\hat{B}^\dagger)^k] = 1 - O(\frac{l}{2^{l-k}}) \,, \tag{2.186}$$

it is enough to prove that $\mathrm{Tr}\,[P^{l,k}_{\boldsymbol{y}\boldsymbol{g}} \hat{B}^k \rho_0 (\hat{B}^\dagger)^k]$ does not depend on $\boldsymbol{g}$ in the cases $k = 1$, and $k = 2$.

Consider the kth iteration of the baker's map

$$\hat{B}^k |\xi^0_{1:N}\rangle = \sum_{\xi^k_{1:N}} C^{k\mathrm{th}}_n(\xi^0_{1:N}, \xi^k_{1:N}) |\xi^k_{1:N}\rangle \,. \tag{2.187}$$

This can also be written in terms of the single iteration coefficients as

$$\hat{B}^k |\xi^0_{1:N}\rangle = \sum_{\xi^{1:k}_{1:N}} \left[\prod_{p=1}^{k} C^{1\mathrm{st}}_n(\xi^{p-1}_{1:N}, \xi^p_{1:N}) \right] |\xi^k_{1:N}\rangle \,, \tag{2.188}$$

which means that

$$C^{k\mathrm{th}}_n(\xi^0_{1:N}, \xi^k_{1:N}) = \sum_{\xi^{1:k-1}_{1:N}} \left[\prod_{p=1}^{k} C^{1\mathrm{st}}_n(\xi^{p-1}_{1:N}, \xi^p_{1:N}) \right] \,. \tag{2.189}$$

Using the results of Ref. [65] we have

$$C_n^{\rm 1st}(\xi_{1:N}^{p-1},\xi_{1:N}^{p}) = \Phi(\xi_1^{p-1},\xi_N^p)\frac{\delta(\xi_{n+2:N}^{p-1}-\xi_{n+1:N-1}^{p})}{2^{n+1}\sin\phi_n^p}\,, \tag{2.190}$$

where $\phi_n^p = \pi(0.\xi_{n+1:1}^{p-1}1 - 0.\xi_{n:1}^{p}10)$ and Φ is a phase factor given by

$$\Phi(\xi_1^{p-1},\xi_N^p) = \frac{1}{\sqrt{2}}[i(-1)^{\xi_N^p} - (-1)^{\xi_1^{p-1}}]\,. \tag{2.191}$$

Since Φ depends only on ξ_1^{p-1} and ξ_N^p we have

$$C_n^{k\rm th}(\xi_{1:N}^0,\xi_{1:N}^k) = \sum_{\xi_1^{1:k-1}}\sum_{\xi_N^{1:k-1}}\left[\left(\prod_{p=1}^{k}\Phi(\xi_1^{p-1},\xi_N^p)\right)\sum_{\xi_{2:N-1}^{1:k-1}}\prod_{p=1}^{k}\frac{\delta(\xi_{n+2:N}^{p-1}-\xi_{n+1:N-1}^{p})}{2^{n+1}\sin\phi_n^p}\right]. \tag{2.192}$$

Focusing on the sum insight the square brackets we notice that the the delta functions and the inverse sines are coupled through $\{\xi_{n+1}^s\}$ variables only, so we can partially factorize the sum

$$\sum_{\xi_{2:N-1}^{1:k-1}}\prod_{p=1}^{k}\frac{\delta(\xi_{n+2:N}^{p-1}-\xi_{n+1:N-1}^{p})}{2^{n+1}\sin\phi_n^p} = \sum_{\xi_{n+1}^{1:k-1}}\left\{\left[\sum_{\xi_{n+2:N-1}^{1:k-1}}\prod_{p=1}^{k}\delta(\xi_{n+2:N}^{p-1}-\xi_{n+1:N-1}^{p})\right]\times\left[\sum_{\xi_{2:n}^{1:k-1}}\prod_{p=1}^{k}\frac{1}{2^{n+1}\sin\phi_n^p}\right]\right\}. \tag{2.193}$$

The sum of the products of the delta functions can be easily computed. Considering $k < N-n$ we have

$$\sum_{\xi_{n+2:N-1}^{1:k-1}}\prod_{p=1}^{k}\delta(\xi_{n+2:N}^{p-1}-\xi_{n+1:N-1}^{p}) = \delta(\xi_{n+1}^{1:k-1}-\xi_{n+2:n+k}^{0})\delta(\xi_{n+1:N-k}^{k}-\xi_{n+k+1:N}^{0})\delta(\xi_{N-k+1:N-1}^{k}-\xi_N^{1:k-1})\,. \tag{2.194}$$

Putting everything back into Eq. (2.192) we have

$$C_n^{k\rm th}(\xi_{1:N}^0,\xi_{1:N}^k) = \frac{\delta(\xi_{n+1:N-k}^{k}-\xi_{n+k+1:N}^{0})}{2^{k(n+1)}}\sum_{\xi_1^{1:k-1}}\left\{\left(\prod_{p=1}^{k}\Phi(\xi_1^{p-1},\xi_{N-k+p}^{k})\right)\sum_{\xi_{2:n}^{1:k-1}}\prod_{p=1}^{k}\frac{1}{\sin\tilde{\phi}_n^p}\right\}, \tag{2.195}$$

where $\tilde{\phi}_n^p$ is equal to ϕ_n^p with ξ_{n+1}^{p-1} replaced by ξ_{n+p}^0

$$\tilde{\phi}_n^p \equiv \pi(0.\xi_{n+p}^0 \xi_{n:1}^{p-1} 1 - 0.\xi_{n:1}^p 1), \qquad p = 1, \ldots k \; . \tag{2.196}$$

Equation (2.195) can be rewritten as

$$C_n^{k\text{th}}(\xi_{1:N}^0, \xi_{1:N}^k) = \sum_{\xi_1^{1:k-1}} \left(\prod_{p=1}^{k} \Phi(\xi_1^{p-1}, \xi_{N-k+p}^k) \right) Z_1(\xi_{1:N}^0, \xi_{1:N-k}^k, \xi_1^{1:k-1}) \; , \tag{2.197}$$

where

$$Z_1(\xi_{1:N}^0, \xi_{1:N-k}^k, \xi_1^{1:k-1}) = \frac{\delta(\xi_{n+1:N-k}^k - \xi_{n+k+1:N}^0)}{2^{k(n+1)}} \sum_{\xi_{2:n}^{1:k-1}} \prod_{p=1}^{k} \frac{1}{\sin \tilde{\phi}_n^p} \; . \tag{2.198}$$

It is important to note here that Z_1 is a real function and $Z_1(\xi_{1:N}^0, \xi_{1:N-k}^k, \xi_1^{1:k-1})$ does not depend on $\xi_{N-k+1:N}^k$. Rearranging the sums and the product of Φ's in (2.197) we compute

$$\begin{aligned} |\langle x_{1:r} \boldsymbol{y} g_{1:k} | \hat{B}^k | \xi_{1:N}^0 \rangle|^2 &= |C_n^{k\text{th}}(\xi_{1:N}^0, x_{1:r} \boldsymbol{y} g_{1:k})|^2 \\ &= |\sum_{\xi_1^1} \Bigg[\Phi(\xi_1^1, g_2) \sum_{\xi_1^2} \Bigg[\Phi(\xi_1^2, g_3) \cdots \sum_{\xi_1^{k-1}} \Bigg[\Phi(\xi_1^{k-1}, g_k) \times \\ &\quad Z_1(\xi_{1:N}^0, x_{1:r} \boldsymbol{y}, \xi_1^{1:k-1}) \Bigg] \cdots \Bigg] \Bigg] |^2 \; , \end{aligned} \tag{2.199}$$

where we put the square brackets to indicate explicitely the order in which we proceed computing this quantity. Considering four possible values that $\Phi(x, g)$ takes for different values of bits x and g we have, for any real function Z of x which is independent of g,

$$\left| \sum_{x=0}^{1} \Phi(x, g) Z(x) \right| = \sqrt{\sum_{x=0}^{1} [Z(x)]^2} \; . \tag{2.200}$$

This implies that

$$\sum_{x=0}^{1} \Phi(x, g) Z(x) = e^{i\nu(g, Z(0), Z(1))} \sqrt{\sum_{x=0}^{1} [Z(x)]^2} \; , \tag{2.201}$$

where the entire dependence on g is absorbed into the phase $\nu(g, Z(0), Z(1))$. Using this property we can simplify Eq. (2.199) by calculating the sum over ξ_1^{k-1}.

This gives

$$\left|\langle x_{1:r}\boldsymbol{y}g_{1:k}|\hat{B}^k|\xi^0_{1:N}\rangle\right|^2 = |\sum_{\xi_1^1}\Big[\Phi(\xi_1^1,g_2)\sum_{\xi_1^2}\Big[\Phi(\xi_1^2,g_3)\cdots\sum_{\xi_1^{k-2}}\Big[\Phi(\xi_1^{k-2},g_{k-1})\times \\ \times Z_2(\xi^0_{1:N},x_{1:r}\boldsymbol{y},\xi_1^{1:k-2})\Big]\cdots\Big]\Big]|^2\,, \tag{2.202}$$

where

$$\left|Z_2(\xi^0_{1:N},x_{1:r}\boldsymbol{y},\xi_1^{1:k-2})\right| = \sqrt{\sum_{\xi_1^{k-1}}\left[Z_1(\xi^0_{1:N},x_{1:r}\boldsymbol{y},\xi_1^{1:k-1})\right]^2}\,. \tag{2.203}$$

We see that Eq. (2.202) has a similar form as Eq. (2.199). Unfortunately, unlike Z_1, the function Z_2 is not real, so we cannot repeat the same trick with the sums over ξ_1^{k-2}, ξ_1^{k-1},..., ξ_1.

2.9 Appendix 2.B

In this appendix we show how some type of sums can be calculated up to a correction term bounded in absolute value. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two binary strings of the same length l. Consider the expression

$$\mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] = \delta_{\boldsymbol{x}}^{\boldsymbol{y}} + O(f)\,, \tag{2.204}$$

where $\{P_{\boldsymbol{x}}\}$ is a set of mutually orthogonal projectors, $\rho_{\boldsymbol{y}}$ is a density matrix and f is some function. Naively calculating the sum over the first k bits of $\boldsymbol{x}$ we would have

$$\sum_{\boldsymbol{x}_{1:k}}\mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] = \delta_{\boldsymbol{x}_{k+1:l}}^{\boldsymbol{y}_{k+1:l}} + 2^k O(f)\,, \tag{2.205}$$

where the error term effectively increased by the factor of 2^k. We can, however, keep the error term at the original level by using the normalization condition

$$\sum_{\boldsymbol{x}}\mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] = 1\,. \tag{2.206}$$

Together with this condition, Eq. (2.204) means that

$$\mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}]\begin{cases}\leq & \epsilon & \boldsymbol{x}\neq\boldsymbol{y}\\ \geq & 1-\epsilon & \boldsymbol{x}=\boldsymbol{y}\,,\end{cases} \tag{2.207}$$

where $\epsilon = \sup |O(f)|$. Noticing that $\mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] \geq 0$ for any x, we have

$$\sum_{\boldsymbol{x}_{1:k}} \mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] \geq 1 - \epsilon\,, \qquad \boldsymbol{x}_{k+1:l} = \boldsymbol{y}_{k+1:l}\,. \tag{2.208}$$

Because of the normalization condition this implies that

$$\sum_{\boldsymbol{x}_{1:k}} \mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] \leq \epsilon\,, \qquad \boldsymbol{x}_{k+1:l} \neq \boldsymbol{y}_{k+1:l}\,, \tag{2.209}$$

or, in summary,

$$\sum_{\boldsymbol{x}_{1:k}} \mathrm{Tr}\,[P_{\boldsymbol{x}}\,\rho_{\boldsymbol{y}}] = \delta^{\boldsymbol{y}_{k+1:l}}_{\boldsymbol{x}_{k+1:l}} + O(f)\,. \tag{2.210}$$

Chapter 3

Conditional evolution in cavity QED

3.1 Introduction

In the previous chapters we introduced various measures of dynamical complexity and gave examples of how these measures can be used to quantify unpredictability of the system behaviour. We showed that a rapid loss of prediction accuracy can occur if a system with complex dynamics interacts with an unknown environment and/or if the initial state of such a system is specified with a finite precision. At first sight it may seem, however, that considerations of this type are either too general or too abstract to be useful in connection with contemporary experiments. The aim of this section is to show how the concept of dynamical complexity can be introduced in a typical experimental setup of single-atom cavity QED [31]. Currently the experiments of this type constitute an indispensable tool for the study of open quantum systems in connection with quantum chaos [41], quantum control [20], quantum computing [66] and many others.

The basic experimental scenario is shown in Fig. 3.1. The setup includes a single two-level atom located insight a high finesse optical resonator and a set of photodetectors arranged to monitor the fields escaping from the system into the en-

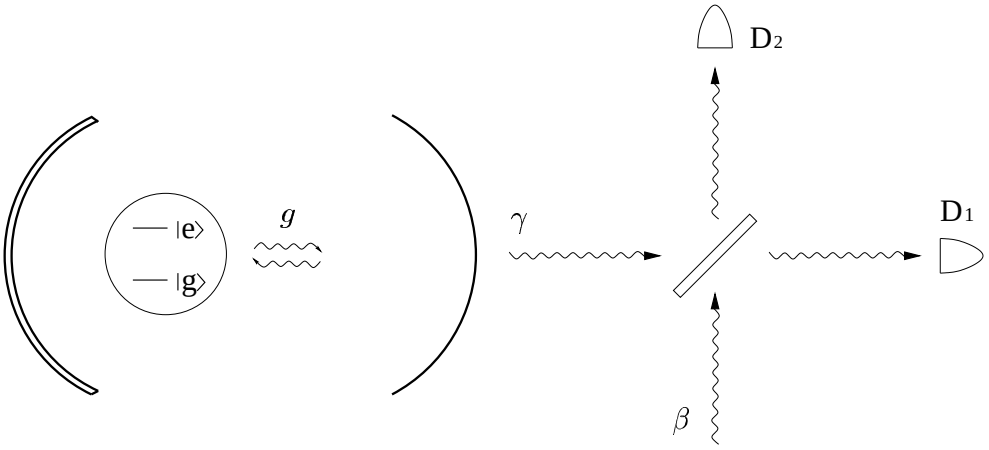


Figure 3.1: Homodyne measurements in cavity QED. Basic parameters of the system include the strength of the atom-cavity coupling g, the rate of atomic spontaneous emission into noncavity optical modes $\gamma_\perp$, and the cavity field decay rate γ. The cavity output field is added to the strong reference field β on the beam-splitter and then analyzed by the detectors D_1 and D_2.

vironment. It is convenient to distinguish two channels through which the system interacts with the environment: i) atomic spontaneous emission into the environment optical modes, and ii) the leakage of photons from the cavity mode through an output mirror. Quantitatively these two channels can be assigned characteristic dissipation rates which we denote $\gamma_\perp$ and γ for the first and the second channel respectively. In what follows we assume that only the second channel is monitored via a continuous (homodyne) measurement. Also, for simplicity, we adjust the cavity length so that the frequency of the resonant optical mode coincides with the frequency of the atomic transition. The dynamics of such a system can be characterized by the dissipation rates γ and $\gamma_\perp$ just introduced, and one extra parameter g which quantifies the strength of the atom-cavity coupling (see the next chapter for details).

In the strong coupling regime ($g \gg \gamma,\ \gamma_\perp$) the atom and the cavity can be viewed as one strongly bonded quantum system, and the effects of the relatively small dissipation can then be viewed as perturbations of the system by the environment.

It is important to emphasize that this situation is already available for experimental investigation [31, 43]. At the same time, this is exactly the situation considered in a somewhat abstract form in the first chapter of this thesis. There we introduced a measure of dynamical complexity and showed how this measure can be used to quantify the sensitivity of the system dynamics to the perturbations by the environment. We argued that this measure can be used to quantify the degree of chaoticity of the system dynamics and gave some examples. Mathematically, two different concepts were used in the definition of dynamical complexity χ. These are algorithmic information and von Neumann entropy. Algorithmic information was used to quantify the information content of the measurement record whereas von Neumann entropy was used to quantify the uncertainty about the system state. Both of these concepts are important in the general definition of dynamical complexity (1.15). We can rewrite this definition as

$$\chi = \frac{\bar{I}_{\min}/\Delta t}{\Delta H/\Delta t}\,, \tag{3.1}$$

where Δt is the interval of time during which the system entropy has changed by ΔH. We see that the problem of calculating χ can be split into two important subproblems. Given the type of measurements, the first subproblem is the rate of entropy production $\Delta H/\Delta t$. This rate gives the behaviour in time of the uncertainty about the system state. The second subproblem is the corresponding rate $\bar{I}_{\min}/\Delta t$ at which the information about the system should be supplied via the measurements to maintain the entropy production at the rate $\Delta H/\Delta t$. In general, both of these subproblems are important, however in particular applications it may be reasonable to concentrate most of the attention on one of the subproblems. For instance, one of the applications of single-atom cavity QED is the Atom-Cavity microscope depicted in Fig. 3.2. The basic idea of the microscope is to infer the position of the atom inside the cavity on the basis of the observed photocurrents which depend on this position. Questions concerning the accuracy of such a microscope are certainly of major importance. This means that the subproblem of system entropy production becomes interesting in its own right. In the field of quantum chaos this subproblem also receives special attention [75]. For these rea-

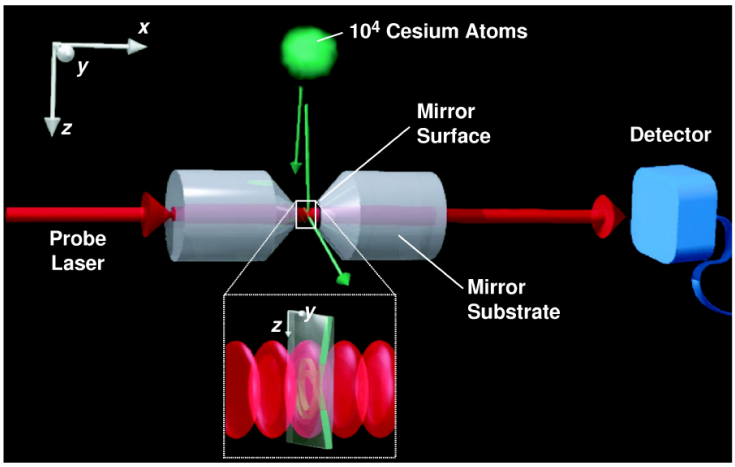


Figure 3.2: Cesium atoms are stored in a magneto-optical trap and dropped through a high-finesse optical cavity. A single atom (green arrow) transiting the cavity mode substantially alters the measured transmission of a probe beam through the cavity. [This figure is taken from [32].]

sons, we will devote this chapter to the problem of calculation of the change of the system entropy in a typical single-atom cavity QED experiment. We defer the information-theoretic analysis of the measurement record as a project for future research (see the conclusions).

The rate of system entropy production quantifies the rate at which the information about the system dissipates into the environment. In our case, the preparation of any state which is significantly different from a steady state is very difficult [21]. To avoid complicated state preparation procedures which are irrelevant to our research it is therefore convenient to wait until the system has reached a steady state and then start observations. In such situations the entropy of the system will be decreasing due to the observations. Formulating the final results we will always assume, unless otherwise stated, that the system is initially in a steady state.

This chapter is organized as follows. In section 3.2 we introduce the mathematical model and explain the main approximations used. In section 3.3 we calculate analytically the rate of entropy reduction $-dH/dt$ using the method of stochastic master equations. The result is the formula for "optical information" conjectured and used by Kimble in [32] to illustrate the accuracy achievable by the Atom-Cavity microscope. The method of stochastic master equations has a number of built-in approximations associated with taking the continuum limit when formulating the differential form of the master equations. These additional approximations implicitly assume that all the detectors have zero response time: photocounts can be resolved during any infinitesimally small interval of time. To avoid this unphysical assumption we show how the rate of system entropy reduction can be calculated from first principles, i.e. without using the method of stochastic master equations. These calculations are presented in section 3.4, which also contains results concerning the steady state of the system.

3.2 Mathematical model and main approximations

Let $|\mathrm{g}\rangle$ and $|\mathrm{e}\rangle$ be the ground and excited states of the atom. For simplicity, we choose the cavity length so that the frequency of the resonant optical mode coincides with the frequency of the atomic transition. In the dipole approximation the interaction of the two-level atom with the electromagnetic field inside the cavity is described by the Hamiltonian [42]

$$H_{\mathrm{int}} = i\frac{g}{2}(a + a^{\dagger})(\sigma - \sigma^{\dagger})\,, \tag{3.2}$$

where $\sigma = |\mathrm{g}\rangle\langle\mathrm{e}|$, g is the strength of the atom-cavity coupling, and a is the annihilation operator for the intracavity field. Including dissipation and on-resonant driving of the cavity mode, the total unconditional master equation in a frame rotating at the driving laser frequency reads:

$$\begin{aligned}\dot{\rho} &= [-iH_{\rm int} + E(a^\dagger - a), \rho\,] \\ &\quad + \frac{\gamma}{2}(2a^\dagger \rho a - a^\dagger a \rho - \rho a^\dagger a) + \frac{\gamma_\perp}{2}(2\sigma^\dagger \rho \sigma - \sigma^\dagger \sigma \rho - \rho \sigma^\dagger \sigma)\,, \end{aligned} \tag{3.3}$$

where E is the strength of the driving, $\gamma_\perp$ is the rate of atomic spontaneous emission into noncavity optical modes and γ is the energy loss rate due to the leakage of photons from the cavity mode through an output mirror ($\gamma = Tc/2L$, where T is the mirror transmission coefficient, L is the length of the cavity and c is the speed of light). In the rotating wave approximation, this equation becomes what is well known as the driven Jaynes-Cummings model with dissipation[1]. This approximation is based on ignoring rapidly changing terms in the master equation, which is equivalent to ignoring the terms proportional to $a\sigma$ and $a^\dagger\sigma^\dagger$ in $H_{\rm int}$[2]. In what follows we will always ignore the spontaneous emission into noncavity modes by setting $\gamma_\perp = 0$.

From the experimental point of view the question of the steady state is very important. In fact, using contemporary techniques it is very difficult to prepare the system in question in any other state. Using the Jaynes-Cummings model, Alsing and Carmichael [3] have shown numerically that in the case of small $\gamma_\perp$ and in the strong driving limit $E \gg g$ the system approaches a steady state of the form

$$\rho_{\rm ss} = \frac{1}{2}(|\alpha;+\rangle\langle\alpha;+| + |\alpha^*;-\rangle\langle\alpha^*;-|)\,, \tag{3.4}$$

where $|\alpha;+\rangle$ and $|\alpha^*;-\rangle$ are two orthogonal quantum states

$$\begin{aligned} |\alpha;+\rangle &= \frac{1}{\sqrt{2}}|\alpha\rangle(|{\rm g}\rangle + i|{\rm e}\rangle)\,, \\ |\alpha^*;-\rangle &= \frac{1}{\sqrt{2}}|\alpha^*\rangle(|{\rm g}\rangle - i|{\rm e}\rangle)\,, \end{aligned} \tag{3.5}$$

where $|\alpha\rangle$ is the coherent field state with amplitude

$$\alpha = (2E + ig)/\gamma\,. \tag{3.6}$$

[1] Consult Ref. [3] and references therein

[2] Consult Ref. [42], p. 324 for details.

This result has recently been confirmed in another numerical simulation [44]. In section 3.4 we will show *analytically* that for $\gamma_\perp = 0$ and for any driving E equation (3.3) has a steady state as defined by (3.4). We thus show the validity of Eq. (3.4) without using either the rotating wave approximation or the strong driving limit.

In conclusion of this section we introduce some new notation which will be used in the following sections. We introduce a basis

$$|+\rangle \equiv \frac{1}{\sqrt{2}}(|\mathrm{g}\rangle + i|\mathrm{e}\rangle) \quad \text{and} \quad |-\rangle \equiv \frac{1}{\sqrt{2}}(|\mathrm{g}\rangle - i|\mathrm{e}\rangle)\,. \tag{3.7}$$

In this basis and assuming $\gamma_\perp = 0$ equation (3.3) becomes

$$\dot{\rho} = [\, ig\frac{\mu_z}{2}(a^\dagger + a) - E(a - a^\dagger)\,, \rho\,] + \frac{\gamma}{2}(2a\rho a^\dagger - a^\dagger a\rho - \rho a^\dagger a) \tag{3.8}$$

where

$$\mu_z \equiv |+\rangle\langle +| - |-\rangle\langle -|\,. \tag{3.9}$$

For further reference it is convenient to rewrite (3.4) using matrix notation. In the basis given by (3.5) we have

$$\rho_{\mathrm{ss}} = \frac{1}{2}\mathbb{1}_2\,, \tag{3.10}$$

where $\mathbb{1}_2$ is a two-dimensional identity matrix. Alternatively, we can use the matrix notation only for the intra-atomic degrees of freedom and rewrite ρ_{ss} as

$$\rho_{\mathrm{ss}} = \frac{1}{2}\begin{pmatrix} |\alpha\rangle\langle\alpha| & 0 \\ 0 & |\alpha^*\rangle\langle\alpha^*| \end{pmatrix}\,. \tag{3.11}$$

The first of these representations, Eq. (3.10), is convenient when calculating the infinitesimal change of the system entropy directly from the stochastic master equation (section 3.3). The second representation, Eq. (3.11), makes a clear distinction between the atomic and the field degrees of freedom. This is convenient in the derivation of the analytic expression for the change of the conditional density matrix conditioned on the results homodyne measurements during an arbitrary finite interval of time (section 3.4).

3.3 Calculations of $\Delta H/\Delta t$ using the method of stochastic master equations

3.3.1 The method of stochastic master equations

Consider the case when the cavity output undergoes a continuous homodyne measurement[3] as explained in Fig. 3.1. In this case the master equation for the conditional density matrix $\rho_{\rm c}$, conditional on the measurement records, can be obtained from the unconditional equation (3.8) by adding a nonlinear stochastic term [44, 73]. Apart from the dependence on a particular measurement history this term also depends on the phase of the reference laser. This phase as well as the intensity of the laser are parameterized by one complex number which in Fig. 3.1 is denoted by β. This is the only free parameter of the measurement scheme. If the measurement record consists of the difference photocurrent $I_- = I_2 - I_1$, where I_1 and I_2 are the photocurrents detected by the first and the second detectors respectively, then the conditional master equation becomes [72, 73]

$$\begin{aligned}\dot{\rho}_{\rm c} &= [\, ig\frac{\mu_z}{2}(a^\dagger + a) - E(a - a^\dagger)\,, \rho_{\rm c}\,] + \frac{\gamma}{2}(2a\rho_{\rm c}a^\dagger - a^\dagger a\rho_{\rm c} - \rho_{\rm c}a^\dagger a) \\ &\quad + \sqrt{\gamma\eta}\,\frac{dW}{dt}\Big(e^{-i\phi}a\rho_{\rm c} + e^{i\phi}\rho_{\rm c}a^\dagger - {\rm tr}\big[\rho_{\rm c}(e^{-i\phi}a + e^{i\phi}a^\dagger)\big]\,\rho_{\rm c}\Big)\,, \qquad (3.12)\end{aligned}$$

where η is the efficiency of the photodetection, ϕ is the phase $\phi = \arg\beta$, and dW/dt is the time derivative of the random Wiener variable which, in practice, should be taken from experimental observations of the difference photocurrent I_- via the relation:

$$I_- = \frac{dQ}{dt} = \gamma\eta|\beta|\,{\rm tr}\big[\rho_{\rm c}(e^{i\phi}a^\dagger + e^{-i\phi}a)\big] + \sqrt{\gamma\eta}|\beta|\frac{dW}{dt}\,. \qquad (3.13)$$

Because of the stochastic Wiener variable W, which connects the conditional master equation (3.12) with the corresponding photocurrent, we shall refer to this method as the method of stochastic master equations. This method relies on the assumption that during a however small interval of time dt there exists a nonzero

[3] For detailed review consult [51] and references therein

photocount $dQ = I_- dt$. This assumption is valid only for ideal detectors with infinitesimally small response time. In practice, we should improve the theory to account for the case of more realistic detectors.

For the reasons explained in the introduction, we consider the case when the system reaches the steady state before any measurements begin at the time t_0. In this case, at t_0 the first two terms in (3.12) vanish and the infinitesimal change of the conditional density matrix $d\rho_{\rm c}(t_0)$ caused by the measurements over the infinitesimal time dt is given by

$$d\rho_{\rm c}(t_0) = \sqrt{\gamma\eta}\left(e^{-i\phi}a\rho_{\rm ss} + e^{i\phi}\rho_{\rm ss}a^\dagger - {\rm tr}\left[\rho_{\rm ss}(e^{-i\phi}a + e^{i\phi}a^\dagger)\right]\ \rho_{\rm ss}\right)dW, \qquad (3.14)$$

This change of the density matrix corresponds to the detection of the charge increment

$$dQ(t_0) = \gamma\eta|\beta|\ {\rm tr}\left[\rho_{\rm ss}(e^{i\phi}a^\dagger + e^{-i\phi}a)\right]\,dt + \sqrt{\gamma\eta}|\beta|dW. \qquad (3.15)$$

In the next section we will prove that the steady state of the unconditional equation (3.8) is given by (3.11). Later we will be interested in the two special cases $\phi = \pi/2$ and $\phi = 0$. Substituting (3.11) into Eq. (3.14) gives for the case of $\phi = \pi/2$:

$$d\rho_{\rm c}(t_0) = g\sqrt{\frac{\eta}{\gamma}}\begin{pmatrix} -|\alpha\rangle\langle\alpha| & 0 \\ 0 & |\alpha^*\rangle\langle\alpha^*| \end{pmatrix} dW, \quad \phi = \pi/2\,. \qquad (3.16)$$

In the basis (3.5) this can be rewritten as

$$d\rho_{\rm c}(t_0) = g\sqrt{\frac{\eta}{\gamma}}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} dW, \quad \phi = \pi/2\,. \qquad (3.17)$$

The corresponding change of the measured charge is given by Eq. (3.15) which now becomes

$$dQ(t_0) = \sqrt{\gamma\eta}\,dW, \quad \phi = \pi/2\,. \qquad (3.18)$$

In the case $\phi = 0$ we get

$$d\rho_{\rm c}(t_0) \propto \rho_{\rm ss}\,. \qquad (3.19)$$

By normalization, this means that the conditional density matrix coincides with the initial steady state $\rho_{\rm ss}$ irrespective of the measurement results. This observation will guide our intuition in section 3.4 in an otherwise formal proof that $\rho_{\rm ss}$ is a steady state of equation (3.8).

3.3.2 Entropy change

By definition, the entropy of a system described by a density matrix ρ is

$$\begin{aligned} H(\rho) &= -\mathrm{tr}(\rho \ln \rho) \\ &= \mathrm{tr}[\sum_{k=1}^{\infty} \frac{(-1)^k}{k} \rho(\rho - \mathbb{1})^k] \,. \end{aligned} \tag{3.20}$$

The average entropy change in the presence of continuous observations is given by

$$\overline{\Delta H} = \langle H(\rho_0 + \Delta\rho_{\rm c}) - H(\rho_0) \rangle \,, \tag{3.21}$$

where the average is taken over all possible measurement outcomes observed over the time Δt, and $\Delta\rho_{\rm c}$ is the difference between the conditional density matrix, conditioned on the measurement outcomes, and the initial state ρ_0 just before the measurements.

Consider the case when the observations during an infinitesimal interval of time dt result in the change of the density matrix

$$d\rho_{\rm c} = A\, dW, \tag{3.22}$$

where A is some operator and dW is an infinitesimal Wiener increment which satisfies the rules of Ito stochastic differential calculus [22]. In particular, dW has the following two properties

$$\langle dW \rangle = 0 \,, \quad (dW)^2 = dt \,. \tag{3.23}$$

After substituting (3.20) into (3.21) and retaining the terms to the second order in dW, we can use the above properties to show that

$$\overline{dH} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \qquad \left\{ \sum_{l=0}^{k-2} \mathrm{tr}[\rho_0 A\, (\rho_0 - \mathbb{1})^l A\, (\rho_0 - \mathbb{1})^{k-l-2}] \right.$$

$$+\sum_{l=0}^{k-1}\mathrm{tr}[A\,(\rho_0-\mathbb{1})^l A\,(\rho_0-\mathbb{1})^{k-l-1}]\Bigg\}\,dt\;, \qquad (3.24)$$

where we adopt the usual convention $\sum_{k=a}^{b} c_k = 0$ if $b < a$.

Kimble's formula

If the steady state of the system is given by Eq. (3.10) we have by comparing (3.17) with (3.22) that

$$A = -g\sqrt{\frac{\eta}{\gamma}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\;. \qquad (3.25)$$

Substituting this into Eq. (3.24) we have

$$\overline{dH} = -4\frac{g^2\eta}{\gamma}\sum_{k=1}^{\infty}\frac{1}{2^k k}\,dt\;. \qquad (3.26)$$

Using the identity

$$\sum_{k=1}^{\infty}\frac{1}{2^k k} = \ln 2 \qquad (3.27)$$

we finally have

$$\overline{dH} = -4\frac{g^2\eta}{\gamma}\ln 2\;dt\,. \qquad (3.28)$$

It is customary to measure entropy in bits which can be done by replacing the natural logarithm in the definition (3.20) by $\log_2$ to give

$$\overline{dH}_2 \equiv \frac{1}{\ln 2}\overline{dH} = -4\frac{g^2\eta}{\gamma}\,dt\,. \qquad (3.29)$$

Up to the prefactor of four, this expression is identical to the measure of "optical information" which was used as a heuristic quantity in Ref. [32]. It is relevant to mention that this method of calculating the system entropy reduction caused by the continuous homodyne monitoring of the cavity output field is rather general. Knowledge of the steady state of the system is enough to perform the calculations (provided of course that the approximations assumed by the method of stochastic master equation are satisfied).

3.4 Analytic solution for the conditional evolution conditioned on a discrete measurement record

3.4.1 General considerations

In this section we will focus our attention on equation (3.8) which we rewrite in the form:

$$\dot{\rho} = \mathcal{L}\rho\,, \tag{3.30}$$

where superoperator $\mathcal{L}$ is defined as

$$\mathcal{L}\rho \equiv [\, ig\frac{\mu_z}{2}(a^\dagger + a) - E(a - a^\dagger), \rho\,] + \frac{\gamma}{2}(2a\rho a^\dagger - a^\dagger a\rho - \rho a^\dagger a)\,. \tag{3.31}$$

In what follows it is convenient to use the notation of Eq. (3.11). In this notation we have

$$\mu_z = \begin{pmatrix} \mathbb{1}_a & 0 \\ 0 & -\mathbb{1}_a \end{pmatrix}\,, \tag{3.32}$$

where $\mathbb{1}_a$ is an identity operator on the field degrees of freedom. Given superoperators $\mathcal{S}$, $\mathcal{J}_1$ and $\mathcal{J}_2$ such that

$$\mathcal{S}(t) = e^{(\mathcal{L}-\mathcal{J}_1-\mathcal{J}_2)t}\,, \tag{3.33}$$

and the initial condition $\rho(0) = \rho_0$, the solution to equation (3.30) can be written using a Dyson expansion,

$$\rho(\Delta t) = \sum_{m=0}^{\infty} \sum_{k_1,\ldots,k_m} p(k_1,\ldots,k_m;\Delta t)\rho_{\rm c}(k_1,\ldots,k_m;\Delta t)\,, \tag{3.34}$$

where $\mathrm{tr}\rho_{\rm c}(k_1,\ldots,k_m;\Delta t) = 1$ and

$$\begin{aligned} &p(k_1,\ldots,k_m;\Delta t)\rho_{\rm c}(k_1,\ldots,k_m;\Delta t) \\ &\quad = \int_0^{\Delta t} dt_m \cdots \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \mathcal{S}(\Delta t - t_m)\mathcal{J}_{k_m}\mathcal{S}(t_m - t_{m-1})\mathcal{J}_{k_{m-1}}\cdots\mathcal{S}(t_1)\rho_0\,. \end{aligned} \tag{3.35}$$

Following [73] we define the "smooth evolution" operator $\mathcal{S}$ as

$$\mathcal{S}(t)\rho \equiv N(t)\rho N^\dagger(t)\ , \text{ where} \tag{3.36}$$

$$N(t) \equiv \exp[ig\frac{\mu_z}{2}(a^\dagger + a)t + E(a^\dagger - a)t - \frac{\gamma}{2}(a^\dagger a + |\beta|^2)t]\ , \tag{3.37}$$

and the "jump" operators $\mathcal{J}_1$ and $\mathcal{J}_2$ as

$$\mathcal{J}_k\rho \equiv C_k\rho C_k^\dagger\ , \quad \text{where } C_k \equiv \sqrt{\gamma/2}e^{i\pi(k-1)/2}[a + (-1)^k\beta]\ . \tag{3.38}$$

Lemma 1

The above definitions satisfy the requirement

$$\mathcal{S}(t) = e^{(\mathcal{L}-\mathcal{J}_1-\mathcal{J}_2)t} \tag{3.39}$$

and therefore Eqs. (3.34) and (3.35) indeed give a solution to (3.30).

Proof

Keeping the terms to the first order in δt we have

$$\begin{aligned}
\mathcal{S}(\delta t)\rho &= N(\delta t)\rho[N(\delta t)]^\dagger \\
&= \rho + \Big([ig\frac{\mu_z}{2}(a + a^\dagger) + E(a^\dagger - a), \rho] \\
&\qquad - \frac{\gamma}{2}(a^\dagger a\rho + \rho a^\dagger a) - \gamma|\beta|^2\rho\Big)\delta t + O(\delta t^2) \\
&= (\mathbb{1} + \delta t\cdot\mathcal{L})\rho - \gamma(a\rho a^\dagger + |\beta|^2\rho)\delta t + O(\delta t^2)\ .
\end{aligned} \tag{3.40}$$

On the other hand by direct calculation we have

$$\mathcal{J}_k\rho = \frac{\gamma}{2}[a\rho a^\dagger + (-1)^k(\beta\rho a^\dagger + \beta^* a\rho) + |\beta|^2\rho] \tag{3.41}$$

which implies that

$$(\mathcal{J}_1 + \mathcal{J}_2)\rho = \gamma(a\rho a^\dagger + |\beta|^2\rho)\ . \tag{3.42}$$

Equation (3.40) therefore becomes

$$\mathcal{S}(\delta t)\rho = (\mathbb{1} + \delta t\cdot[\mathcal{L} - (\mathcal{J}_1 + \mathcal{J}_2)])\rho + O(\delta t^2) \tag{3.43}$$

Taking the limit $\delta t \to 0$ we have Eq. (3.39) as required. □

There are many different definitions of $\mathcal{S}$, $\mathcal{J}_1$ and $\mathcal{J}_2$ which satisfy the above lemma. However, definitions (3.36) and (3.38) are somewhat special: the quantities $\rho_{\rm c}(k_1,\ldots,k_m;\Delta t)$ and $p(k_1,\ldots,k_m;\Delta t)$ which they define have an important physical meaning. Suppose that the continuous measurements were performed over the interval of time Δt and recorded as a sequence $(k_1,\ldots,k_m;\Delta t)$ of photodetector labels in the order of photodetections[4]. Then, the conditional density matrix conditioned on the measurement record $(k_1,\ldots,k_m;\Delta t)$ is given by $\rho_{\rm c}(k_1,\ldots,k_m;\Delta t)$, and the corresponding probability is $p(k_1,\ldots,k_m;\Delta t)$.

Now we shall prove some auxiliary theorems which can be used for the direct calculation of $\rho_{\rm c}(k_1,\ldots,k_m;\Delta t)$ from Eq. (3.35). It may be helpful to note that necessary simplification of Eq. (3.35) can be achieved if we know how to swap the smooth evolution $\mathcal{S}$ with the jump operators $\mathcal{J}_k$.

Theorem 6

The operator $M(t)\equiv e^{\gamma|\beta|^2t/2}N(t)$ can be factorized as

$$M(t)=e^{Z_1}e^{-\frac{\gamma t}{2}a^\dagger a}e^{Z_2a^\dagger}e^{Z_3a}\ , \tag{3.44}$$

where

$$\begin{aligned} Z_1 &= \frac{4E^2+g^2}{\gamma^2}(1-e^{-\gamma t/2}-\gamma t/2)\\ Z_2 &= \frac{2E+ig\mu_z}{\gamma}(e^{\gamma t/2}-1)\\ Z_3 &= \frac{2E-ig\mu_z}{\gamma}(e^{-\gamma t/2}-1) \end{aligned} \tag{3.45}$$

Proof

By definition

$$M(t)=\exp[ig\frac{\mu_z}{2}(a^\dagger+a)t+E(a^\dagger-a)t-\frac{\gamma t}{2}a^\dagger a]\ . \tag{3.46}$$

Because a, $a^\dagger$, $a^\dagger a$ and $\mathbb{1}$ span a Lie algebra, $M(t)$ can be factorized in a systematic way as follows. First we find a function $x(t)$ such that

$$M(t)=e^{x(t)a^\dagger a}\tilde{M}(t)\ , \tag{3.47}$$

[4]For example, $k_j=1$ would mean that the jth photodetection was registered by the first detector.

where $\tilde{M}(t)$ is an exponential of a linear combination of a and $a^\dagger$. We will then repeat the same procedure factorizing $\tilde{M}$ which will conclude the prove of the theorem.

Equation (3.47) gives

$$\dot{M} = \dot{x}a^\dagger a\, e^{xa^\dagger a}\tilde{M} + e^{xa^\dagger a}\dot{\tilde{M}} \; . \tag{3.48}$$

On the other hand, equation (3.46) gives

$$\dot{M} = [ig\frac{\mu_z}{2}(a^\dagger + a) + E(a^\dagger - a) - \frac{\gamma}{2}a^\dagger a]e^{xa^\dagger a}\tilde{M} \; . \tag{3.49}$$

Comparing this expression with the previous one we have

$$\dot{\tilde{M}} = [g(x) - (\dot{x} + \gamma/2)a^\dagger a]\tilde{M} \; , \tag{3.50}$$

where

$$g(x) = e^{-xa^\dagger a}[ig\frac{\mu_z}{2}(a^\dagger + a) + E(a^\dagger - a)]e^{xa^\dagger a} \; . \tag{3.51}$$

According to Identities 1 and 2 (Eqs. (3.101) and (3.105) in the Appendix), the above equation can be rewritten as

$$g(x) = (E + ig\frac{\mu_z}{2})e^{-x}a^\dagger - (E - ig\frac{\mu_z}{2})e^{x}a \; . \tag{3.52}$$

Looking at equation (3.50) we demand that

$$\dot{x} + \gamma/2 = 0 \; , \tag{3.53}$$

thereby making $\tilde{M}$ independent of $a^\dagger a$. From equation (3.46) we see that $M(0) = \mathbb{1}$ and therefore, equation (3.47) implies that

$$x(0) = 0 \;\; \text{and} \;\; \tilde{M}(0) = \mathbb{1} \; . \tag{3.54}$$

With these restrictions equations (3.53) and (3.50) can be integrated to give, according to Eq. (3.47),

$$M(t) = e^{-\frac{\gamma t}{2}a^\dagger a}\tilde{M}(t) \; , \tag{3.55}$$

where

$$\tilde{M} = \text{texp}[(2E + ig\mu_z)\frac{a^\dagger}{\gamma}(e^{\gamma t/2} - 1) + (2E - ig\mu_z)\frac{a}{\gamma}(e^{-\gamma t/2} - 1)] \, . \tag{3.56}$$

The proof of the theorem will be completed if we repeat the same procedure for factorizing $\tilde{M}$. As before we introduce a function $y(t)$ such that

$$\tilde{M}(t) = e^{y(t)a^\dagger}\tilde{M}'(t)\,. \tag{3.57}$$

We therefore have

$$\dot{\tilde{M}} = \dot{y}a^\dagger e^{ya^\dagger}\tilde{M}' + e^{ya^\dagger}\dot{\tilde{M}}'\,. \tag{3.58}$$

On the other hand, equation (3.56) gives

$$\dot{\tilde{M}} = [(E + ig\frac{\mu_z}{2})e^{\gamma t/2}a^\dagger - (E - ig\frac{\mu_z}{2})e^{-\gamma t/2}a]e^{ya^\dagger}\tilde{M}'\,. \tag{3.59}$$

The last two expressions imply

$$\dot{\tilde{M}}' = [(E + ig\frac{\mu_z}{2})e^{\gamma t/2}a^\dagger - (E - ig\frac{\mu_z}{2})e^{-\gamma t/2}e^{-ya^\dagger}ae^{ya^\dagger} - \dot{y}a^\dagger]\tilde{M}'\,. \tag{3.60}$$

Using Identity 3 (Eq. (3.106) in the Appendix) we rewrite this expression as

$$\dot{\tilde{M}}' = \left([(E + ig\frac{\mu_z}{2})e^{\gamma t/2} - \dot{y}]a^\dagger - (E - ig\frac{\mu_z}{2})e^{-\gamma t/2}(a + y)\right)\tilde{M}'\,. \tag{3.61}$$

We eliminate $a^\dagger$ from this expression by setting

$$\dot{y} = (E + ig\frac{\mu_z}{2})e^{\gamma t/2}\,. \tag{3.62}$$

Equation (3.57) gives the boundary conditions

$$y(0) = 0 \quad \text{and} \quad \tilde{M}'(0) = \mathbb{1}\,. \tag{3.63}$$

Performing integration in (3.61) and in (3.62) using these boundary conditions and the fact that $\mu_z^2 = \mathbb{1}$ we have according to (3.57)

$$\tilde{M}(t) = \exp[\frac{2E + ig\mu_z}{\gamma}(e^{\gamma t/2} - 1)a^\dagger]\tilde{M}'(t)\,, \tag{3.64}$$

where

$$\tilde{M}'(t) = \exp[\frac{4E^2 + g^2}{\gamma^2}(1 - e^{-\gamma t/2} - \gamma t/2)]\exp[-\frac{2E - ig\mu_z}{\gamma}(1 - e^{-\gamma t/2})a]\,. \tag{3.65}$$

This completes the proof of the theorem. □

Theorem 7

Using the definition

$$f_k \equiv \sqrt{\gamma/2}\, e^{i\pi(k-1)/2} \tag{3.66}$$

we have

$$C_k M(t) = M(t) f_k [e^{-\gamma t/2} a + \frac{1 - e^{-\gamma t/2}}{\gamma}(2E + ig\mu_z) + (-1)^k \beta] \,. \tag{3.67}$$

Proof

By definition [Eqs. (3.38) and (3.66)] and using Theorem 6 we have

$$C_k M(t) = f_k e^{Z_1} [a + (-1)^k \beta] e^{-\frac{\gamma t}{2} a^\dagger a} e^{Z_2 a^\dagger} e^{Z_3 a} \,, \tag{3.68}$$

where Z_1, Z_2 and Z_3 are specified in the statement of Theorem 6. Using Identity 2 and then Identity 3, given in the Appendix as equations (3.105) and (3.106), we have

$$\begin{aligned} C_k M(t) &= f_k e^{Z_1} e^{-\frac{\gamma t}{2} a^\dagger a} [e^{-\gamma t/2} a + (-1)^k \beta] e^{Z_2 a^\dagger} e^{Z_3 a} \\ &= f_k e^{Z_1} e^{-\frac{\gamma t}{2} a^\dagger a} e^{Z_2 a^\dagger} [e^{-\gamma t/2}(a + Z_2) + (-1)^k \beta] e^{Z_3 a} \\ &= f_k M(t) [e^{-\gamma t/2}(a + Z_2) + (-1)^k \beta] \,. \end{aligned} \tag{3.69}$$

Putting the value of Z_2 from Theorem 6 we have Eq. (3.67) as required. □

Using these theorems we can proceed with the calculation of the conditional density matrix $\rho_c(k_1, \ldots, k_m; \Delta t)$. We have from Eqs. (3.35), (3.36) and (3.38) that

$$\begin{aligned} & p(k_1, \ldots, k_m; \Delta t) \rho_c(k_1, \ldots, k_m; \Delta t) \\ &= \frac{1}{m!} \int_0^{\Delta t} dt_m \cdots \int_0^{\Delta t} dt_2 \int_0^{\Delta t} dt_1 [N(\Delta t - t_m) C_{k_m} \cdots N(t_2 - t_1) C_{k_1} N(t_1)] \rho_0 [\cdots]^\dagger \,. \end{aligned} \tag{3.70}$$

We can now use Theorem 7 to compute the operator in the square brackets. We have, for instance,

$$C_{k_1} N(t_1) = N(t_1) f_k [e^{-\gamma t_1/2} a + \frac{1 - e^{-\gamma t_1/2}}{\gamma}(2E + ig\mu_z) + (-1)^{k_1} \beta] \,. \tag{3.71}$$

Then, using the identity $N(t_2-t_1)N(t_1)=N(t_2)$, we see that repeating the same type of calculations we have

$$N(\Delta t-t_m)C_{k_m}\cdots N(t_2-t_1)C_{k_1}N(t_1)=N(\Delta t)\prod_{p=1}^{m}f_{k_p}[e^{-\gamma t_p/2}a+\frac{1-e^{-\gamma t_p/2}}{\gamma}(2E+ig\mu_z)+(-1)^{k_p}\beta \tag{3.72}$$

Using the identity $f_k f_k^*=\gamma/2$ we therefore have

$$p(k_1,\ldots,k_m;\Delta t)\rho(k_1,\ldots,k_m;\Delta t)=\frac{\gamma^m}{2^m m!}N(\Delta t)R(\rho_0,\beta)N^\dagger(\Delta t)\,, \tag{3.73}$$

where

$$R(\rho_0,\beta)=\int_0^{\Delta t}dt_m\int_0^{\Delta t}dt_{m-1}\cdots\int_0^{\Delta t}dt_1 \\ \Big(\prod_{p=1}^{m}e^{-\gamma t_p/2}[a+\frac{e^{\gamma t_p/2}-1}{\gamma}(2E+ig\mu_z)+(-1)^{k_p}\beta e^{\gamma t_p/2}]\Big)\rho_0\Big(\cdots\Big)^\dagger\,. \tag{3.74}$$

For notation convenience, we do not indicate explicitely the dependence of $R(\rho_0,\beta)$ on the measurement record $(k_1,\ldots,k_m;\Delta t)$ which, however, should always be remembered. In conclusion of this section we consider the case $\rho_0=\rho_{\rm ss}$, where $\rho_{\rm ss}$ is given by equation (3.11), and derive a simplified expression for $R(\rho_{\rm ss},\beta)$ keeping β arbitrary. We define

$$f(\mu_z,\beta)\equiv\frac{e^{\gamma t_p/2}-1}{\gamma}(2E+ig\mu_z)+(-1)^{k_p}\beta e^{\gamma t_p/2}\,, \tag{3.75}$$

so that

$$R(\rho_{\rm ss},\beta)=\int_0^{\Delta t}dt_m\int_0^{\Delta t}dt_{m-1}\cdots\int_0^{\Delta t}dt_1\Big(\prod_{p=1}^{m}e^{-\gamma t_p/2}[a+f(\mu_z,\beta)]\Big)\rho_{\rm ss}\Big(\cdots\Big)^\dagger\,. \tag{3.76}$$

We note that

$$[\mu_z,\rho_{\rm ss}]=0\quad\text{and}\quad(\mu_z)^2=\mathbb{1}\,. \tag{3.77}$$

Using the first of these properties and the expression for $\rho_{\rm ss}$ as given by Eq. (3.11), we have by direct calculation

$$[a+f(\mu_z,\beta)]\rho_{\rm ss}[a+f(\mu_z,\beta)]^\dagger = \Big(f(\mu_z,\beta)[f(\mu_z,\beta)]^\dagger+2{\rm Re}[f(\mu_z,\beta)]{\rm Re}(\alpha) \\ +2{\rm Im}[f(\mu_z,\beta)]{\rm Im}(\alpha)\mu_z+|\alpha|^2\Big)\rho_{\rm ss}\,, \tag{3.78}$$

where $\alpha=(2E+ig)/\gamma$. This equation will be useful in the following sections where we will analyze (3.76) in greater details.

3.4.2 Looking for the steady state

In this section we show that the state $\rho_{\rm ss}$ defined by (3.11) is a steady state of equation (3.30). First of all we note that it is *sufficient* to find a measurement such that for any measurement record $(k_1, \ldots, k_m; \Delta t)$ the conditional density matrix

$$\rho_{\rm c}(k_1, \ldots, k_m; \Delta t) = \rho_{\rm ss}\,. \tag{3.79}$$

In this case solution (3.34) of the unconditional master equation (3.30) becomes

$$\rho(\Delta t) = \sum_{m=0}^{\infty} \sum_{k_1,\ldots,k_m} p(k_1, \ldots, k_m; \Delta t)\rho_{\rm ss} = \rho_{\rm ss}\,, \tag{3.80}$$

for any Δt as required. It should be said that requirement (3.79) may not be necessary in the most general case. Intuitively, we would expect that being subjected to a measurement from a nontrivial class the system would depart from the steady state. In our case, however, we will find such β that equation (3.79) is satisfied.

Before we proceed with a rigorous analysis it is helpful to develop some intuition about the dependence of the experimental results on β. We have seen in section 3.3.1 that within the formalism of stochastic master equations the conditional and unconditional equations differ only by the term

$$\sqrt{\gamma\eta}\,\frac{dW}{dt}\left(e^{-i\phi}a\rho_{\rm c} + e^{i\phi}\rho_{\rm c}a^\dagger - {\rm tr}\left[\rho_{\rm c}(e^{-i\phi}a + e^{i\phi}a^\dagger)\right]\rho_{\rm c}\right). \tag{3.81}$$

This term is present only in the conditional master equation. For real β and at the initial condition $\rho_{\rm c}(0) = \rho_{\rm ss}$ this term is proportional to $\rho_{\rm ss}$. This means that if $\rho_{\rm ss}$ is indeed a steady state of equation (3.8), then, for real β, it is also a steady state of the corresponding conditional master equation (3.12). In this section we aim for a much more general consideration which does not rely on any built-in assumptions of the formalism of stochastic master equations. The above arguments, however, can be used to conjecture that if the system is initially prepared in the state $\rho_{\rm ss}$ then in the case of real β the conditional density matrix should coincide with $\rho_{\rm ss}$ at all times. If this conjecture is true then, as we explained at the beginning of this chapter, $\rho_{\rm ss}$ must be a steady state of the unconditional master equation (3.8).

Using the fact that $(\mu_z)^2 = \mathbb{1}$, we have that in this case $\beta = \beta_0$, where $\beta_0 = |\beta|$:

$$f(\mu_z)[f(\mu_z)]^\dagger = \frac{4E^2+g^2}{\gamma^2}[e^{\gamma t_p/2}-1]^2 + \beta_0^2 e^{\gamma t_p} + \frac{4E\beta_0}{\gamma}(-1)^{k_p}(e^{\gamma t_p} - e^{\gamma t_p/2}), \quad (3.82)$$

and

$$\mathrm{Re}[f(\mu_z)]\mathrm{Re}(\alpha) + \mathrm{Im}[f(\mu_z)]\mathrm{Im}(\alpha)\mu_z = \frac{4E^2+g^2}{\gamma^2}(e^{\gamma t_p/2}-1) + \frac{2E\beta_0}{\gamma}(-1)^{k_p}e^{\gamma t_p/2}. \quad (3.83)$$

Because $|\alpha|^2 = |2E+ig|^2/\gamma^2 = (4E^2+g^2)/\gamma^2$ we therefore have according to Eq. (3.78):

$$[a+f(\mu_z,\beta_0)]\rho_{\rm ss}[a+f(\mu_z,\beta_0)]^\dagger = e^{\gamma t_p}[\frac{4E^2+g^2}{\gamma^2} + (-1)^{k_p}\frac{4E\beta_0}{\gamma} + \beta_0^2]. \quad (3.84)$$

Substituting this into (3.76) we obtain:

$$R(\rho_{\rm ss},\beta_0) = (\Delta t)^m \prod_{p=1}^{m}\Big(\frac{4E^2+g^2}{\gamma^2} + (-1)^{k_p}\frac{4E\beta_0}{\gamma} + \beta_0^2\Big)\rho_{\rm ss}\,. \quad (3.85)$$

Therefore, according to Eq.(3.73),

$$\rho_{\rm c}(k_1,\ldots,k_m;\Delta t) \propto N(\Delta t)\rho_{\rm ss}N^\dagger(\Delta t), \quad \text{for any real } \beta. \quad (3.86)$$

As a final step linking this statement to equation (3.79) we prove the following lemma:[5]

Lemma 2

Smooth evolution leaves $\rho_{\rm ss}$ invariant in the following sense:

$$N(\Delta t)\rho_{\rm ss}N^\dagger(\Delta t) \propto \rho_{\rm ss}\,. \quad (3.87)$$

This is true for any value of the parameter β.

Proof

Because μ_z and $\rho_{\rm ss}$ commute, we can see from Eq. (3.11) that the smooth evolution leaves $\rho_{\rm ss}$ diagonal:

$$N(\Delta t)\rho_{\rm ss}[N(\Delta t)]^\dagger = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}. \quad (3.88)$$

[5]Eq. (3.79) follows from (3.86), the following lemma, and the normalization of the density matrix.

Using Theorem 6 we have

$$2e^{-2Z_1}\Lambda_1 = \left(e^{-\frac{\gamma\Delta t}{2}a^\dagger a}e^{Z_2^+ a^\dagger}e^{Z_3^- a}\right)|\alpha\rangle\langle\alpha|\left(\cdots\right)^\dagger, 2e^{-2Z_1}\Lambda_2 = \left(e^{-\frac{\gamma\Delta t}{2}a^\dagger a}e^{Z_2^- a^\dagger}e^{Z_3^+ a}\right)|\alpha^*\rangle\langle\alpha^*|\left(\cdots\right)^\dagger \tag{3.89}$$

where

$$Z_2^\pm \equiv \frac{2E \pm ig}{\gamma}(e^{\gamma\Delta t/2} - 1)\,, Z_3^\pm \equiv \frac{2E \pm ig}{\gamma}(e^{-\gamma\Delta t/2} - 1)\,. \tag{3.90}$$

In order to calculate Λ_1 we use Identity 4 (equation (3.108) in the Appendix) which gives

$$2e^{-2Z_1}\Lambda_1 = |e^{Z_3^-\alpha}|e^{|\alpha+Z_2^+|^2-|\alpha|^2}\left(e^{-\frac{\gamma\Delta t}{2}a^\dagger a}|\alpha+Z_2^+\rangle\langle\alpha+Z_2^+|e^{-\frac{\gamma\Delta t}{2}a^\dagger a}\right). \tag{3.91}$$

Now, using Identity 5 (equation (3.111) in the Appendix) we have

$$2e^{-2Z_1}\Lambda_1 = |e^{Z_3^-\alpha}|e^{|\alpha+Z_2^+|^2e^{-\gamma\Delta t}-|\alpha|^2}|(\alpha+Z_2^+)e^{-\frac{\gamma\Delta t}{2}}\rangle\langle(\alpha+Z_2^+)e^{-\frac{\gamma\Delta t}{2}}|\,. \tag{3.92}$$

Using the definition of Z_2^+ and the value of $\alpha = (2E+ig)/\gamma$ we see that

$$(\alpha+Z_2^+)e^{-\frac{\gamma\Delta t}{2}} = \alpha. \tag{3.93}$$

Therefore

$$2e^{-2Z_1}\Lambda_1 = |e^{Z_3^-\alpha}||\alpha\rangle\langle\alpha|\,. \tag{3.94}$$

Repeating the same arguments for Λ_2 we have from Eq. (3.89):

$$2e^{-2Z_1}\Lambda_2 = |e^{Z_3^+\alpha^*}||\alpha^*\rangle\langle\alpha^*|\,. \tag{3.95}$$

Because $|e^{Z_3^-\alpha}| = |e^{Z_3^+\alpha^*}|$ we can now see that

$$\begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \propto \begin{pmatrix} |\alpha\rangle\langle\alpha| & 0 \\ 0 & |\alpha^*\rangle\langle\alpha^*| \end{pmatrix} = 2\rho_{\rm ss}\,. \tag{3.96}$$

Together with Eq. (3.88) this completes the proof. □

3.4.3 Conditional evolution

In the previous section we identified a measurement which does not give any information about the system once it has reached the steady state $\rho_{\rm ss}$. Although this

fact was useful in confirming that $\rho_{\rm ss}$ is indeed a steady state of the system, our ultimate goal is to gain nonzero information about the system and quantify this information.

It can be seen, within the formalism of stochastic master equations, that the information gain about the system is maximized for purely imaginary β. We therefore concentrate on the case $\beta = i\beta_0$, where $\beta_0 = |\beta|$ and calculate the corresponding conditional density matrix. Direct calculations give

$$f(\mu_z)[f(\mu_z)]^\dagger = \frac{4E^2+g^2}{\gamma^2}(e^{\gamma t_p/2}-1)^2 + |\beta|^2 e^{\gamma t_p} + (-1)^{k_p}\frac{2g|\beta|}{\gamma}(e^{\gamma t_p} - e^{\gamma t_p/2}), \quad (3.97)$$

and

$$\mathrm{Re}[f(\mu_z)]\mathrm{Re}(\alpha) + \mathrm{Im}[f(\mu_z)]\mathrm{Im}(\alpha)\mu_z = \frac{4E^2+g^2}{\gamma^2}(e^{\gamma t_p/2}-1) + (-1)^{k_p}\frac{g|\beta|}{\gamma}e^{\gamma t_p/2}\mu_z\,. \quad (3.98)$$

Therefore

$$R(\rho_{\rm ss}, i\beta_0) = \int_0^{\Delta t} dt_m \int_0^{\Delta t} dt_{m-1} \cdots \int_0^{\Delta t} dt_1 \prod_{p=1}^{m}\Big(\frac{4E^2+g^2}{\gamma^2} + \beta_0^2 + (-1)^{k_p}\frac{2g\beta_0}{\gamma}[\mathbb{1} + (\mu_z - \mathbb{1})e^{-\gamma t_p/2}]\Big)\rho_{\rm ss}\,. \quad (3.99)$$

Performing the integration we obtain

$$R(\rho_{\rm ss}, i\beta_0) = \prod_{p=1}^{m}\Big[\Delta t\cdot\Big(\frac{4E^2+g^2}{\gamma^2} + \beta_0^2 + (-1)^{k_p}\frac{2g\beta_0}{\gamma}\Big) + (-1)^{k_p}4g\beta_0(\mu_z-\mathbb{1})\frac{1-e^{-\gamma\Delta t/2}}{\gamma^2}\Big]\rho_{\rm ss}\,. \quad (3.100)$$

Because $\rho_{\rm ss}$ and μ_z are both diagonal in the used basis[6] this essentially gives us the eigenvalues of the conditional density matrix. We omit the calculations as the eigenvalues follow immediately from the above equation and normalization. Knowing these eigenvalues we can calculate the von Neumann entropy of the conditional

[6] See comments to equation (3.9)

density matrix $H[\rho_c(k_1,\ldots,k_m;\Delta t)]$, from which the reduction of the system entropy is simply $-\Delta H = H(\rho_{\rm ss}) - H[\rho_c(k_1,\ldots,k_m\Delta t)]$, where $H(\rho_{\rm ss})$ is equal to 1 bit. We therefore succeeded to compute the reduction of the system entropy under the action of a continuous homodyne measurement.

3.5 Appendix 3.A

Here we prove some identities used throughout this chapter

Identity 1

$$e^{-xa^\dagger a}a^\dagger e^{xa^\dagger a} = a^\dagger e^{-x} \tag{3.101}$$

Proof

Define a function

$$g_1(x) \equiv e^{-xa^\dagger a}a^\dagger e^{xa^\dagger a}\,. \tag{3.102}$$

We have $g_1(0) = a^\dagger$, and moreover

$$\begin{aligned}\frac{dg_1(x)}{dx} &= e^{-xa^\dagger a}[a^\dagger, a^\dagger a]e^{xa^\dagger a}\\ &= e^{-xa^\dagger a}(-a^\dagger)e^{xa^\dagger a}\\ &= -g_1(x)\,.\end{aligned} \tag{3.103}$$

Performing integration with respect to x we therefore obtain

$$g_1(x) = a^\dagger e^{-x}\,. \tag{3.104}$$

as required. □

Identity 2

$$e^{-xa^\dagger a}ae^{xa^\dagger a} = ae^{x} \tag{3.105}$$

Proof Follows trivially from Eq. (3.101). □

Identity 3

$$e^{-ya^\dagger}ae^{ya^\dagger} = a + y \tag{3.106}$$

Proof This identity is a particular case of a more general result

$$e^{-ya^\dagger} f(a, a^\dagger) e^{ya^\dagger} = f(a + y, a^\dagger) \tag{3.107}$$

which holds for any function $f(a, a^\dagger)$ that can be expanded in a power series in a and $a^\dagger$ (consult Ref. [42] p.151 for the proofs). □

Identity 4

$$e^{\lambda a^\dagger} |\alpha\rangle\langle\alpha| e^{\lambda^* a} = e^{|\alpha+\lambda|^2 - |\alpha|^2} |\alpha + \lambda\rangle\langle\alpha + \lambda| \tag{3.108}$$

Proof

For any coherent state $|\alpha\rangle$ we have

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{(\alpha a^\dagger)^n}{n!} |0\rangle = e^{-\frac{1}{2}|\alpha|^2} e^{\alpha a^\dagger} |0\rangle \ . \tag{3.109}$$

Therefore

$$\begin{aligned} e^{\lambda a^\dagger} |\alpha\rangle &= e^{-\frac{1}{2}|\alpha|^2} e^{(\alpha+\lambda)a^\dagger} |0\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} e^{\frac{1}{2}|\alpha+\lambda|^2} \left(e^{-\frac{1}{2}|\alpha+\lambda|^2} e^{(\alpha+\lambda)a^\dagger} |0\rangle \right) \\ &= e^{\frac{1}{2}(|\alpha+\lambda|^2 - |\alpha|^2)} |\alpha + \lambda\rangle \ , \end{aligned} \tag{3.110}$$

from where Eq.(3.108) trivially follows. □

Identity 5

$$e^{-\lambda a^\dagger a} |\alpha\rangle\langle\alpha| e^{-\lambda a^\dagger a} = e^{|\alpha|^2 (e^{-2\lambda} - 1)} |\alpha e^{-\lambda}\rangle\langle\alpha e^{-\lambda}| \tag{3.111}$$

Proof

Using the well-known decomposition of a coherent state in the Fock basis

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \ , \tag{3.112}$$

we have

$$e^{-\lambda a^\dagger a} |\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-\lambda n} |n\rangle$$

$$
\begin{aligned}
&= e^{-\frac{1}{2}|\alpha|^2} e^{\frac{1}{2}|\alpha e^{-\lambda}|^2} \left(e^{-\frac{1}{2}|\alpha e^{-\lambda}|^2} \sum_{n=0}^{\infty} \frac{(\alpha e^{-\lambda})^n}{\sqrt{n!}} |n\rangle \right) \\
&= e^{\frac{1}{2}|\alpha|^2 (e^{-2\lambda}-1)} |\alpha e^{-\lambda}\rangle \ , \qquad (3.113)
\end{aligned}
$$

which is equivalent to (3.111). □

Conclusions and outlook

In this thesis we have introduced and developed the notion of dynamical complexity as a measure of complexity of system dynamics. Using several examples we demonstrated how this notion can be used to quantify the accuracy to which we can predict the system evolution. This led us to consider some applications in the theory of quantum chaos and more generally in the theory of open quantum systems subject to a continuous measurement. At the same time, the presented research has opened a number of important questions. In this final section we will discuss some of these questions as an outlook for future research.

- In Chapter 2 we have considered the quantum baker's map which is an example of a quantum system whose classical counterpart is chaotic. We have found that the symbolic dynamics of the classical baker's transformation is recovered in the classical limit. Because of the prominent role of symbolic dynamics in the theory of classical chaos this allowed us to conclude that unpredictability of the system behaviour can occur even when the system is described by the linear equations of quantum mechanics. We employed the rate of entropy production as a measure of dynamical complexity which was used to quantify the growing uncertainty about the system state relative to a set of measurements with fixed accuracy. Our results, however, were limited to the nearly classical regime where we could use the asymptotic relations obtained for the classical limit. A much more general theory of quantum symbolic dynamics is required to understand the origin of unpredictability in quantum systems to the degree provided by the theory of classical chaos.

This would bring us closer to the solution of another fundamental problem, namely the analytic demonstration of a hypersensitivity to perturbation in a purely quantum system.

- In chapters 1 and 3 we investigated applications of the concept of dynamical complexity to open quantum systems. Comparison of the information content in the measurement record to the reduction of the system entropy plays a significant role in judging the complexity of system dynamics. Unfortunately, this has been achieved only for some rather artificial examples: the problem is that it is difficult to estimate the algorithmic information contained in a real experimental record. In the realistic case of the cavity QED experiment we therefore restricted our attention to the rate of system entropy reduction caused by the measurements. Although this reduction is in itself a good measure of simplicity of the system dynamics, a future project is to consider the problem of estimating the average information content of the measurement record and complete our investigation of this important example.

- The results of Chapter 3, such as Kimble's formula, are general enough to account for limited detector efficiency $\eta < 1$, which is the fraction of detected photons. There is, however, a number of detector imperfections which cannot be accounted for by a single parameter such as detector efficiency. For instance, dark counts or any other type of noise which results in false events in the measurement record cannot be modeled as unregistered photodetections. Such imperfections, however, are common in present-day experiments and they do impose limitations on the rate at which information about the system state can be obtained. Both the limited detector efficiency and the detector imperfections mentioned above lead to a partial misinterpretation of the measurement record. A plan for future research is to introduce a concept of *detector accuracy*, which would accommodate all of the above detector imperfections, and generalize our theory for the case of limited de-

tector accuracy.

- Finally, our information theoretic analysis of the cavity QED experiment can be a starting point for an investigation of *quantum control*. A typical experimental setup in this case would include an additional feedback loop which can be used to manipulate and stabilize the system state. In the presence of such feedback the system behaviour would become highly predictable, or which is equivalent, have low dynamical complexity.

Bibliography

[1] Abramowitz, M., and Stegun, A. *Handbook of Mathematical Functions.* Dover Publications, New York, 1972.

[2] Alekseev, V. M., and Yakobson, M. V. Symbolic dynamics and hyperbolic dynamic systems. *Phys. Reports 75* (1981), 287–325.

[3] Alsing, P., and Carmichael, H. J. Spontaneous dressed-state polarization of a coupled atom and cavity mode. *Quantum Opt. 3*, 1 (February 1991), 13–32.

[4] Arnold, V. I., and Avez, A. *Ergodic Problems of Classical Mechanics.* Benjamin, New York, 1968.

[5] Badii, R., and Politi, A. *Complexity.* Cambridge University Press, Cambridge, 1997.

[6] Balazs, N. L., and Voros, A. The quantized baker's transformation. *Ann. Phys. 190* (1989), 1–31.

[7] Balian, R. *From Microphysics to Macrophysics. Volume 1.* Springer, Berlin, 1991.

[8] Berry, M. V. In *Chaos and Quantum Physics* (Amsterdam, 1991), M.-J. Giannoni, A. Voros, and J. Zinn-Justin, Eds., North-Holland.

[9] Berry, M. V. True quantum chaos? an instructive example. In *New Trends in Nuclear Collective Dynamics* (Berlin, 1992), Y. Abe, H. Horiuchi, and K. Matsuyanagi, Eds., Springer, pp. 183–186.

[10] Bowen, R. *Equilibrium states and the ergodic theory of Anosov Diffeomorphisms. Lecture notes in mathematics Vol 470.* Springer, Berlin, 1975.

[11] Breslin, J. K., Milburn, G. J., and Wiseman, H. M. Optimal quantum trajectories for continuous measurement. *Phys. Rev. Lett. 74*, 24 (1995), 4827–4830.

[12] Brudno, A. A. Entropy and the complexity of the trajectories of a dynamical system. *Trans. Mosc. Math. Soc. 44*, 2 (1983), 127–151. [Russian original: *Trudy. Moskov. Matem. Obshch. 44* (1982), 124–149].

[13] Brun, T. A., and Hartle, J. B. Entropy of classical histories. *LANL e-print* (1999). quant-ph/9808024.

[14] Brun, T. A., and Schack, R. Realizing the quantum baker's map on a NMR quantum computer. *Phys. Rev. A 59*, 4 (1999), 2649–2658.

[15] Caves, C. M., and Drummond, P. D. Quantum limits on bosonic communication rates. *Rev. Mod. Phys. 66*, 2 (April 1994), 481–537.

[16] Chaitin, G. J. *Algorithmic Information Theory.* Cambridge University Press, Cambridge, England, 1987.

[17] Chernoff, H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat. 23* (1952), 493–507.

[18] da Luz, M. G. E., and Ozorio de Almeida, A. M. Path integral for the quantum baker's map. *Nonlinearity 8* (1995), 43–64.

[19] Dittes, F. M., Doron, E., and Smilansky, U. Long-time behavior of the semiclassical baker's map. *Phys. Rev. E 49*, 2 (1994), R963–R966.

[20] Doherty, A. C., Habib, S., Jacobs, K., Mabuchi, H., and Tan, S. M. Quantum feedback control and classical control theory. *Phys. Rev. A 62* (2000), 012105.

[21] Doherty, A. C., and Hood, C. J. Private communication.

[22] Gardiner, C. W. *Handbook of Stochastic Methods*, 2nd ed. Springer, Berlin, 1985.

[23] Gell-Mann, M., and Hartle, J. B. Quantum mechanics in the light of quantum cosmology. In *Complexity, Entropy, and the Physics of Information* (Redwood City, CA, 1990), W. H. Zurek, Ed., Addison Wesley.

[24] Gell-Mann, M., and Hartle, J. B. Classical equations for quantum systems. *Phys. Rev. D. 47* (April 1993), 3345.

[25] Gell-Mann, M., and Hartle, J. B. Strong decoherence. In *Quantum Classical Correspondence: proceedings of the 4th Drexel symposium on quantum non-integrability* (1997), D. H. Feng and H. B. L, Eds., International Press, pp. 3 – 35.

[26] Gibbons, G. W. Typical states and density matrices. *J. Geom. Phys. 8* (1992), 147.

[27] Gradshteyn, I. S., and Ryzhik, I. M. *Table of Integrals, Series, and Products.* Academic Press, Boston, 1980.

[28] Griffiths, R. *J. Stat. Phys. 36* (1984), 219.

[29] Hannay, J. H., Keating, J. P., and Ozorio de Almeida, A. M. Optical realization of the baker's transformation. *Nonlinearity 7*, 5 (1994), 1327–1342.

[30] Hartle, J. B. Quantum past and the utility of history. *Physica Scripta T76* (1998), 67–77.

[31] Hood, C. J., Chapman, M. S., Lynn, T. W., and Kimble, H. J. Real-time cavity QED with single atoms. *Phys. Rev. Lett. 80* (1998), 4157–4160.

[32] Hood, C. J., Lynn, T. W., Doherty, A. C., Parkins, A. S., and Kimble, H. J. The atom-cavity microscope: single atoms bound in orbit by single photons. *Science 287* (2000), 1447–1453.

[33] Horn, R. A., and Johnson, C. R. *Matrix analysis.* Cambridge University Press, Cambridge, 1996.

[34] Kaplan, L., and Heller, E. J. Overcoming the wall in the semiclassical baker's map. *Phys. Rev. Lett. 76*, 9 (1996), 1453–1456.

[35] Koch, C., and Laurent, G. Complexity in the nervous system. *Science 284* (1999), 96–98.

[36] Kraus, K. *States, Effects, and Operations. Fundamental Notions of Quantum Theory.* Springer, Berlin, 1983. Lecture Notes in Physics Vol. 190.

[37] Lakshminarayan, A. On the quantum bakers map and its unusual traces. *Ann. Phys. 239*, 2 (1995), 272–295.

[38] Lesniewski, A., Rubin, R., and Salwen, N. Classical limits for quantum maps on the torus. *J. Math. Phys. 39*, 4 (1998), 1835–1847.

[39] Li, M., and Vitányi, P. *An Introduction to Kolmogorov Complexity and its Applications.* Springer, Berlin, 1997.

[40] Lichtenberg, A. J., and Lieberman, M. A. *Regular and Stochastic Motion.* Springer, New York, 1983.

[41] Liu, X. M., Hug, M., and Milburn, G. J. Sensitivity to measurement perturbation of single-atom dynamics in cavity QED. *Phys. Rev. A 62* (2000), 043801.

[42] Louisell, W. H. *Quantum Statistical Properties of Radiation.* Wiley, New York, 1990.

[43] Mabuchi, H., Turchette, Q. A., Chapman, M. S., and Kimble, H. J. Real-time detection of individual atoms falling through a high-finesse optical cavity. *Opt. Lett. 21* (1996), 1393–1395.

[44] Mabuchi, H., and Wiseman, H. M. Retroactive quantum jumps in a strongly coupled atom-field system. *Phys. Rev. Lett. 81*, 21 (1998), 4620–4623.

[45] Mattson, H. F. *Discrete Mathematics with Applications*. Wiley, New York, 1993.

[46] Omnes, R. *J. Stat. Phys. 53* (1988), 893, 933, 957.

[47] Ozorio de Almeida, A. M., and Saraceno, M. Periodic orbit theory for the quantized baker's map. *Ann. Phys. 210*, 1 (1991), 1–15.

[48] Pakonski, P., Ostruszka, A., and Życzkowski, K. Quantum baker map on the sphere. *Nonlinearity 12*, 2 (1999), 269–284.

[49] Peres, A. *Quantum Theory: Concepts and Methods.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[50] Pesin, Y. B. Characteristic Lyapunov exponents and smooth ergodic theory. *Russ. Math. Surveys 32*, 4 (1977), 55–114. [Russian original: *Usp. Mat. Nauk 32*, 4 (1977), 55–112].

[51] Plenio, M. B., and Knight, P. L. The quantum-jump approach to dissipative dynamics in quantum optics. *Rev. Mod. Phys. 70* (1998), 101–144.

[52] Rind, D. Complexity and climate. *Science 284* (1999), 105–107.

[53] Rubin, R., and Salwen, N. A canonical quantization of the baker's map. *Ann. Phys. 269*, 2 (1998), 159–181.

[54] Saraceno, M. Classical structures in the quantized baker's transformation. *Ann. Phys. 199* (1990), 37–60.

[55] Saraceno, M., and Voros, A. Towards a semiclassical theory of the quantum bakers map. *Physica D 79*, 2–4 (1994), 206–268.

[56] Schack, R. Algorithmic information and simplicity in statistical physics. *Int. J. Theor. Phys. 36*, 1 (January 1997), 209–226.

[57] Schack, R. Using a quantum computer to investigate quantum chaos. *Phys. Rev. A 57*, 3 (March 1998), 1634–1635.

[58] Schack, R., and Caves, C. M. Information and entropy in the baker's map. *Phys. Rev. Lett. 69* (December 1992), 3413–3416.

[59] Schack, R., and Caves, C. M. Hypersensitivity to perturbations in the quantum baker's map. *Phys. Rev. Lett. 71*, 4 (July 1993), 525–528.

[60] Schack, R., and Caves, C. M. Chaos for Liouville probability densities. *Phys. Rev. E 53* (1996), 3387.

[61] Schack, R., and Caves, C. M. Information-theoretic characterization of quantum chaos. *Phys. Rev. E 53* (1996), 3257.

[62] Schack, R., and Caves, C. M. Hypersensitivity to perturbation: An information-theoretical characterization of classical and quantum chaos. In *Quantum Communication, Computing, and Measurement* (New York, 1997), O. Hirota, A. S. Holevo, and C. M. Caves, Eds., Plenum Press, pp. 317–330.

[63] Schack, R., and Caves, C. M. Shifts on a finite qubit string: A class of quantum baker's maps. *Applicable Algebra in Engineering, Communication and Computing, AAECC 10* (2000), 305–310.

[64] Schack, R., D'Ariano, G. M., and Caves, C. M. Hypersensitivity to perturbation in the quantum kicked top. *Phys. Rev. E 50*, 2 (August 1994), 972–987.

[65] Soklakov, A. N., and Schack, R. Classical limit in terms of symbolic dynamics for the quantum baker's map. *Phys. Rev. E 61*, 5 (2000), 5108–5114.

[66] Steane, A. Quantum computing. *Rep. Prog. Phys. 61* (February 1998), 117 – 173.

[67] Wehrl, A. General properties of entropy. *Rev. Mod. Phys. 50*, 2 (April 1978), 221–260.

[68] Weng, G., Bhalla, U. S., and Iyengar, R. Complexity in biological signaling systems. *Science 284* (1999), 92–96.

[69] Werner, B. T. Complexity in natural landform patterns. *Science 284* (1999), 102–104.

[70] Weyl, H. *The Theory of Groups and Quantum Mechanics.* Dover, New York, 1950.

[71] Whitesides, G. M., and Ismagilov, R. F. Complexity in chemistry. *Science 284* (1999), 89–92.

[72] Wiseman, H. M., and Milburn, G. J. Interpretation of quantum jump and diffusion processes illustrated on the bloch sphere. *Phys. Rev. A 47*, 3 (March 1993), 1652–1666.

[73] Wiseman, H. M., and Milburn, G. J. Quantum theory of field quadrature measurements. *Phys. Rev. A 47* (January 1993), 642–662.

[74] Wootters, W. K. Random quantum states. *Found. Phys. 20* (1990), 1365–1378.

[75] Zurek, W. H., and Paz, J. P. Decoherence, chaos, and the second law. *Phys. Rev. Lett. 72*, 16 (April 1994), 2508–2511.