Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského

Quantum Channels With Memory

(Dizertačná práca)

Mgr. Tomáš Rybár

Štúdijný odbor:	4.1.2 Všeobecná fyzika a matematická fyzika
Externá vzdelávacia inštitúcia:	Centrum pre výskum kvantovej informácie Fyzikálny ústav Slovenskej akadémie vied
Školitelia:	Prof. RNDr. Vladimír Bužek, DrSc. (CVKI SAV) Doc. Mário Ziman, PhD. (CVKI SAV, ITP ETH)

Bratislava, Máj 2012

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Mgr. Tomáš Rybár
Študijný program: všeobecná fyzika a matematická fyzika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: 4.1.2. všeobecná fyzika a matematická fyzika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Kvantové procesy s pamäťou

Školiteľ: prof. RNDr. Vladimír Bužek, DrSc.
Katedra: FMFI.KTFDF - Katedra teoretickej fyziky a didaktiky fyziky

Spôsob sprístupnenia elektronickej verzie práce:
bez obmedzenia

Dátum zadania: 27.01.2011

Dátum schválenia: 03.05.2012

prof. RNDr. Melánia Babincová, CSc.
garant študijného programu

.....
študent

.....
školiteľ

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Mgr. Tomáš Rybár
Study programme:	General Physics and Mathematical Physics (Single degree study, Ph.D. III. deg., full time form)
Field of Study:	4.1.2. General Physics And Mathematical Physics
Type of Thesis:	Dissertation thesis
Language of Thesis:	English
Secondary language:	Slovak

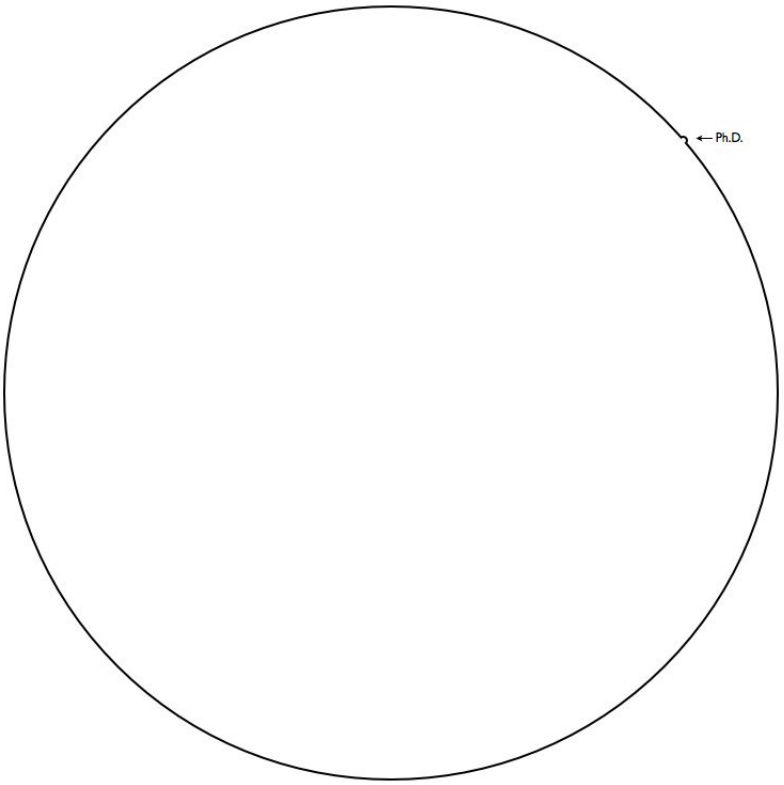
Title: Quantum Channels with Memory

Tutor:	prof. RNDr. Vladimír Bužek, DrSc.
Department:	FMFI.KTFDF - Department of Theoretical Physics and Didactics of Physics
Assigned:	27.01.2011
Approved:	03.05.2012

prof. RNDr. Melánia Babincová, CSc.
Guarantor of Study Programme

.....
Student

.....
Tutor


Ph.D.

Abstract

Quantum memory channels represent a very general, yet simple and comprehensible model for causal processes. As such they have attracted considerable research interest, mostly aimed on their transfer capabilities and structure properties. Most notably it was shown that memory channels can be implemented via physically naturally motivated collision models.

Quantum cellular automata are a compelling model for simulation of physical processes. In this work we establish a profound connection between quantum cellular automata and memory channels. We show that the locally computable invariant of an automaton is equal to the lowest dimension of ancillary system required for a memory channel implementing this automaton. Further we define the class of memory channels with finite depth and show that all causal quantum cellular automata can be implemented with such memory channels. For two dimensional memory system and qubit input we show that the depth of such memory channels is less than three.

We also define the concept of repeatable channels and show that only unital channels can be implemented repeatably with pure memory channels. In the special case of qubit channels we also show that every unital qubit channel has a repeatable implementation. We also briefly explore the possibilities of stroboscopical simulation of channels and show that all random unitary channels can be stroboscopically simulated. Particularly in qubit case, all indivisible qubit channels are also random unitary, hence for qubit all indivisible channels can be stroboscopically simulated.

Memory channels also naturally capture the framework of correlated experiments. We develop methods to gather and interpret data obtained in such setting and in detail examine the two qubit case. We also show that for control unitary interactions the measured data will never contradict a simple unitary evolution. Thus no memory effects can be spotted then.

Abstrakt

Kvantové kanály s pamäťou predstavujú veľmi všeobecný a predsa jednoduchý a zrozumiteľný model popisujúci kauzálne procesy. Pamäťové kanály sú predmetom mnohých štúdií, zväčša zameraných na ich transportačné vlastnosti a štruktúru. Bolo ukázané, že pamäťové kanály môžu byť vytvorené pomocou fyzikálne prirodzene motivovaných zrážkových modelov.

Kvantové bunkové automaty patria medzi presvedčivé modely pre simuláciu fyzikálnych procesov. V tejto práci predstavíme hlboké prepojenie medzi kvantovými bunkovými automatmi a pamäťovými kanálmi. Ukážeme, že lokálne spočítateľný invariant bunkového automatu je rovný najmenšej dimenzii pamäte potrebnej pre pamäťový kanál implementujúci tento automat. Ďalej definujeme triedu pamäťových kanálov s konečnou hĺbkou pamäte a ukážeme, že všetky kvantové bunkové automaty môžu byť vytvorené pomocou takýchto pamäťových kanálov. Pre dvojdimenzionálny systém pamäte aj vstupu ukážeme, že hĺbka takéhoto pamäťového kanála je menšia ako tri.

Taktiež definujeme pojem opakovateľných kanálov a ukážeme, že iba unitálne kanály môžu byť implementované opakovateľne pomocou unitárnych pamäťových kanálov. V špeciálnom prípade qubitových kanálov tiež ukážeme, že každý unitálny kanál má opakovateľnú implementáciu. V krátkosti preskúmame, aj možnosti stroboskopickej simulácie kanálov a dokážeme, že všetky náhodné unitárne kanály su stroboskopicky simulovateľné. Pre qubit sú všetky nedeliteľné kanály aj náhodné unitárne, a teda všetky nedeliteľné qubitové kanály sú stroboskopicky simulovateľné.

Pamäťové kanály prirodzene popisujú rámec korelovaných experimentov. Vyvinieme metódu na zhromažďovanie a interpretáciu dát získaných v takomto modeli a podrobne analyzujeme dvojqubitový prípad. Taktiež ukážeme, že pre kontrolované unitárne interakcie namerané dáta nikdy nebudú odporovať obyčajnému unitárnemu vývoju. V takom prípade nie je možné pozorovať žiadne pamäťové efekty.

Contents

Chapter 1

Introduction

Quantum information science is young and rapidly developing research field with many implications towards applications and fundamental questions. It connects physics with computation and information processing as well as raises questions about how and what can be read out from a physical state.

Any physical process can be viewed as some sort of communication between sender and receiver. Such information processing protocol is always affected by noise whose source is the environment. The noise usually hinders us from acquiring certain information about the process or message. It is thus crucial to identify the noise, and its consequences. When repeating some experiment, the noise affecting each try is taken independently. This is usually justified by assuming that the environment is not affected by the experiment. Such assumption allows us to approach the noise in a systematic way. However in many real world applications this assumption cannot be justified.

If the relaxation of environment is not perfect, the information from previous experiments may still roam around inside the environment. This greatly hardens the evaluation of the experiment, because repeated experiments cannot be treated independently anymore. The experiment explicitly depends on the past experiments but naturally is still independent of the upcoming experiments in future. Such scheme is called causal.

Quantum channels successfully describe independent experiments, whereas memory channels stand as model for the latter case when the experiments are affected by the history of the experiment. The important structure of causality underlies, and defines the concept of memory channels as stands proven in the paper of Werner and Kretschmann [37]. They show that every causal process is a collision model. That is a repetition of simpler processes, e.g. the single experiments, which share a common system, the environment. The environment serves then as the memory, it "remembers" the actions of previous processes. Memory channels are a more general concept for describing repeating processes and experiments as compared to quantum (memoryless) channels, and may better reflect the reality inside the laboratories. Most of the recent research was naturally aimed towards the information transfer capacity of certain classes of memory channels (see references in the introduction to Chapter 4). Our aim was to investigate the structural properties of quantum memory channels and analyze the possibilities of their estimation. In memoryless setting the information leaked into environment is irrevocably lost,

whereas in memory seeting, this leaked information can leak back in future experiemnts, thus is in principle available to the experimenter. On the other hand, the experimenter can obtain less information about the initial state of environment, since he has only one copy of it, as opposed to the memoryless case, where he has in principle infinitely many copies of the environment.

In this work we will also make a connection between memory channels and quantum cellular automata. Quantum cellular automata, originally proposed by Feynman in [22], offer a promising way towards quantum computations and simulations. This seemingly different concept, whose key feature is the finite speed of information propagation, is deeply connected with memory channels. We will show that every quantum cellular automaton can be consistently mapped to a subclass of memory channels. It turns out then that memory channels, collision models and quantum cellular automata are roughly the same objects motivated by different questions.

The work is organized as follows. Chapters 2 and 3 establish the language and basic notions of quantum information and estimation methods. Chapter 4 introduces the model of memory channels and quantum cellular automata. In Chapter 5 structural properties of memory channels are proposed and analyzed and connection between memory channels and quantum cellular automata is made. Some of this material was published in [58, 72] and some is in preparation [27, 71]. Chapter 6 develops estimating protocols for memory channels, and illustrates them on a simple two dimensional example. Results from this chapter are also in preparation to be published in near future in [73].

Chapter 2

Basics of quantum information

In this we will introduce the basic language of quantum information. All this can be found in standard textbooks on this subject [44, 49] and [28, 48, 5] for more in depth mathematical foundations. The notation and language will be mostly compatible with [28].

2.1 Notations and symbols

Definition of these symbols can be found in Appendix A.1.
$\mathcal{H}, \mathcal{H}_i, \mathcal{H}_{[a,b]}, \mathcal{M}$ - Hilbert space of some system,
$\mathcal{T}(\mathcal{H})$ - set of trace class operators on Hilbert space $\mathcal{H}$
$\mathcal{L}(\mathcal{H})$ - set of bounded operators on $\mathcal{H}$
T^*, T - some transformation in Schrödinger picture/ Heisenberg picture
$\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{M}$ - a C^*-algebra

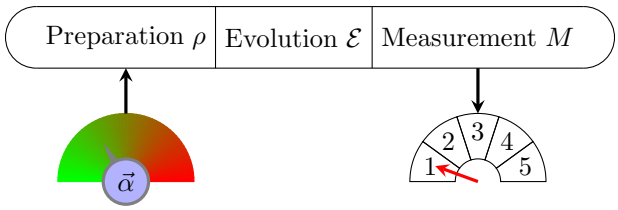


Figure 2.1: General experimenting scheme. Experimenter sets with some knob parameters $\vec{\alpha}$ of preparation procedure which in turn specifies a quantum mechanical state ρ. The state is then optionally evolved with evolution $\mathcal{E}$ and finally measured with measurement M with specific outcome, in this case 1. The borders between preparation evolution and measurement can be laid almost arbitrarily, for the purpose of description.

2.2 States and effects

A quantum mechanical experiment (see Figure 2.1) can be divided into three parts. Preparation, evolution and measurement. In preparation stage we feed our experimental apparatus with some set of (classical) parameters $\vec{\lambda}$ in order to produce some quantum mechanical state ρ. The same set of these parameters will always produce the same quantum mechanical state of the system $\rho(\vec{\lambda})$. In the sense that whenever we push the preparation button, a quantum state will be produced, independent of previous preparations and the ensemble of all states prepared with this fixed preparation is described by $\rho(\lambda)$. The evolution stage is to much extent arbitrary and can be made a part of preparation or measurement, depending on our needs. For now we let the evolution to be trivial, until the next section. At the end we decide which measurement to use and read out the outcome $O \in \Omega$, where Ω is the set of all possible outcomes. Since QM is a statistical theory, it predicts only probabilities $p(O|\rho)$. We have to repeat such experiment many times to acquire relevant statistics.

A quantum state ρ is then associated with certain preparation procedure and an effect E_O will then attach to this state the probability of outcome O when we conduct the measurement $E_O(\rho) = p(O|\rho)$. The situation when a preparation procedure ρ was used and outcome O was obtained is called an *event*.

Example 2.2.1 (Identity and zero effect)**.** The identity effect I assigns probability 1 to every state ρ, $I(\rho) = 1$. This describes a measurement with single outcome that will be always registered. Similarly zero effect assigns zero probability to an event that never happens.

Suppose that we fix the measurement procedure but we alternate between two different preparation procedures randomly. This should also define a possibly different but valid preparation procedure. Lets say that we choose the first preparation procedure with associated state ρ_1 with probability q and the other procedure with state ρ_2 with probability $1 - q$. The resulting preparation procedure will be associated with some state $q\rho_1 + (1 - q)\rho_2$, which is a mixture of the two states. The probabilities assigned with effect E should be consistent with probabilities measured with preparation procedures ρ_1 and ρ_2, thus

$$E(q\rho_1 + (1 - q)\rho_2) = qE(\rho_1) + (1 - q)E(\rho_2) \tag{2.1}$$

for any effect E and arbitrary $0 \leq q \leq 1$. To summarize:

- an effect is an affine mapping from the set of states to the interval $[0, 1]$
- the set of states is a convex set.

2.2.1 States

Now we can define the mathematical representation of the set of physical states. At this point actual quantum mechanics comes into play. The set of states, in the Hilbert space formulation of QM, is described by positive trace class operators of unit trace, also called density matrices

$$\mathcal{S}(\mathcal{H}) := \{\rho \in \mathcal{T}(\mathcal{H}) | \rho \geq O, \mathrm{Tr}(\rho) = 1\}, \tag{2.2}$$

we identify the set of states in quantum mechanics with $\mathcal{S}(\mathcal{H})$. This is a convex set. Every state from this set has a canonical decomposition

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{2.3}$$

where $|\psi_i\rangle$ are eigenvectors and $\sum_i p_i = 1$.

A convex set has the property that whenever $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$ then also their convex linear combination $\lambda\rho_1 + (1-\lambda)\rho_2$ for every $\lambda \in (0,1)$ is from this set. Every convex set is then fully described by its extremal elements. Those are the elements for which no nontrivial decomposition into a convex combination of different elements is possible.

The extremal points of $\mathcal{S}(\mathcal{H})$ are the 1-dimensional projections $|\psi\rangle\langle\psi|$ also called the *pure states*. Each $1-d$ projection can be identified with a normalized vector in $\mathcal{H}$, thus pure states can be also viewed as vectors on $\mathcal{H}$ with unit norm up to global phase. Every other state is then a *mixed state*. As can be seen from the canonical decomposition every state can be written as convex combination of 1-dimensional orthogonal projections. This is however not the only convex decomposition possible. In fact there are uncountably many other decompositions into convex combination of non-orthogonal projections or some non extremal elements of the state space which yield the same state.

Remark 2.2.2 (Classical state)**.** In classical statistical mechanics we describe the state space in a similar fashion but with an added requirement that all physically relevant states have to be diagonal in some specific basis. While this may seem as a subtle change, it is the essence of quantum versus classical debate.

In finite d-dimensional $\mathcal{H}$ the states of quantum system constitute of self-adjoint $d \times d$ square matrices with unit trace. Self-adjoint matrices form a vector space over real numbers. Given a basis in $\mathcal{H}$ as $\{|i\rangle\}$, there is a standard way how to construct basis in $\mathcal{L}_S(\mathcal{H}) \supset \mathcal{S}(\mathcal{H})$ called a *traceless operator basis*.

Example 2.2.3 (Self-adjoint operator basis)**.** Operators τ_{mn} form a basis in $\mathcal{L}_S(\mathcal{H})$, where

$$\tau_{mm} = \frac{1}{\sqrt{m(m+1)}}\left(\sum_{k=0}^{m-1} |k\rangle\langle k| - m|m\rangle\langle m|\right), \tag{2.4}$$

and for $m = 0$

$$\tau_{00} = \frac{1}{\sqrt{d}}\left(\sum_{k=0}^{d-1} |k\rangle\langle k|, \right. \tag{2.5}$$

and for $m < n$

$$\tau_{mn} = \frac{1}{\sqrt{2}}(|m\rangle|n\rangle + |n\rangle|m\rangle) \tag{2.6}$$

$$\tau_{nm} = \frac{1}{\sqrt{2}} i(-|m\rangle|n\rangle + |n\rangle|m\rangle). \tag{2.7}$$

One can check that $\tau_{ab} = \tau_{ab}^\dagger$, $\mathrm{Tr}(\tau_{ab}) = 0\ \forall (a,b) \neq (0,0)$ and that $\langle \tau_{ab} | \tau_{cd} \rangle_{\mathrm{HS}} = \mathrm{Tr}(\tau_{ab}\tau_{cd}) = \delta_{ac}\delta bd$. Every state ρ in $\mathcal{S}(\mathcal{H})$ can then be written as

$$\rho = \frac{1}{\sqrt{d}}(\tau_0 + \sum_{a=1}^{d^2-1} r_a \tau_a) \tag{2.8}$$

where $\tau_{00} =: \tau_0$ and τ_a are τ_{ij} in some fixed order.

Example 2.2.4 (Qubit)**.** The simplest nontrivial Hilbert space is of dimension 2. Let the basis of this space be denoted by normalized orthogonal vectors $|0\rangle, |1\rangle$. The traceless operator basis then is

$$\tau_{00} = \frac{1}{\sqrt{2}}\mathbb{I} \quad \tau_{01} = \frac{1}{\sqrt{2}}X \quad \tau_{10} = \frac{1}{\sqrt{2}}Y \quad \tau_{11} = \frac{1}{\sqrt{2}}Z \tag{2.9}$$

where

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.10}$$

are the Pauli matrices. A state ρ is then expressed as $\frac{1}{\sqrt{2}}(\tau_{00} + r_{01}\tau_{01} + r_{10}\tau_{10} + r_{11}\tau_{11})$ or more common as $\rho = \frac{1}{2}(\mathbb{I} + \vec{r}\vec{\sigma})$, where $\vec{\sigma} = (X, Y, Z)$ is the vector of Pauli matrices. In order to $\rho \geq O$ we find out that $\| r \| \leq 1$. The qubit state is defined by 3 parameters r_i such that $\sum_i r_i r_i \leq 1$. The set of all qubit states thus forms a ball better known as Bloch ball where each state is specified by a vector lying within this unit ball called a Bloch vector, see Figure 2.2. Pure states are the states for which $\| r \| = 1$ and lie on the surface of the ball. Two pure states are orthogonal if the Hilbert-Schmidt product of their density matrices is zero. Let $\rho_1 = \frac{1}{2}(\mathbb{I} + \vec{r}\vec{\sigma})$ and $\rho_2 = \frac{1}{2}(\mathbb{I} + \vec{s}\vec{\sigma})$.

$$\langle \rho_1 | \rho_2 \rangle_{\mathrm{HS}} = \mathrm{Tr}(\rho_1\rho_2) = \frac{1}{2}(1 + r_i s_j \mathrm{Tr}(\sigma_i\sigma_j)) = \frac{1}{2}(1 + \vec{r}\vec{s}), \tag{2.11}$$

therefore two pure states are orthogonal if their Bloch vectors are antipodal $\vec{r} = -\vec{s}$.

Definition 2.2.5 (von Neumann entropy)**.** The *von Neumann entropy* $S(\rho)$ of a state $\rho \in \mathcal{S}(\mathcal{H})$ is given by following

$$S(\rho) := -\mathrm{Tr}(\rho \log \rho) = -\sum_k p_k \log p_k, \tag{2.12}$$

where p_k are the nonzero eigenvalues of ρ.

It is the quantum analog of the classical Shannon entropy. The von Neumann entropy has following important properties:

- S is a concave function on the set of states, i.e. $S(q\rho_1 + (1-q)\rho_2) \geq qS(\rho_1) + (1-q)S(\rho_2)$;
- S is invariant under unitary conjugation $S(U\rho U^\dagger) = S(\rho)$;

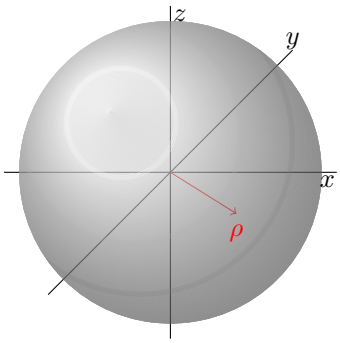


Figure 2.2: States of qubit form a Bloch ball. Each point of this ball is a valid physical state from a 2-dimensional Hilbert space.

- $S(\rho) = 0$ iff ρ is a pure state.

A unique state ρ_M which maximizes the von Neumann entropy is called a *maximally mixed state*. Such state is proportional to identity $\rho_M = \frac{1}{d}\mathbb{I}$ and the von Neumann entropy of such state is

$$S(\rho_M) = -\sum_{k=0}^{d-1} \frac{1}{d} \log \frac{1}{d} = \log d, \tag{2.13}$$

where d is the dimension of underlying Hilbert space.

Pure states have an additional structure called the superposition. Having two pure states $|\psi\rangle$ and $|\phi\rangle$ we can form another pure state $|\theta\rangle = \alpha|\psi\rangle + \beta|\phi\rangle$ which is called a superposition of $|\psi\rangle$ and $|\phi\rangle$. Such combination of pure states is not allowed in classical mechanics and gives rise to many if not all quantum phenomena.

2.2.2 Effects

When we defined the state space in previous subsection we also made a step toward defining the effects. From the beginning of this section we know that an effect is a linear mapping from $\mathcal{S}(\mathcal{H})$ to the interval $[0, 1]$. Thus an effect is a linear functional on $\mathcal{S}(\mathcal{H})$ which by linearity can be smeared over the whole $\mathcal{T}(\mathcal{H})$. The dual space of $\mathcal{T}(\mathcal{H})$ is $\mathcal{L}(\mathcal{H})$ and that is the native space of effects. So every effect can be associated with an operator $E \in \mathcal{L}(\mathcal{H})$ such that

$$E(\rho) = \mathrm{Tr}(E\rho), \tag{2.14}$$

where $\mathbb{I} \geq E \geq O$ in order to obtain a valid probability from interval $[0, 1]$. Let denote the set of effects by $\mathcal{E}(\mathcal{H})$.

Example 2.2.6 (Qubit effects)**.** Since effects are positive operators we can expand them in the traceless operator basis, thus every qubit effect can be written as

$$E = (\mathbb{I}\alpha + \vec{a}\vec{\sigma}). \tag{2.15}$$

Eigenvalues of this operator are $\lambda_\pm = \alpha \pm \parallel a \parallel$ therefore E is effect if and only if $\lambda_- \geq 0$ and $\lambda_+ \leq 1$. Notice that this also implies $\parallel a \parallel \leq 1$.

2.2.3 Observables

Each measurement device can have n possible different outcomes. Each outcome is represented by an effect E_j such that the probability of outcome j when $\rho \in \mathcal{S}(\mathcal{H})$ was prepared is $\mathrm{Tr}(E_j\rho)$. Formally let Ω be a nonempty set of outcomes. A *σ-algebra* on Ω is a collection $\mathcal{F}$ of subsets of Ω such that

- $\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$
- if $X \in \mathcal{F}$ then $\Omega \setminus X \in \mathcal{F}$
- if $X_1, X_2, \ldots \in \mathcal{F}$ then $\bigcup_i X_i \in \mathcal{F}$.

$\mathcal{F}$ represents all possible questions "Was the observed outcome in this subset of possible outcomes? ".

Definition 2.2.7 (POVM)**.** A *positive operator valued measure* is a mapping $A : \mathcal{F} \mapsto \mathcal{E}(\mathcal{H})$ such that

- $A(\emptyset) = O$
- $A(\Omega) = \mathbb{I}$
- $A(\bigcup_i X_i) = \sum_i A(X_i)$ for any set X_i of disjoint sets in $\mathcal{F}$.

A POVM is thus a prescription which assigns to every possible set of outcomes appropriate effect. However for all practical purposes in this work it is enough to think of POVM as a direct outcome - effect assignment. In this sense Ω is just the set of possible outcomes. Hence since $A(\Omega) = \mathbb{I}$ also $\sum_{i=1}^{n} E_i = \mathbb{I}$ for a measurement where every outcome is registered. An observable is then simply a POVM[1].

2.2.4 Bipartite systems

So far we have considered only one system in Hilbert space with arbitrary dimension. Let us have two independent quantum systems A and B in separate Hilbert spaces $\mathcal{A}$ and $\mathcal{B}$. The Hilbert space of composite system A+B is then obtained by tensor product $\mathcal{A} \otimes \mathcal{B}$. If the state of system A is $\rho_{\mathcal{A}}$ and of the system B, $\rho_{\mathcal{B}}$, the joint system is then again obtained by tensor product $\omega_{\mathcal{A}B} = \rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}$. We call such states factorized. Similarly for effects. If we have an effect $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ assigned to outcomes O_A, O_B the composite effect on $\mathcal{A} \otimes \mathcal{B}$ is $E_{\mathcal{A}} \otimes E_{\mathcal{B}}$. The probability assigned to $E_{\mathcal{A}} \otimes E_{\mathcal{B}}$ should respect the independent nature of the events $p(O_A, O_B|\rho_{\mathcal{A}}, \rho_{\mathcal{B}}) = p(O_A|\rho_{\mathcal{A}})p(O_B|\rho_{\mathcal{B}})$ and this is true:

$$\begin{aligned} p(O_A, O_B|\rho_{\mathcal{A}}, \rho_{\mathcal{B}}) &= \mathrm{Tr}[(E_{\mathcal{A}} \otimes E_{\mathcal{B}})(\rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}})] = \mathrm{Tr}(E_{\mathcal{A}}\rho_{\mathcal{A}})\mathrm{Tr}(E_{\mathcal{B}}\rho_{\mathcal{B}}) \\ &= p(O_A|\rho_{\mathcal{A}})p(O_B|\rho_{\mathcal{B}}). \end{aligned} \tag{2.16}$$

[1] Observables as self-adjoint operators are then POVMs with all effects being orthogonal projections. They are sometimes referred to as projection valued measures, PVMs.

However with the structure of the tensor product we have also introduced a state space which is by far larger than a state space of factorized density matrices. Not every density matrix in $\mathcal{A} \otimes \mathcal{B}$ is of the aforementioned form. States which cannot by written in this form describe correlated systems. If the state can be written as convex combination of factorized states, then it is called *separable*. Finally states which are not separable or factorized are called *entangled* and describe systems with quantum correlations Pure states can be only entangled or factorized.

We have thus a way how to describe joint systems, however through this journey we have lost the knowledge of how to describe single systems which are part of a larger system. If a state is not factorized the legitimate question is then what are the *local* states of respective subsystems?

Given that we have some $\omega_{\mathcal{A}B} \in \mathcal{S}(\mathcal{A} \otimes \mathcal{B})$ which is not factorized we denote $\omega_{\mathcal{A}}$ and $\omega_{\mathcal{B}}$ the local states of respective subsystems even if we do not know now what they are. We require that any measurement outcome probability obtained only on one of the subsystems has to be recovered by the density matrix of the subsystem:

$$\mathrm{Tr}[\omega_{\mathcal{A}B}(E_{\mathcal{A}} \otimes \mathbb{I})] = \mathrm{Tr}(\omega_{\mathcal{A}} E_{\mathcal{A}}). \tag{2.17}$$

The state $\omega_{\mathcal{A}}$ is called a partial trace of $\omega_{\mathcal{A}B}$ over system B and is denoted by

$$\omega_A = \mathrm{Tr}_{\mathcal{B}} \omega_{AB} = \sum_i \langle i|_{\mathcal{B}} \omega_{\mathcal{A}B} |i\rangle_{\mathcal{B}} \tag{2.18}$$

with some basis in $\mathcal{B}$, $\{|i\rangle_{\mathcal{B}}\}$.

Let the basis of $\mathcal{A} \otimes \mathcal{B}$ be $\{|i\rangle_{\mathcal{A}} \otimes |j\rangle_{\mathcal{B}}\}$ where $\{|i\rangle_{\mathcal{A}/\mathcal{B}}\}$, are bases of respective subspaces and let the state $\omega_{\mathcal{A}B}$ have an expansion in this basis

$$\omega_{\mathcal{A}B} = \sum_{klmn} \omega_{kl,mn} |k\rangle_{\mathcal{A}} \langle l| \otimes |m\rangle_{\mathcal{B}} \langle n|, \tag{2.19}$$

then the density matrix $\omega_{\mathcal{A}}$ is

$$\begin{aligned} \sum_i \langle i|_{\mathcal{B}} \omega_{\mathcal{A}B} |i\rangle_{\mathcal{B}} &= \sum_{iklmn} \omega_{kl,mn} |k\rangle_{\mathcal{A}} \langle l| \otimes \langle i|m\rangle_{\mathcal{B}} \langle n|i\rangle \\ &= \sum_{ikl} \omega_{kl,ii} |k\rangle_{\mathcal{A}} \langle l|. \end{aligned} \tag{2.20}$$

Example 2.2.8 (Schmidt decomposition)**.** Let us have a pure state $|\psi\rangle$ in $\mathcal{S}(\mathcal{H} \otimes \mathcal{H})$ with a basis in (both) $\mathcal{H}$: $\{|i\rangle\}$, such that $|\psi\rangle = \sum_j \sqrt{p}_j |jj\rangle$. Let denote the partial trace over the first and second subsystem Tr_1 and Tr_2 respectively. The density matrix of $|\psi\rangle$ is

$$|\psi\rangle\langle\psi| = \sum_{ij} \sqrt{p}_{ij} |ii\rangle\langle jj|, \tag{2.21}$$

and the local density matrix of both subsystems is

$$\mathrm{Tr}_1(|\psi\rangle\langle\psi|) = \sum_i p_i |i\rangle\langle i| = \mathrm{Tr}_2(|\psi\rangle\langle\psi|). \tag{2.22}$$

This is a so called *Schmidt decomposition*. Every pure state in $\mathcal{A} \otimes \mathcal{B}$ (where $\mathcal{A}$ doesn't have to be equal to $\mathcal{B}$) can be written for a suitably chosen basis in $\mathcal{A}$ and $\mathcal{B}$ $\{|i\rangle_{\mathcal{A}/\mathcal{B}}\}$ as(2.21) and the density matrices of subsystems are diagonal in these bases. Moreover for any density matrix $\rho \in \mathcal{S}(\mathcal{H})$ there exists such $|\psi\rangle \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H})$ of the aforementioned form.

The entropy of a pure state is zero. However if this state is bipartite, the entropy of its separate parts needs not be zero if the composite state is not factorized. This reflects that the information inside the local states is not enough to describe the whole information of the bipartite state. The reason is that in the description of local states we cannot include correlations which might be present. This property of von Neumann entropy is called *subadditivity*. Mathematically

$$S(\omega_{\mathcal{A}B}) \leq S(\rho_{\mathcal{A}}) + S(\rho_{\mathcal{B}}), \tag{2.23}$$

end the equality arises only when systems A and B are not correlated.

2.3 Evolution

In quantum mechanics we have learned that Schrödinger equation governs the time evolution of a quantum state

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle, \tag{2.24}$$

where H is a self-adjoint operator, a hamiltonian. If this hamiltonian is time independent we can give a formal solution to this

$$|\psi(t)\rangle = e^{iHt/\hbar}|\psi\rangle =: U(t)|\psi\rangle, \tag{2.25}$$

where $U(t)$ is some unitary operator for every t. For a general $\rho \in \mathcal{S}(\mathcal{H})$ this yields unitary conjugation $\rho(t) = U(t)\rho U^\dagger(t)$. This is the Schrödinger picture. Alternatively we can look on the evolution of observables: $A(t) = U^\dagger(t)AU(t)$ in Heisenberg picture. Both pictures are equivalent, however some times it is more insightful to use Heisenberg picture, mostly when infinite dimensional systems are considered and sometimes the Schrödinger picture is more intuitive. Another asymmetry which favors the Heisenberg picture is that the states in big Hilbert spaces are hard to describe in terms of localization. This is in general also true for observables, however physically for us it usually makes sense to think about measurements and measurement apparatuses which are precisely localized, and physically feasible to construct. It seems that nonlocal states are far more easily produced than nonlocal measurements. In this work we will use both pictures at our advantage. A map in Heisenberg picture T will have a Schrödinger picture equivalent T^* with an asterisk connected via identity

$$\mathrm{Tr}[T^*(\rho)B] = \mathrm{Tr}[\rho T(B)], \tag{2.26}$$

for any observable B and any state ρ.

If we have our state space defined as $\mathcal{S}(\mathcal{H})$ then possibly every map $T^* : \mathcal{S}(\mathcal{H}) \mapsto \mathcal{S}(\mathcal{H})$ could be a valid evolution. Such evolution should have thus following properties

- linearity; $T^*(\rho_1 + \lambda\rho_2) = T^*(\rho_1) + \lambda T^*(\rho_2)$
- trace preserving; $\mathrm{Tr}[T^*(\rho)] = \mathrm{Tr}\rho$
- positive; $T^*(\rho) \geq O$.

This turns out to be not enough. Lets say we have a map T^* which is positive. This guarantees that $T^*(\rho) \geq O$ for every positive $\rho \in \mathcal{H}$. This does not guarantee us that $\mathbb{I}_\mathcal{K} \otimes T^*(\omega) \geq O$ for every $\omega \in \mathcal{K} \otimes \mathcal{H}$ where $\mathcal{K}$ is a Hilbert space with arbitrary dimension.

Maps that satisfy this condition $\mathbb{I}_\mathcal{K} \otimes T^*(\omega)$ for arbitrary $\mathcal{K}$ are called *completely positive*. Notice that due to linearity the action can be shifted from the state space $\mathcal{S}(\mathcal{H})$ to the trace class operator $\mathcal{T}(\mathcal{H})$.

Definition 2.3.1 (Channel)**.** A linear trace preserving mapping in Schrödinger picture $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ which is at the same time completely positive is called a channel.

In Heisenberg picture the trace preserving condition translates to unitality on observables. We call a mapping T *unital* if it preserves the identity operator: $T(\mathbb{I}) = \mathbb{I}$.

Definition 2.3.2 (Channel)**.** A linear unital mapping in Heisenberg picture $T : \mathcal{L}(\mathcal{B}) \mapsto \mathcal{L}(\mathcal{A})$ which is at the same time completely positive is called a channel.

In previous lines we have adopted convention that states entering some evolution (inputs) live in Hilbert space $\mathcal{A}$ and states which are produced after the evolution (outputs) live in Hilbert space $\mathcal{B}$. The Hilbert spaces $\mathcal{A}$ and $\mathcal{B}$ can be different, but unless explicitly stated we will consider them as isomorphic. However we will still maintain the distinction in naming the input Hilbert space as $\mathcal{A}$ and the output space $\mathcal{B}$ for better readability. Note that Heisenberg picture describes the evolution *against* the flow of time from observables on outputs in $\mathcal{B}$ to observables on inputs in $\mathcal{A}$.

To show that not every positive map is also completely positive see next example.

Example 2.3.3 (Positive but not completely positive)**.** Let us have a linear qubit map $E : \mathcal{S}(\mathcal{H}) \mapsto \mathcal{S}(\mathcal{H})$ defined on (not normed) traceless operator basis:

$$E(\mathbb{I}) = \mathbb{I} \quad E(X) = X \quad E(Y) = Y \quad E(Z) = O. \tag{2.27}$$

This map projects the states in Bloch ball onto its xy-plane and thus is positive. Lets take a bipartite pure state $|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ whose density matrix is

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \frac{1}{4}(\mathbb{I} + X \otimes X + Y \otimes Y - Z \otimes Z). \tag{2.28}$$

Then

$$\mathbb{I} \otimes E(|\psi\rangle\langle\psi|) = \frac{1}{4}(\mathbb{I} + X \otimes X + Y \otimes Y) = \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}. \tag{2.29}$$

However this is not positive, hence projection to plane is not completely positive. As it will turn out later a projection to a line is completely positive.

We do not know whether all channels are actually physical. Schrödinger equation gives us only unitary evolution on single system, however the structure of state space does not put such hard constraints on evolution. Stinespring's key result [64] tells us that every completely positive linear map T^* can be thought of as a unitary evolution on a larger Hilbert space.

Theorem 2.3.4 (Stinespring's dilation theorem)**.** *Let $\mathfrak{B}$ be a unital C^*-algebra and let $T : \mathfrak{B} \mapsto \mathcal{L}(\mathcal{A})$ be a completely positive unital map. Then there exists a Hilbert space $\mathcal{K}$, a unital $\dagger$-homomorphism[2] $\pi : \mathfrak{B} \mapsto \mathcal{L}(\mathcal{K})$ and an isometry $V : \mathcal{A} \mapsto \mathcal{K}$ such that*

$$T(b) = V^\dagger \pi(b) V. \tag{2.30}$$

for all $b \in \mathfrak{B}$. The triple $(\pi, V, \mathcal{K})$ is called Stinespring representation of channel T. If the closed linear span of $\pi(\mathfrak{B})V\mathcal{A}$ equals $\mathcal{K}$ then such representation is called minimal.

The Stinespring's dilation theorem is very powerful and does not even require separability of Hilbert space.

For finite d-dimensional $\mathcal{B}$ where we identify $\mathcal{L}(\mathcal{B})$ with $\mathfrak{B}$ we get that $\mathcal{K} = \mathcal{M} \otimes \mathcal{B}$, for some finite dimensional Hilbert space $\mathcal{M}$, isometry V can be written as $V = U|0\rangle$ where U is unitary[3], $|0\rangle$ is some pure state in $\mathcal{M}$ and the mapping $\pi(b)$ will become $\pi(b) = \mathbb{I}_\mathcal{M} \otimes b$ so that

$$\begin{aligned} \mathrm{Tr}[\rho T(b)] &= \mathrm{Tr}[(|0\rangle\langle 0| \otimes \rho) U^\dagger (\mathbb{I}_\mathcal{M} \otimes b) U] \\ &= \mathrm{Tr}[U(|0\rangle\langle 0| \otimes \rho) U^\dagger (\mathbb{I}_\mathcal{M} \otimes b)] = \mathrm{Tr}[T^*(\rho) b]. \end{aligned} \tag{2.31}$$

In Schrödinger picture this has a nice interpretation. Any channel $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ can be realized as a composition of three maps:

1. take input state ρ and attach to it a pure state of some environment $\mathcal{M}$: $|0\rangle$, to get $|0\rangle\langle 0| \otimes \rho$;
2. let the composite system evolve with unitary interaction U, $U(|0\rangle\langle 0| \otimes \rho)U^\dagger$;
3. trace out the environmental degrees of freedom to obtain the output, $T^*(\rho) = \mathrm{Tr}_\mathcal{M}[U(|0\rangle\langle 0| \otimes \rho)U^\dagger]$.

The last step can can be seen from

$$\begin{aligned} &\mathrm{Tr}[T^*(\rho) b] = \mathrm{Tr}[U(|0\rangle\langle 0| \otimes \rho) U^\dagger (\mathbb{I}_\mathcal{M} \otimes b)] \\ &= \sum_{i\mu} \langle \mu i | U(|0\rangle\langle 0| \otimes \rho) U^\dagger (\mathbb{I}_\mathcal{M} \otimes b) | \mu i \rangle \\ &= \sum_{i\mu} \langle \mu i | U(|0\rangle\langle 0| \otimes \rho) U^\dagger (\mathbb{I}_\mathcal{M} \otimes b) | \mu i \rangle \\ &= \mathrm{Tr}[\mathrm{Tr}_\mathcal{M}\big(U(|0\rangle\langle 0| \otimes \rho) U^\dagger\big) b], \end{aligned} \tag{2.32}$$

for some basis $\{|\mu\rangle\}$ in $\mathcal{M}$ and $\{|i\rangle\}$ in $\mathcal{B}$, hence

$$T^*(\rho) = \mathrm{Tr}_\mathcal{M}[U(|0\rangle\langle 0| \otimes \rho) U^\dagger]. \tag{2.33}$$

[2] $\pi(b^\dagger) = \pi(b)^\dagger$

[3] In case $\mathcal{A}$ and $\mathcal{B}$ were not isomorphic U would remain an isometry.

Remark 2.3.5 (Kraus representation)**.** The unitary $U : \mathcal{M} \otimes \mathcal{A} \mapsto \mathcal{M} \otimes \mathcal{B}$ can be written as

$$U = |i\rangle\langle j| \otimes A_{ij}, \tag{2.34}$$

where A_{ij} are some linear operators $A_{ij} : \mathcal{A} \mapsto \mathcal{B}$. Then

$$\begin{aligned}
&\mathrm{Tr}_{\mathcal{M}}[U(|0\rangle\langle 0| \otimes \rho)U^\dagger] = \sum_i \langle i|(U(|0\rangle\langle 0| \otimes \rho)U^\dagger)|i\rangle \\
&= \sum_{iklmn} \langle i|k\rangle\langle l|0\rangle\langle 0|m\rangle\langle n|i\rangle A_{kl}\rho A_{nm}^\dagger \\
&= \sum_i A_{i0}\rho A_{i0}^\dagger = T^*(\rho).
\end{aligned} \tag{2.35}$$

This is the so called *Kraus representation* of a channel $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ with *Kraus operators* $A_k := A_{k0}$. This way we can completely remove the explicit environment from the description of channel and instead use only the set of Kraus operators $\{A_k\}$ which has to satisfy the trace preserving condition of channel

$$\mathrm{Tr}[\rho] = \mathrm{Tr}[T^*(\rho)] = \mathrm{Tr}[\sum_k A_k\rho A_k^\dagger] \Rightarrow \sum_k A_k^\dagger A_k = \mathbb{I}_{\mathcal{A}}. \tag{2.36}$$

In Heisenberg picture we get $T : \mathcal{L}(\mathcal{B}) \mapsto \mathcal{L}(\mathcal{A})$ through the duality relation $\sum_k \mathrm{Tr}[A_k\rho A_k^\dagger b] = \sum_k \mathrm{Tr}[\rho A_k^\dagger b A_k]$:

$$T(b) = \sum_k A_k^\dagger b A_k, \tag{2.37}$$

where we see that the trace preserving condition transforms to unitality

$$\sum_k A_k^\dagger \mathbb{I}_{\mathcal{B}} A_k = \mathbb{I}_{\mathcal{A}}. \tag{2.38}$$

The Kraus operators are not unique. As can be seen from the Stinespring's representation the unitary U can be replaced for $(w \otimes \mathbb{I})U$ with an arbitrary unitary $w : \mathcal{M} \mapsto \mathcal{M}$,

$$\begin{aligned}
&\mathrm{Tr}[(w \otimes \mathbb{I})U(|0\rangle\langle 0| \otimes \rho)U^\dagger(w^\dagger \otimes \mathbb{I})(\mathbb{I}_{\mathcal{M}} \otimes b)] \\
&= \mathrm{Tr}[(w^\dagger \otimes \mathbb{I})(w \otimes \mathbb{I})U(|0\rangle\langle 0| \otimes \rho)U^\dagger(\mathbb{I}_{\mathcal{M}} \otimes b)] \\
&= \mathrm{Tr}[U(|0\rangle\langle 0| \otimes \rho)U^\dagger(\mathbb{I}_{\mathcal{M}} \otimes b)].
\end{aligned} \tag{2.39}$$

This freedom translates to freedom on Kraus operators:

$$\hat{A}_k = \sum_i w_{ki} A_i. \tag{2.40}$$

The new Kraus operators $\hat{A}_k$ form the same channel as operators A_k. It was shown by Kraus [36, 35] that any channel $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ can have a Kraus representation with number of Kraus operators n such that $n \leq d^2$ where $d = \max\big(\dim(\mathcal{A}), \dim(\mathcal{B})\big)$.

Remark 2.3.6 (Uniqueness of minimal Stinespring dilation)**.** Steinspring's representation is also highly non-unique. However it can be shown that minimal Stinespring's representation is unique up to isometry in the following sense. Let us have two different Stinespring's representations of the same channel T, $(\pi, V, \mathcal{K})$ and $(\pi_1, V_1, \mathcal{K})_1$,

$$T(b) = V^\dagger \pi(b) V = V_1^\dagger \pi_1(b) V_1. \tag{2.41}$$

Assume that the first representation is minimal. Then we can conclude that $\dim \mathcal{K} \leq \dim \mathcal{K}_1$ and there exist a well defined isometry $W : \mathcal{K} \mapsto \mathcal{K}_1$ with prescription

$$W(\pi(b)V\psi) := \pi_1(b)V_1\psi, \tag{2.42}$$

for all $b \in \mathfrak{B}$ and $\psi \in \mathcal{A}$. By setting $b = \mathbb{I}$ we get

$$WV = V_1, \tag{2.43}$$

and the intertwining relation

$$W\pi = \pi_1 W \tag{2.44}$$

as bonus.

Definition 2.3.7 (Choi-Jamiolkowski operator)**.** Let us have a channel T^* on d-dimensional Hilbert space $\mathcal{H}$ with some basis $\{|i\rangle\}$. We call the operator $\chi(T^*) = \sum_{i,j=0}^{d-1} |i\rangle\langle j| \otimes T^*(|i\rangle\langle j|)$ the *Choi-Jamiolkowski operator* of the channel T^*.

Note that

$$\mathrm{Tr}_2[\chi(T^*)] = \sum_{i,j=0}^{d-1} |i\rangle\langle j| \mathrm{Tr}[T^*(|i\rangle\langle j|)], \tag{2.45}$$

due to the trace preserving quality of T^* we get

$$\begin{aligned} \sum_{i,j=0}^{d-1} |i\rangle\langle j| \mathrm{Tr}[T^*(|i\rangle\langle j|)] &= \sum_{i,j=0}^{d-1} |i\rangle\langle j| \delta_{ij} \\ &= \sum_{i=0}^{d-1} |i\rangle\langle i| = \mathbb{I}, \end{aligned} \tag{2.46}$$

where by Tr_2 we denote the partial trace over the second subsystem. Since T^* is completely positive, the Choi-Jamiolkowski operator is positive[4]. The virtue of Choi-Jamiolkowski operator is that it is unique, what makes it a good representation for optimization in the space of quantum channels. Moreover, the relation is isomorphic, meaning that every positive operator χ in $\mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ whose partial trace over one subsystem is $\mathrm{Tr}_2\chi = \mathbb{I}$ is a Choi-Jamiolkowski operator of some channel.

[4]The operator $\sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|$ is positive, and remains positive whenever completely positive mapping occurs on its subparts.

A channel is a linear map acting on a real vector space of self-adjoint operators. Lets denote by $\{\tau_j\}$ the traceless operator basis. We can view a channel as an affine mapping on the generalized Bloch vector $\vec{r}$ from equation (2.8),

$$\vec{r} \mapsto \vec{a} + \hat{A}\vec{r}, \tag{2.47}$$

with some vector $\vec{a}$ and operator A. It can be further encapsulated in one operator

$$\begin{pmatrix} 1 \\ \vec{r} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ \vec{a} & \hat{A} \end{pmatrix} \begin{pmatrix} 1 \\ \vec{r} \end{pmatrix}. \tag{2.48}$$

Let

$$A = \begin{pmatrix} 1 & 0 \\ \vec{a} & \hat{A} \end{pmatrix}, \tag{2.49}$$

then

$$A_{ij} = \mathrm{Tr}[\tau_i T^*(\tau_j)]. \tag{2.50}$$

For unitary channels A is from special orthonormal group, $A^T A = \mathbb{I}, \det(A) = 1$ and $\vec{a} = 0$ because of unitality.

Example 2.3.8 (Contraction to a line)**.** Let us have a qubit channel $E^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ defined on the operator basis

$$E^*(\mathbb{I}) = \mathbb{I} \quad E^*(X) = X \quad E^*(Y) = O \quad E^*(Z) = O. \tag{2.51}$$

The vector representation is by far the easiest to acquire, let $\tau = 1/\sqrt{2}(\mathbb{I}, X, Y, Z)$, then

$$\begin{gathered} A_{ij} = \mathrm{Tr}[\tau_i E^*(\tau_j)] \Rightarrow \\ A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{gathered} \tag{2.52}$$

The Choi-Jamiolkowski operator is also quite straightforward. Knowing that

$$\sum_{i,j=0}^{1} |i\rangle\langle j| \otimes |i\rangle\langle j| = \frac{1}{2}(\mathbb{I} \otimes \mathbb{I} + X \otimes X - Y \otimes Y + Z \otimes Z), \tag{2.53}$$

we get

$$\begin{aligned} \chi(E^*) &= \frac{1}{2}(\mathbb{I} \otimes \mathbb{I} + X \otimes X) \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.54}$$

To get the Kraus representation is a bit more involving. Expanding the Kraus operators in operator basis

$$A_k = \sum_a a_{ka}\tau_a, \tag{2.55}$$

we get

$$\begin{aligned} E^*(\rho) &= \sum_k \sum_{a,b=0}^{3} a_{ka}\overline{a_{kb}}\tau_a\rho\tau_b \\ &= \sum_{a,b=0}^{3} x_{ab}\tau_a\rho\tau_b. \end{aligned} \tag{2.56}$$

where $x_{ab} = \sum_k a_{ka}\overline{a_{kb}}$. Then observe

$$\begin{aligned} \chi(E^*) &= \sum_k \sum_{a,b=0}^{3} a_{ka}\overline{a_{kb}} \sum_{i,j=0}^{1} |i\rangle\langle j| \otimes (\tau_a|i\rangle\langle j|\tau_b) \\ &= \sum_{a,b=0}^{3} x_{ab} \sum_{i,j=0}^{1} |i\rangle\langle j| \otimes (\tau_a|i\rangle\langle j|\tau_b) \\ &= \sum_{a,b=0}^{3} x_{ab}|\psi_a\rangle\langle\psi_b|, \end{aligned} \tag{2.57}$$

where $|\psi_a\rangle = \sum_{i=0}^{1} |i\rangle \otimes \tau_a|i\rangle$ are orthonormal vectors. The matrix x is the Choi-Jamiolkowski operator of channel E^* in basis $\{|\psi_a\rangle\}$, therefore it is positive and diagonalizable, thus we can write

$$x_{ab} = \sum_{k=0}^{3} U_{ak}\lambda_k\overline{U_{bk}} \tag{2.58}$$

where λ_k are eigenvalues of $\chi(E^*)$, and

$$E^*(\rho) = \sum_{a,b=0}^{3} x_{ab}\tau_a\rho\tau_b = \sum_{k,a,b=0}^{3} \lambda_k U_{ak}\tau_a\rho\overline{U_{bk}}\tau_b. \tag{2.59}$$

We then get

$$A_k = \sum_{a=0}^{3} \sqrt{\lambda_k}U_{ak}\tau_a. \tag{2.60}$$

For channel E^* we obtain x_{ab} from the Choi-Jamiolkowski operator, $x_{ab} = \langle\psi_a|\chi(E^*)\psi_b\rangle$:

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.61}$$

Since this is already diagonal with only one nonzero eigenvalue 1 with multiplicity 2, we know that Kraus operators will be first two members of operator basis:

$$E^*(\rho) = \tau_0\rho\tau_0 + \tau_1\rho\tau_1 = \frac{1}{2}(\rho + X\rho X). \tag{2.62}$$

Now we can write the Stinespring dilation of this channel. We have only two Kraus operators, thus a qubit Hilbert space $\mathcal{M}$ is sufficient. The isometry V will then be

$$V = \begin{pmatrix} \tau_0 \\ \tau_1 \end{pmatrix}$$
$$E^*(\rho) = \mathrm{Tr}_2[V\rho V^\dagger] \tag{2.63}$$

It can be checked that this dilation is minimal. This is not very surprising since the environmental overhead is the smallest nontrivial Hilbert space of dimension 2. To this V many unitaries can be formed with different $|O\rangle$ such that

$$U|O\rangle = V. \tag{2.64}$$

One such interesting unitary interaction is U_{cnot} together with $|O\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ where

$$\begin{aligned} U_{\mathrm{cnot}} &= |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes X \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \end{aligned} \tag{2.65}$$

Remark 2.3.9. If the state of environment in some (not minimal) unitary Stinespring dilation of channel $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ is complete mixture, then the channel T^* is unital. Let $T^*(\rho) := \mathrm{Tr}_{\mathcal{M}}\big(U(1/d_{\mathcal{M}}\mathbb{I}_{\mathcal{M}} \otimes \rho)U^\dagger\big)$ for some unitary $U : \mathcal{M} \otimes \mathcal{A} \mapsto \mathcal{M} \otimes \mathcal{B}$, where $d_{\mathcal{M}} = \dim \mathcal{M}$. Then $T^*(\mathbb{I}_{\mathcal{A}}) = 1/d_{\mathcal{M}}\mathrm{Tr}_{\mathcal{M}}\big(U(\mathbb{I}_{\mathcal{M}} \otimes \mathbb{I}_{\mathcal{A}})U^\dagger\big) = 1/d_{\mathcal{M}}\mathrm{Tr}_{\mathcal{M}}(\mathbb{I}_{\mathcal{M}} \otimes \mathbb{I}_{\mathcal{A}}) = \mathbb{I}_{\mathcal{A}}$.

Definition 2.3.10 (Contractive map)**.** a map T^* is called *contractive* if

$$\| T^*(\rho_1 - \rho_2) \|_{\mathrm{tr}} \leq k \| \rho_1 - \rho_2 \|_{\mathrm{tr}}, \tag{2.66}$$

for some $k < 1$. If this inequality holds only for $k \leq 1$, then the map is *non-expansive*.

Every channel is non-expansive, but not all channels are contractive. Due to Banach fixed point theorem, contractive channels have exactly one unique fixed point.

An important norm can be defined on the set of all linear maps.

Definition 2.3.11 (Norm of complete boundedness)**.** Let us have a linear mapping $T^* : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$. The norm of complete boundedness, or the *cb*-norm is then defined as

$$\| T^* \|_{\mathrm{cb}} = \max_{\rho} \frac{\| \mathbb{I}_{\mathcal{H}} \otimes T^*(\rho) \|_{\mathrm{tr}}}{\| \rho \|_{\mathrm{tr}}}, \tag{2.67}$$

where $\rho \in \mathcal{T}(\mathcal{H} \otimes \mathcal{A})$ for arbitrary $\mathcal{H}$ in Schrödinger picture and

$$\| T \|_{\mathrm{cb}} = \max_{G} \frac{\| \mathbb{I}_{\mathcal{H}} \otimes T(G) \|_{\infty}}{\| G \|_{\infty}}, \tag{2.68}$$

where $G \in \mathcal{L}(\mathcal{H} \otimes \mathcal{A})$ for arbitrary $\mathcal{H}$ in Heisenberg picture.

The *cb*-norm can be also defined as the largest difference between the overall probabilities in two statistical quantum experiments differing only by one use of T^*. These experiments may involve entangling the systems on which the channels act with arbitrary further systems.

2.4 Measurement

Untill now we have considered only measurement as an observable which maps states to probability distributions on measurement outcomes. We might be interested in the state after the measurement has occurred. For this we will need the notions of a *measurement* model and *instruments*. These ideas were first formalized by Ozawa [46, 45]. A comprehensive reference on quantum measurements is the monograph by Busch et al [9]. A measurement of some quantum system can be realized by coupling the system to some other quantum system, called a probe, interaction and then measurement of the probe.

Definition 2.4.1 (Measurement model)**.** Let $A : \mathcal{F} \mapsto \mathcal{E}(\mathcal{H})$ be an observable (PPOVM) on our system of concern. A *measurement model* M is a quadruple $M = (\mathcal{K}, \rho_1, V^*, F)$, where

- $\mathcal{K}$ is a Hilbert space of the probe.
- ρ_1 is the initial state of the probe.
- V^* is a channel, $V^* : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}) \mapsto \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$, which describes the measurement interaction between the probe and the system.
- F is an observable of the probe with the outcome space $(\Omega, \mathcal{F})$ taking values in $\mathcal{E}(\mathcal{K})$. We call it the *pointer observable* describing the measurement of probe.

If the following *probability reducibility* condition holds, then M is a measurement model (memo) for the observable A:

$$\mathrm{Tr}[\rho A(X)] = \mathrm{Tr}[V^*\big(\rho \otimes \rho_1\big)\big(\mathbb{I} \otimes F(X)\big)] \quad \forall X \in \mathcal{F}, \rho \in \mathcal{S}(()\mathcal{H}). \tag{2.69}$$

We could have just measured the system itself, without attaching the probe, however this is usually how the measurements work in reality. A direct measurement of a particle often completely destroys the particle, e.g. photon absorption.

A measurement model also defines the state of system after measurement. Let us have a measurement model $M = (\mathcal{K}, \rho_1, V^*, F)$ which defines an observable A with outcome space $(\Omega, \mathcal{F})$. The state of system after measuring event $X \in \mathcal{F}$ is

$$\widetilde{\rho}_X = \frac{1}{\mathrm{Tr}(V^*(\rho \otimes \rho_1)\mathbb{I} \otimes F(X))} \mathrm{Tr}_{\mathcal{K}}(V^*(\rho \otimes \rho_1)\mathbb{I} \otimes F(X)). \tag{2.70}$$

We call the $\widetilde{\rho}_X$ a *conditional output state*. We can introduce an operation

$$\mathcal{I}_X^M(\rho) = \mathrm{Tr}_{\mathcal{K}}(V^*(\rho \otimes \rho_1)\mathbb{I} \otimes F(X)), \tag{2.71}$$

and the conditional output state is

$$\widetilde{\rho}_X = \frac{1}{\mathrm{Tr}(\mathcal{I}_X^M(\rho))}\mathcal{I}_X^M(\rho). \tag{2.72}$$

The mapping $\mathcal{I}_X^M$ has following properties

(C1) for each $X \in \mathcal{F}$, $\mathcal{I}_X^M$ is linear, completely positive and trace non-increasing

(C2) $\mathrm{Tr}(\mathcal{I}_\Omega^M(\rho)) = 1$ and $\mathcal{I}_\emptyset^M(\rho) = O$ for all $\rho \in \mathcal{S}(\mathcal{H})$

(C3) If $\{X_j\}$ is a sequence of mutually disjoint sets, then

$$\mathrm{Tr}(\mathcal{I}_{\cup_j X_j}^M) = \sum_j \mathrm{Tr}(\mathcal{I}_{X_j}^M).$$

These properties can be abstracted and lead to following definition.

Definition 2.4.2 (Instrument)**.** A mapping $\mathcal{I}$ from outcome space $(\Omega, \mathcal{F})$ to the set of trace non-increasing CP-maps on $\mathcal{H}$ is called an *instrument* if it satisfies the properties (C1)-(C3).

We see that every measurement model M defines an instrument $\mathcal{I}^M$ and we say that $\mathcal{I}^M$ is instrument induced by M. Furthermore, due to Ozawa's theorem [46], for every instrument there exists a corresponding measurement model.

Chapter 3

Process estimation

In this short chapter we will introduce some tools for interpreting the data obtained from quantum mechanical experiments. We will start by inverting the statistics to assess the preparation procedure, i.e. the state estimation. Then we will extend this notion to process estimation. The estimation of processes is essentially the same as estimation of states since every completely positive map can be via Choi-Jamiolkowski isomorphism connected with appropriate (unnormalized) state on a larger Hilbert space. Thus the estimation of quantum process can be viewed as estimation of its Choi-Jamiolkowski state. Since the statistics obtained in experiments are finite, we don't have exact probabilities, the inverted map is not necessarily completely positive. If we consider this just to be the effect of finite statistics we can use Bayesian approach and search for a channel with highest likelihood in the set of all channels. This will naturally yield a valid physical map. In reality this is a hard problem of finding global maximum/minimum and has to be solved numerically, with all disadvantages of this approach. It is important that all these approaches assume to have in principle infinitely many independent copies of the state or process. For more information and references about estimations see Ref. [47].

3.1 State estimation

Assume that we have a experimental setup capable of preparing some unknown state $\rho \in \mathcal{S}(\mathcal{H})$. To reconstruct the density matrix of this state we need to measure some observable A to get probability distribution over all effects of the observable. Let E_k correspond to primitive effects of the observable A. Then we try to measure the probabilities $p(\rho)_k = \mathrm{Tr}(\rho E_k)$ by counting the events when particular outcome has occurred. If we want to completely determine the state ρ we need to have a one to one correspondence between the state and the measured probability distribution. This leads to a notion of an informationally complete observable introduced by Prugovečki [53]

Definition 3.1.1 (Informationally complete observable)**.** Observable A is informationally complete if

$$\Phi_A(\rho_1) = \Phi_A(\rho_2) \Rightarrow \rho_1 = \rho_2, \tag{3.1}$$

where $\Phi_A(\rho)$ is probability distribution over outcomes of observable A for state ρ.

Informational completeness assures us that whenever we measure some particular probability distribution, then this probability distribution corresponds to a unique state. In such case we are able to invert the probabilities to obtain ρ. Suppose A is informationally complete. Every effect can be expanded in the traceless operator basis

$$E_k = \sum_j Q_{kj}\tau_j \tag{3.2}$$

and so can be the state ρ

$$\rho = \sum_j r_j\tau_j. \tag{3.3}$$

The probability p_k is then

$$p(\rho)_k = \mathrm{Tr}(\rho E_k) = \sum_{jl} Q_{kj}r_l\mathrm{Tr}(\tau_l\tau_j) = \sum_l Q_{kl}r_l, \tag{3.4}$$

where we used the orthonormality of the traceless operator basis. This can be written in matrix form as $\vec{p}(\rho) = Q.\vec{r}$. To reconstruct the state we make a left inversion

$$Q^{-1}\vec{p}(\rho) = \vec{r}. \tag{3.5}$$

The condition of informational completeness ensures that Q is left invertible.

Example 3.1.2 (S-G along three axes)**.** A common measurement is a Stern Gerlach measurement along some axis, lets say z-axis. In ideal case this is a PVM with two outcomes with effects being the density matrices of eigenstates of Pauli matrix Z:

$$E_{z+} = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{z-} = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.6}$$

This is not an informationally complete observable. For example the eigenstates of X, $|x+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ and $|x-\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$ yield equal probability distributions over E_{z+}, E_{z-}, $p_{z+}(x+) = p_{z+}(x-) = p_{z-}(x+) = p_{z-}(x-) = 1/2$. We can add two S-G measurements along remaining orthogonal axes to get an informationally complete observable. The added effects are

$$E_{x+} = |x+\rangle\langle x+| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad E_{x-} = |x-\rangle\langle x-| = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{3.7}$$

along the x-axis and

$$E_{y+} = |y+\rangle\langle y+| = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad E_{y-} = |y-\rangle\langle y-| = \frac{1}{2}\begin{pmatrix} 0 & i \\ -i & 1 \end{pmatrix}, \tag{3.8}$$

along the y axis. We have to normalize the effects when combining more measurements into one observable. The resulting observable will have the effects $\{q_xE_{x\pm}, q_yE_{y\pm}, q_zE_{z\pm}\}$, where q_k

represents the portion of measurements along each axis. Let us use the qubit traceless operator basis from (2.9) $\{\tau_0, \tau_1, \tau_2, \tau_3\}$. The matrix Q then is

$$Q = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \tag{3.9}$$

where we assume the order

$$\vec{p} = (p_{x+}, p_{x-}, p_{y+}, p_{y-}, p_{z+}, p_{z-})^T. \tag{3.10}$$

This observable is in fact informationally overcomplete and we need any four linearly independent rows from the matrix. This is a bit inconvenient since in real application various rows will give slightly different results, depending on the size of your statistics.

The class of informationally complete observables is broad. Some observables might be better then the other simply because they use less outcomes. Minimal informationally complete observables are those which use the smallest outcome space. This is for d-dimensional Hilbert space, the smallest informationally complete observable has d^2 outcomes, that is the dimension of the operator space. From the set of all minimal informationally complete observables we may select a special class of symmetric informationally complete observables.

Definition 3.1.3 (SIC observable)**.** An observable A on Hilbert space $\mathcal{H}$ is *symmetric informationally complete* (SIC) if:

1. A is minimal, ie. has d^2 outcomes $k \in (1, 2, \dots d^2)$;
2. each effect E_k is a rank-1 operator;
3. $\mathrm{Tr}(E_k) = \alpha$, where α is constant for all k;
4. $\mathrm{Tr}(E_k E_l) = \beta$, where β is constant for all $k \neq l$.

SIC observables are interesting for the high symmetry and efficiency in estimation protocols. It is not known whether there exists a SIC observable in every finite dimensional Hilbert space. Numerical studies have shown that there are SIC observables in all dimensions up to 67 [63]. The constants α and β are fixed by the dimension of Hilbert space $\mathcal{H}$: $\alpha = 1/d$ and $\beta = 1/(d^2(d+1))$.

Example 3.1.4 (SIC qubit observable)**.** The most efficient qubit observable needs at least four effects, three of them are independent, for three independent parameters characterizing the qubit state. The effects have trace $\mathrm{Tr}E_k = 1/2$ and have rank 1, thus they can be parametrized as $E_k = \frac{1}{4}(\mathbb{I} + \vec{r}_k\vec{\sigma})$. The second condition $\mathrm{Tr}(E_k E_l) = 1/12$ gives

$$\frac{1}{12} = \mathrm{Tr}(E_k E_l) = \frac{1}{8}(1 + \vec{r}_k\vec{r}_l) \Rightarrow \vec{r}_k\vec{r}_l = -\frac{1}{3}. \tag{3.11}$$

As a result vectors $\vec{r}_k$ point into corners of a tetrahedron. One possible choice can be

$$\begin{array}{ll} \vec{r}_1 = \frac{1}{\sqrt{3}}(1,1,1) & \vec{r}_2 = \frac{1}{\sqrt{3}}(1,-1,-1) \\ \vec{r}_3 = \frac{1}{\sqrt{3}}(-1,1,-1) & \vec{r}_2 = \frac{1}{\sqrt{3}}(-1,-1,1), \end{array} \tag{3.12}$$

and the matrix Q from (3.2) is $Q_{kj} = \vec{r}_{kj}$.

Sometimes we don't need to fully estimate a state, but instead just test a discrete set of alternatives. For example we have a state ρ which is known to be either ρ_1 or ρ_2. The problem is to answer which of these two it is. If we have only single copy of this state, we can either decide for a strategy where the probability of a wrong identification will be minimal or the probability of wrong identification will be zero, but the probability of inconclusive outcome will be greater than zero. The former strategy is called *minimal error discrimination* and the latter *unambiguous discrimination*.

In minimal error discrimination strategy the probability of error is proportional to the trace norm of operator $\rho_1 - \rho_2$ (see cf. chapter 3 in [28])

$$p_{\text{error}} = \frac{1}{2}(1 - \frac{1}{2}\| \rho_1 - \rho_2 \|_{\text{tr}}), \tag{3.13}$$

where

$$\frac{1}{2}\| \rho_1 - \rho_2 \|_{\text{tr}} =: D(\rho_1, \rho_2) \tag{3.14}$$

is the trace distance between two states. Trace distance has a clear operational meaning. The larger the trace distance is, the smaller the probability of error becomes and hence the easier it is to discriminate between the two states. Naturally, if we are given more copies of the state ρ, it becomes easier to discriminate between the two possibilities, hence

$$D(\rho_1^{\otimes n}, \rho_2^{\otimes n}) \geq D(\rho_1, \rho_2), \tag{3.15}$$

where n is the number of copies of ρ we are given.

Another common way of measuring the difference between quantum states is the *fidelity* of two states ρ_1 and ρ_2

$$F(\rho_1, \rho_2) = \text{Tr}(|\sqrt{\rho_1}\sqrt{\rho_2}|) = \text{Tr}\sqrt{\sqrt{\rho_1}\sqrt{\rho_2}}. \tag{3.16}$$

This quantity provides an upper bound on the optimal success probability for the unambiguous discrimination of pair of states.

3.2 Process estimation

Let us have an unknown process $T^* : \mathcal{T}(A) \mapsto \mathcal{T}(B)$. If we can prepare state $|I\rangle \in \mathcal{S}(\mathcal{A} \otimes \mathcal{A})$ such that

$$|I\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle|j\rangle, \tag{3.17}$$

then

$$\mathbb{I} \otimes T^*(|I\rangle\langle I|) = \frac{1}{d}\chi(T^*) \tag{3.18}$$

is the scaled Choi-Jamiolkowski operator of channel T^*. This can be measured using state estimation on the bipartite system. More generally, any estimation procedure of $T^* : \mathcal{T}(A) \mapsto \mathcal{T}(B)$ consists of preparation of a state $\rho \in \mathcal{S}(\mathcal{K} \otimes \mathcal{A})$, application of the map $\mathbb{I}_\mathcal{K} \otimes T^*(\rho)$ and some measurement of observable A on the Hilbert space $\mathcal{K} \otimes \mathcal{B}$ with effects E_n. Repeating this procedure many times for various input states ρ_k and measurement observables A_l with effects E_{kn}, one then measures conditional probabilities of events $p(E_{ln}|\rho_k, A_l)$

$$p(E_{ln}|\rho_k, A_l) = \mathrm{Tr}(\mathbb{I}_\mathcal{K} \otimes T^*(\rho_k)E_{kl} = \frac{p(E_{ln,}\rho_k, A_l)}{p(\rho_k, A_l)}, \tag{3.19}$$

where $p(\rho_k, A_l)$ is the probability that we have chosen preparation k and measurement l and $p(E_{ln,}\rho_k, A_l)$ is the overall probability of observing event composed of preparation k and effect E_{ln}. The overall probabilities still contain information about details of the estimation scheme. You may choose to repeat the procedure for certain preparations more often then for the other as well as you can use for each preparation a different measurement. The conditional probabilities $p(E_{ln}|\rho_k, A_l)$ are free of this information, they depend only on the process. From these probabilities we can then reconstruct desired parameters.

It is possible to do the complete process estimation using only local states and one local, informationally complete measurement. Let us have some fixed set of states $\{\rho_k\} \in \mathcal{S}(\mathcal{A})$ and an informationally complete observable A in $\mathcal{B}$ with effects E_n. Using the channel many times one measures the probabilities

$$p_{n|k} := p(E_n|\rho_k) = \mathrm{Tr}\big(T^*(\rho_k)E_n\big). \tag{3.20}$$

Again we can define the matrix Q as in (3.2) and a matrix S such that

$$\begin{aligned} E_n &= \sum_j Q_{nj}\tau_j \\ \rho_k &= \sum_i S_{ki}\tau_i \end{aligned} \tag{3.21}$$

where $\{\tau_i\}$ is the traceless operator basis in $\mathcal{A}$ and in $\mathcal{B}$ if they have same dimensions. One can then write

$$\begin{aligned} p_{n|k} &= \sum_{ij} Q_{nj}S_{ki}\mathrm{Tr}\big(T^*(\tau_i)\tau_j\big) = \sum_{ij} Q_{nj}S_{ki}A_{ij} \\ &\Rightarrow \quad p = SAQ^T \end{aligned} \tag{3.22}$$

where A_{ij} is the vector representation of channel T^*. The matrix A can be reconstructed by inverting the matrices Q and S as in previous section. For this we need that the set of states $\{\rho_k\}$ spans the whole operator space, then the matrix S is invertible.

As with states, we can also do process discrimination. Minimum error discrimination of two processes, T_1^*, T_2^* then falls down to discrimination of the output states $T_1^*(\rho)$ and $T_2^*(\rho)$. We can use a strategy when we prepare an input state $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{A})$ and discriminate states $\mathbb{I}_{\mathcal{H}} \otimes T_1^*(\rho)$ and $\mathbb{I}_{\mathcal{H}} \otimes T_2^*(\rho)$. Thus the distinguishability of two processes will be proportional to the largest trace distance of states $\mathbb{I}_{\mathcal{H}} \otimes T_1^*(\rho)$ and $\mathbb{I}_{\mathcal{H}} \otimes T_2^*(\rho)$, that is the *cb*-norm (2.67).

3.3 Maximum likelihood

Both methods from previous two sections require measuring probabilities. This can never be done perfectly with finite statistics. In real experiment we measure the number of occurrences n_k of some event k. For large enough statistics the probability is then assumed to be n_k/N where N is the number of all events. Due to the small errors in probabilities the resulting inversion might not be a physical state or process. We can then ask what process or state describes the measured data best. In such case we can use the maximum likelihood method to estimate the parameters of a process or state. We construct the likelihood function as

$$L(R, D) = p(D \mid R) \tag{3.23}$$

where R is the set of parameters of the statistical model we try to estimate and D represents the measured data. Likelihood function is then the probability of measuring data D given that the statistical model has parameters R. Maximizing this function over the allowed set of model parameters will give us the parameters R such that this probability is maximal.

Main problem with this method is that the likelihood function tends to be quite complex, and finding the global maximum is a very hard task, especially when the maximum lies on the border of the parameter space. On the other side it has lot of nice properties. It uses all measured data, no post selection is needed. It behaves well asymptotically, for large statistics it converges to the real parameters of the model. By construction, the result is always a physical map or state.

3.4 Teleportation experiment

In an actual experiment of group in Heidelberg under supervision of Prof. Jian-Wei Pan, a scheme for teleportation was tested. Teleportation is a scheme for communicating one quantum state from site A to site B. Without going into details of the protocol, the result is an ideal channel from sender to receiver. At the end one has to assess how much the practical realization did deviate from an ideal channel, and thus how successful the teleportation was. In this experiment, six preparations were used: $\{H, V, P, N, R, L\}$, corresponding to various light polarizations. Ideally ρ_H, ρ_V should be the horizontal and vertical polarizations, $|0\rangle$ and $|1\rangle$, ρ_P, ρ_N should be the $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle), |+\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$ and finally ρ_R, ρ_L the $|R\rangle = 1/\sqrt{2}(|0\rangle + i|1\rangle), |L\rangle = 1/\sqrt{2}(|0\rangle - i|1\rangle)$. However the preparations are not ideal and we have to estimate them. Three projective measurements with outcomes H, V, P, N and R, L polarizations were used where $E_H = 1/2(\mathbb{I} + Z)$, $E_V = 1/2(\mathbb{I} - Z)$, $E_P = 1/2(\mathbb{I} + X)$,$E_N = 1/2(\mathbb{I} - X)$, $E_R = 1/2(\mathbb{I} + Y)$ and $E_L = 1/2(\mathbb{I} - Y)$, this is equivalent to the ones in example 3.1.2 . The measured data is summarized on figure 3.1.

Prep.	Out.	Count	Tel. count	Prep.	Out.	Count	Tel. count
H	H	1305	53	V	H	33	7
	V	29	0		V	1067	56
	P	613	30		P	709	27
	N	670	20		N	516	42
	R	718	30		R	497	30
	L	592	28		L	652	38
P	H	759	25	N	H	548	20
	V	596	23		V	557	32
	P	1426	55		P	26	7
	N	25	6		N	1052	47
	R	674	36		R	586	45
	L	686	51		L	542	20
R	H	635	21	L	H	657	25
	V	672	29		V	506	41
	P	777	45		P	588	33
	N	576	26		N	591	40
	R	1369	49		R	39	4
	L	49	7		L	1084	49

Figure 3.1: Experimental data for teleportation protocol. (H,V), (P,N) and (R,L) correspond to qubit preparations and measurements along three perpendicular axes. In column "Count" data for measurement on preparations is placed, $n_{k,l}$, and in column "Tel. count" measured data for teleported preparations, $n_{(k,l)}$.The teleported data have much lower count due to probabilistic nature of the protocol used.

Two datasets are shown. One for estimating the preparations, where no teleportation of the inputs has occurred and one for teleportation of the preparations. The counts for the teleportation are much lower due to probabilistic nature of the protocol.

Since the data for preparation estimation is overcomplete for each preparation, it is best to choose the maximal likelihood method to find the prepared states. Each preparation k is a qubit state parametrized by a three dimensional vector $\vec{r_k}$ such that

$$\rho_k = \frac{1}{2}(\mathbb{I} + r_{k,x}X + r_{k,y}Y + r_{k,z}Z). \tag{3.24}$$

The likelihood function is then the probability that preparation k with parameters $\vec{r_k}$ will yield measured set of counts $\{n_{k,l}\}$, $l \in \{V, H, N, P, R, L\}$:

$$L(\vec{r_k}, \{n_{k,l}\}) = \prod_l p_{l|k}^{n_{k,l}}, \tag{3.25}$$

where

$$p_{l|k} = \mathrm{Tr}(\rho_k E_l) \tag{3.26}$$

is the probability of measuring effect E_l when preparation k with parameters $\vec{r_k}$ was input. Maximizing over the parameters $\vec{r_k}$ we obtain the result[1]. Alternatively it is numerically easier to, instead of maximizing the function L, which contains very small numbers for large counts to minimize $-\log L = \sum_l n_{k,l} \log 1/p_{l|k}$. Either way, using this method we obtain following results

$$\begin{aligned}
\rho_H &= \begin{pmatrix} 0.98 & -0.02 - i0.05 \\ -0.02 + i0.05 & 0.02 \end{pmatrix} \quad \rho_V = \begin{pmatrix} 0.03 & 0.08 + i0.07 \\ 0.08 - i0.07 & 0.97 \end{pmatrix} \\
\rho_P &= \begin{pmatrix} 0.56 & 0.48 \\ 0.48 & 0.44 \end{pmatrix} \quad \rho_N = \begin{pmatrix} 0.5 & -0.48 - i0.02 \\ -0.48 + i0.02 & 0.5 \end{pmatrix} \\
\rho_R &= \begin{pmatrix} 0.49 & 0.07 - i0.47 \\ 0.07 + i0.47 & 0.51 \end{pmatrix} \quad \rho_L = \begin{pmatrix} 0.56 & +i0.47 \\ -i0.47 & 0.44 \end{pmatrix}.
\end{aligned} \tag{3.27}$$

Then we can continue in same manner with process estimation. We parametrize the process T^* by its Choi-Jamiolkowski state

$$\chi(T^*) = \sum_{a,b=0}^{1} 1/2|a\rangle\langle b| \otimes T^*(|a\rangle\langle b|) \tag{3.28}$$

and the probability of event (k, l) is then

$$p(k, l) = \mathrm{Tr}(T^*(\rho_k)E_l) = \mathrm{Tr}(\chi(T^*)\rho_k^T \otimes E_l), \tag{3.29}$$

with the likelihood function

$$L(\chi, n_{(k,l)}) = \prod_{(k,l)} p(k, l)^{n_{(k,l)}}. \tag{3.30}$$

This yields

$$\chi(T^*) = \begin{pmatrix} 0.46 & 0.04 - i0.05 & -0.02 - i0.04 & 0.42 - i0.04 \\ 0.04 + i0.05 & 0.04 & 0.03 + i0.01 & 0.02 + i0.04 \\ -0.02 + i0.04 & 0.03 - i0.01 & 0.06 & -0.04 + i0.05 \\ 0.42 + i0.04 & 0.02 - i0.04 & -0.04 - i0.05 & 0.44 \end{pmatrix}. \tag{3.31}$$

For comparison the Choi-Jamiolkowski state of ideal transfer in this representation is

$$\chi(I) = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{3.32}$$

[1]Actually one can use directly the overall probabilities of events, since this just changes the multiplicative constant in front of the likelihood function, but doesn't change the position of maxima in the parameter space.

For numerical minimization of the function $-\log L(\chi, n_{(k,l)})$ we used downhill simplex method, also known as Nelder-Maeds method.

A separate problem is to introduce "error bars" on the result. A commonly used procedure is to perform Monte-Carlo simulation of the experiment, with events distributed according to Poisson distribution with mean values of the actual measured data. Then from these datasets estimate the state or process, and the error in terms of a distance measure, for example fidelity of the Choi-Jamiolkowski state.

To obtain the error bars of the preparation procedures we made 10000 simulations. Results are summarized in table 3.2.

Preparation	Mean fidelity [%]	5σ interval [%]
H	98	± 2
V	97	± 3
P	98	± 2
N	98	± 2
R	97	± 2
L	97	± 3

Figure 3.2: Fidelity of preparation procedures.

The fidelity of the teleportation was $85 \pm 12\%$(5 sigma) with respect to ideal transfer, where we sampled 1044 simulations to obtain the distribution, see figure 3.3. This is well above the classical threshold of 2/3, which can be achieved without entanglement.

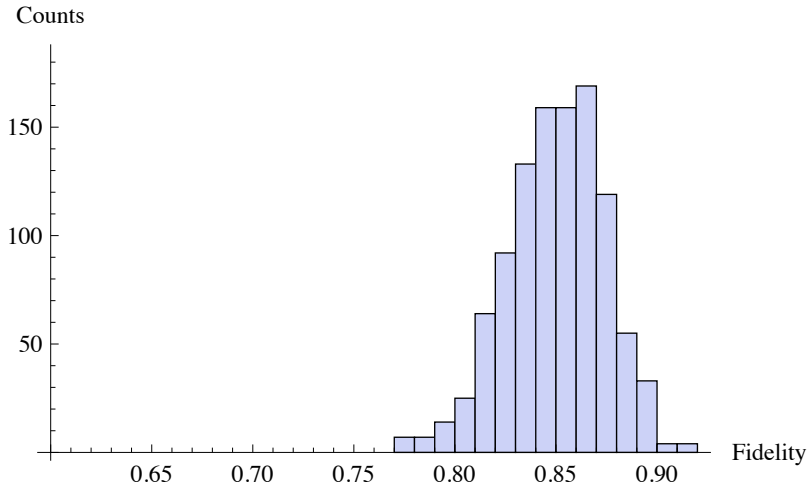


Figure 3.3: Distribution of fidelities of 1000 simulations of the experiment. The 5 sigma interval covers the whole data set.

Chapter 4

Quantum memory channels

Schrödinger equation governs the evolution of a closed quantum system. Such evolution is unitary. It is difficult if not impossible to experimentally achieve such ideal situation. Unavoidable interaction between the system under consideration and environment occurs leading to a non-unitary evolution of the system. Such non-unitary evolution is described by a quantum channel - a linear, completely positive and trace preserving map in Schrödinger picture on quantum states or equivalently by linear, completely positive and unital map in Heisenberg picture on observables. Stinespring theorem gives us full justification for this step, because it tells us that every quantum channel can be extended to a unitary evolution on a larger Hilbert space.

In this chapter we will introduce the model of quantum channels with memory, memory channels. Quantum memory channel is a simple quantum channel of some system, where the system has an inner structure and additional physical requirement of causality is placed on the channel. The system is an ordered, infinite sequence of subsystems and the causality condition requires that information can move only in one way. Information stored in a particular subsystem can then affect evolution only of subsystems which follow in sequence after. This can be interpreted as a simple collision model where some quantum device sequentially processes a sequence of quantum signals. The research of memory channels is mostly focused on investigation of transmission rates for particular classes of memory channels [37, 42, 43, 2, 4, 23, 32, 31, 13, 16, 15, 10, 69, 19]. Recently, attention has been paid to an interesting class of the so-called bosonic memory channels [24, 11, 57, 56, 50, 41, 39, 61, 40] and also to memory effects in the transmission of quantum states over spin chains [51, 3, 52, 55].

4.1 Memory channel as general causal process

As we have already mentioned a memory channel is just a special case of a simple quantum channel on a specially structured input. The aforementioned structure of input is as follows. The input is a chain of ordered quantum systems each with a finite dimensional Hilbert space $\mathcal{A}_k$ where $k \in \mathbb{Z}$ denotes the ordering. This can be understood as infinite sequence of quantum particles or cells as in the case of quantum cellular automata. Since this sequence is potentially infinite, we will start by describing the channel in Heisenberg picture.

Every subsystem comes with an algebra of observables $\mathfrak{A}_k = \mathcal{L}(\mathcal{A}_k)$. If we would like to

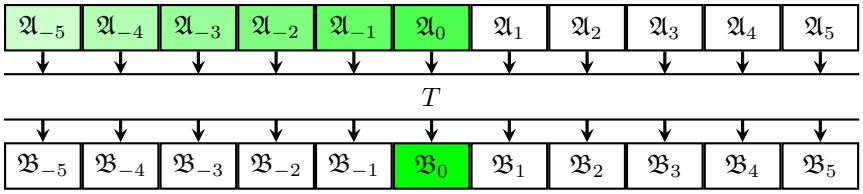


Figure 4.1: A causal process has the property that any information that is stored at the output $\mathfrak{B}_0$(bottom green) could only originate from inputs on the left from this site (upper green region).

denote algebras associated to sets of subsystems $\Lambda \subset \mathbb{Z}$ we follow this notation: if Λ is a finite subset of $\mathbb{Z}$ then $\mathfrak{A}_\Lambda = \bigotimes_{k\in\Lambda} \mathfrak{A}_k$. $\mathfrak{A}_\Lambda$ is then just a simple tensor product of selected subsystems. In case of infinite $\Lambda \subset \mathbb{Z}$ we associate with $\mathfrak{A}_\Lambda$ the C^*-closure of increasing family of finite dimensional algebras $\mathfrak{A}_{\Lambda_f}$ for finite $\Lambda_f \subset \Lambda$. This is also called a *quasi-local algebra* [5]. From now on $\mathfrak{A}_\mathbb{Z}$ will be abbreviated as $\mathfrak{Z}$ and we will also use the shorthand $\mathfrak{A}_-$ and $\mathfrak{A}_+$ for the left and right hand chain halves $\mathfrak{A}_{(-\infty,0]}$ and $\mathfrak{A}_{[1,\infty)}$ when it comes in handy. For clarity we will also make difference between input algebras $\mathfrak{A}$ and output algebras $\mathfrak{B}$ even if we consider them isomorphic. The terms $\mathfrak{B}_{[a,b]}$,$\mathfrak{B}_\pm$ and similar are then defined analogously. Also we will denote the Hilbert space of outputs with $\mathcal{B}_k$ even though it is isomorphic to $\mathcal{A}_k$

Definition 4.1.1 (Causal process)**.** Causal process is a channel (in Heisenberg picture) $T : \mathfrak{Z} \mapsto \mathfrak{Z}$ such that for every $z \in \mathbb{Z}$

$$T(b_{(-\infty,z]} \otimes \mathbb{I}^{\mathcal{B}}_{[z+1,\infty)}) = T(b_{(-\infty,z]}) \otimes \mathbb{I}^{\mathcal{A}}_{[z+1,\infty)}, \tag{4.1}$$

where $B_{(-\infty,z]} \in \mathfrak{B}_{(-\infty,z]}$ and $\mathbb{I}^{\mathcal{B}}_{[z+1,\infty)}$ is the identity operator in $\mathfrak{B}_{[z+1,\infty)}$ and $\mathbb{I}^{\mathcal{A}}_{[z+1,\infty)}$ is the identity operator in $\mathfrak{A}_{[z+1,\infty)}$.

Remark 4.1.2 (Notation)**.** To avoid too many identical subscripts we adopted following convention. For any causal channel $T(\mathfrak{B})$, if we would like to address its transformation on a subpart localized in region $[a, b]$ we could write $T_{[a,b]}(\mathfrak{B}_{[a,b]})$, however we omit the first subscript and write only $T(\mathfrak{B}_{[a,b]})$. When we see $T(\cdot)$ we need to know where the $\cdot$ is localized to correctly understand the transformation. This goes also for Schrödinger picture where the transformation will be specified by the localization of inputs.

The equation (4.1) just reflects the causality condition placed on the channel T. It means that whatever you can measure on the output with $b_{(-\infty,z]}$ can be measured on inputs up to the z-th particle with $T(b_{(-\infty,z]}) \in \mathfrak{A}_{(-\infty,z]}$. This is illustrated on Figure 4.1 where it is shown that a single localized output cell is smeared over infinitely many input cells localized left from the original cell.

We can also write this definition in Schrödinger picture. Since the associated Hilbert space $\mathcal{A}_{(-\infty,\infty)}$ is not separable there are necessarily problems with describing the whole state space. However describing finite sets of subsystems is possible. Thus we can write that a causal process is a sequence of channels (in Schrödinger picture) $T^* : \mathcal{T}(\mathcal{A}_{[1,z]}) \mapsto \mathcal{T}(\mathcal{B}_{[1,z]})$ for every $z \in \mathbb{N}$ such that

$$T^*(\omega_{[1,x]}) = \mathrm{Tr}_{[x+1,y]}[T^*(\omega_{[1,y]})], \tag{4.2}$$

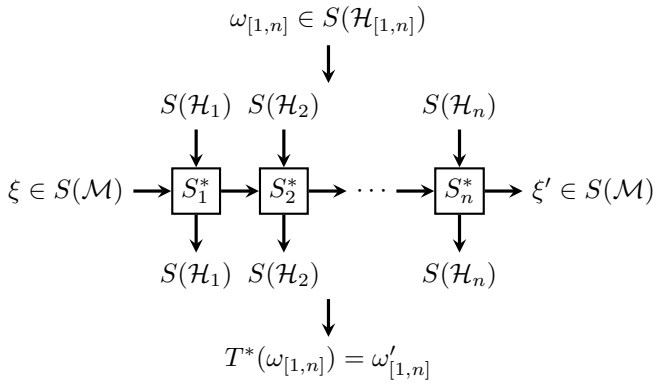


Figure 4.2: Environment particle ξ is colliding causally with n particles of input sequence, where each collision is described by a channel S_i

for every $x < y$, $x, y \in \mathbb{N}$ and every $\omega_{[1,x]}$, $\omega_{[1,y]}$, such that

$$\omega_{[1,x]} = \mathrm{Tr}_{[x+1,y]}(\omega_{[1,y]}), \tag{4.3}$$

where $\mathrm{Tr}_{[a,b]}$ denotes partial trace over the interval.

Example 4.1.3 (Shift of a chain)**.** The channel σ_x (in Heisenberg picture) is defined on a translationally invariant chain of quantum systems with Hilbert spaces $\mathcal{A}_j \equiv \mathcal{A}$ and appropriate observable algebras $\mathfrak{A}_j \equiv \mathfrak{A}$ and $\mathfrak{B}_j \equiv \mathfrak{B}$. The σ_x is defined by

$$\sigma_x(\lambda_j^a) = \lambda_{j-x}^a \tag{4.4}$$

for all $\lambda_y^a \in \mathfrak{A}_y$ where $\{\lambda_y^a\}_a$ form the operator basis in $\mathfrak{A}_y$. This shift channel just shifts the chain by x sites to the left. In Schrödinger picture it can be defined as

$$\sigma_x^*(\omega_{[1,n]}) = \omega_{[1-x,n-x]}. \tag{4.5}$$

A memory channel is a general causal process. A physically more compelling insight might be conveyed through the view of collision models in next section.

4.2 Collision model

Let us look on a situation where an infinite sequence of quantum systems is colliding one at a time with some environment. This situation is illustrated on figure 4.2.

The environment is described by a single quantum system in Hilbert space $\mathcal{M}$ in state ξ. This environment sequentially collides with subparts of the input signal of length n. The i-th collision is described by a channel $S_i^* : \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{A}_i) \mapsto \mathcal{T}(\mathcal{B}_i) \otimes \mathcal{T}(\mathcal{M})$. The input signal is then processed with a channel T^*:

$$T^*(\omega_{[1,n]}) = \mathrm{Tr}_{\mathcal{M}}[S_{[1,n]}^*(\xi \otimes \omega_{[1,n]})], \tag{4.6}$$

where $S^*_{[1,n]}$ is an n-fold concatenation of subsequent collisions,

$$\begin{aligned} S^*_{[1,n]} &= (\mathbb{I}_{[1,n-1]} \otimes S^*_n) \cdots (\mathbb{I}_1 \otimes S^*_2 \otimes \mathbb{I}_{[3,n]}) \cdot (S^*_1 \otimes \mathbb{I}_{[2,n]}) \\ &= S^*_n \cdots S^*_2 \cdot S^*_1. \end{aligned} \tag{4.7}$$

Note that in the definition of collision a swap of environmental system and input is included by default. Though this makes things a tiny bit less readable for the first time, it simplifies significantly the notation for concatenation of collisions as can already be appreciated in equation (4.7).

If $S^*_i = S^*$ for all i the we call such collision model translationally invariant. If all S^*_i are unitary we call such collision model pure. In previous lines we have assumed implicitly that the environment and the input signals are initially uncorrelated. We would like to stress that this does not need to be the case, however it won't change the picture and the message of this section. One thing needs to be said though, that the description by channel T^* is possible only when the memory is initially uncorrelated with the input signal, otherwise T^* wont be a channel, though the way how to acquire the output of the memory channel will remain the same.

It is easy to see that this channel is causal from the construction of the model. Each subsequent collision can change only the environment or the colliding signal, all signals which collided before cannot be affected by subsequent inputs. However the inputs which follow after can be affected, where the effect is mediated through the environmental particle ξ. This is also the motivation to call this environment memory, because it "remembers" the inputs from "past" and mediates their effect in "present". From now on *memory* will refer to the environment system in the appropriate collision model.

Let us continue with a very basic example of a simple swap interaction where the memory and colliding input are simply swapped.

Example 4.2.1 (Swap collision)**.** Let $\sigma^* : \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B}) \otimes \mathcal{T}(\mathcal{M})$ be a translationally invariant pure collision model with a qubit memory and a chain of qubits where the unitary interaction is defined as

$$U_\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.8}$$

with $U_\sigma : \mathcal{A} \otimes \mathcal{M} \mapsto \mathcal{B} \otimes \mathcal{M}$. Let the initial state of memory be $\xi \in \mathcal{T}(\mathcal{M}) \equiv \mathcal{T}(\mathcal{A})$, then the transformation on first qubit in state $\omega_1 \in \mathcal{T}(\mathcal{A}_1)$ reads

$$\begin{aligned} T^*(\omega_1) &= \mathrm{Tr}_{\mathcal{M}}[{\sigma^*}_1(\xi \otimes \omega_1)] = \mathrm{Tr}_{\mathcal{M}}[U_\sigma(\xi \otimes \omega_1)U_\sigma^\dagger] \\ &= \mathrm{Tr}_{\mathcal{M}}[\omega_1 \otimes \xi] = \xi. \end{aligned} \tag{4.9}$$

The state of memory after first collision is

$$\xi' = \mathrm{Tr}_1[U_\sigma(\xi \otimes \omega_1)U_\sigma^\dagger] = \omega_1. \tag{4.10}$$

The transformation on first n qubits is then straightforward:

$$\begin{aligned} T^*(\omega_{[1,n]}) &= \mathrm{Tr}_{\mathcal{M}}[{\sigma^*}_{[1,n]}(\xi \otimes \omega_{[1,n]})] \\ &= \mathrm{Tr}_{\mathcal{M}}[(\mathbb{I}_{[1,n-1]} \otimes FU_\sigma)\dots(FU_\sigma \otimes \mathbb{I}_{[2,n]}) \\ &\quad (\xi \otimes \omega_{[1,n]})(U_\sigma^\dagger F^\dagger \otimes \mathbb{I}_{[2,n]})\dots(\mathbb{I}_{[1,n-1]} \otimes U_\sigma^\dagger F^\dagger)] \\ &= \mathrm{Tr}_{\mathcal{M}}[\xi \otimes \omega_{[1,n]}] = \xi \otimes \omega_{[1,n-1]}, \end{aligned} \tag{4.11}$$

where F swaps the memory and the system and is included for the nice concatenation properties. This way the memory system is the last subsystem after transformation and first at the begining. Note that in this case $F = U = U^\dagger = F^\dagger$ and the interaction seems trivial, however the position of memory system has changed. Thus the state of memory system after n collisions will be $\xi' = \omega_n$. This is in fact the shift channel σ_1^* from Example 4.1.3.

Another interesting example consists of control-U interaction.

Example 4.2.2. [Controlled U collision] Let us have again a translationally invariant pure collision model $\gamma^* : \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B}) \otimes \mathcal{T}(\mathcal{M})$ with memory living in Hilbert space $\mathcal{M}$, $\dim \mathcal{M} = d_{\mathcal{M}}$ and chain of qudits in Hilbert space $\mathcal{A}$, $\dim \mathcal{A} = d$. Where the collision is described by unitary

$$U_\gamma = \sum_{i=0}^{d_{\mathcal{M}}-1} |i\rangle\langle i| \otimes U_i, \tag{4.12}$$

with $U_\gamma : \mathcal{A} \otimes \mathcal{M} \mapsto \mathcal{B} \otimes \mathcal{M}$ and where $\{|i\rangle\}$ is the basis of $\mathcal{M}$ and U_i are some unitaries on $\mathcal{A}$. The transformation on inputs reads

$$\begin{aligned} T^*(\omega_{[1,n]}) &= \mathrm{Tr}_{\mathcal{M}}[{\gamma^*}_{[1,n]}(\xi \otimes \omega_{[1,n]})] \\ &= \mathrm{Tr}_{\mathcal{M}}[({U_i}^{\otimes n}\omega_{[1,n]}{U_j^\dagger}^{\otimes n}) \otimes (|i\rangle\langle i|\xi|j\rangle\langle j|)] \\ &= \sum_{i=0}^{d_{\mathcal{M}}-1} \langle i|\xi|i\rangle {U_i}^{\otimes n}\omega_{[1,n]}{U_i^\dagger}^{\otimes n}, \end{aligned} \tag{4.13}$$

and the state of ξ after nth collision will be

$$\langle i|\xi'|j\rangle = \langle i|\xi|j\rangle \mathrm{Tr}[{U_i}^{\otimes n}\omega_{[1,n]}{U_j^\dagger}^{\otimes n}]. \tag{4.14}$$

Note that the diagonal elements of memory do not change because

$$\mathrm{Tr}[{U_i}^{\otimes n}\omega_{[1,n]}{U_i^\dagger}^{\otimes n}] = \mathrm{Tr}\omega_{[1,n]} = 1. \tag{4.15}$$

Example 4.2.3 (Memoryless channel)**.** Let us have a collision model $\mathcal{E}_k^* : \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k) \otimes \mathcal{T}(\mathcal{M})$ with memory living in Hilbert space $\mathcal{M}$, $\dim \mathcal{M} = d_{\mathcal{M}}$ where

$$\mathcal{E}_k^*(\xi \otimes \omega_k) = (\mathbb{I}_k \otimes P_k^*)\big(U_k(\xi \otimes \omega_k)U_k^\dagger\big), \tag{4.16}$$

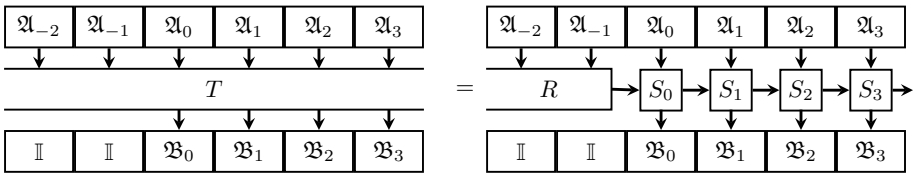


Figure 4.3: Representation of structure theorem where causal process T is constructed as a collision model with memory initializer R. Note that the figure is in Heisenberg picture and therefore has to be read *against* the arrows which represent time.

for all $\xi \in \mathcal{T}(\mathcal{M})$, $\omega_k \in \mathcal{T}(\mathcal{A}_k)$ and some $U_k : \mathcal{M} \otimes \mathcal{A}_k \mapsto \mathcal{B}_k \otimes \mathcal{M}$ and P^* a completely depolarizing channel on $\mathcal{M}$ such that

$$P_k^*(m) = \xi_{k+1}, \tag{4.17}$$

for all $m \in \mathcal{T}(\mathcal{M})$ where ξ_k is a predefined sequence of states in $\mathcal{T}(\mathcal{M})$.

This interaction will produce a transformation on the input sequence

$$T^*(\omega_{[1,n]}) = (T_1^* \otimes \ldots \otimes T_n^*)(\omega_{[1,n]}), \tag{4.18}$$

where

$$T_k^*(\omega_k) = \mathrm{Tr}_{\mathcal{M}}[U_k(\xi_k \otimes \omega_k)U_k^\dagger] \tag{4.19}$$

are channels which act on every subsystem of input signal independently. We call such action memoryless since the behavior of channel is not affected by any of the inputs and the action is independent on every subsystem. If we would like to make such model pure, we would need to provide for each P_k^* a separate dilation space and that would result into infinite resources needed for memory system.

We see that every collision model then defines a causal process on input signals therefore a memory channel. It would be interesting to know whether all causal processes can be connected to an appropriate collision model. In [37] structure theorem is proved which tells us that such collision model exists for every translationally invariant causal process. A weaker structure theorem for finite sequences was also proved in [21] and this proof does not require translational invariance. In the next section we will present the structure theorem.

4.3 Structure theorem

We will present the structure theorem and its proof as it is stated in [37] with a minor change, we drop the assumption of translational invariance.

Theorem 4.3.1 (Structure theorem)**.** *Let $T : \mathfrak{Z} \mapsto \mathfrak{Z}$ be a causal channel. Ignore the outputs on its left half chain $\mathfrak{B}_-$. Then there exists a memory observable algebra $\mathfrak{M}$ and an initializing channel $R : \mathfrak{M} \mapsto \mathfrak{A}_-$ such that for all $n \in \mathbb{N}$*

$$T(\mathbb{I}_- \otimes b_{[1,n]}) = (R \otimes \mathbb{I}_{[1,n]})S_{[1,n]}(b_{[1,n]} \otimes \mathbb{I}^{\mathcal{M}}) \tag{4.20}$$

for all $b_{[1,n]} \in \mathfrak{B}_{[1,n]}$, where $S_{[1,n]}$ is the n-fold concatenation of a collision model (in Heisenberg picture) $S_i : \mathfrak{B}_i \otimes \mathfrak{M} \mapsto \mathfrak{M} \otimes \mathfrak{A}_i$

Proof. Let $\mathcal{H}_-$ be the Hilbert space associated with universal representation of the left half chain $\mathfrak{B}_-$. Further let $(\mathcal{K}, \pi, V)$ be the minimal Stinespring dilation for $T\,|_{\mathfrak{B}_-}$:

$$T(b) = V^\dagger \pi(b_-)V \quad \forall b_- \in \mathfrak{B}_- \tag{4.21}$$

with Stinespring isometry $V : \mathcal{H}_- \mapsto \mathcal{K}$. From Stinespring's representation (4.21) and the causality property (4.1) we can conclude that

$$\begin{aligned} V^\dagger \pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]})V &= T(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]}) = T(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]} \\ &= (V^\dagger \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]})(\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]})(V \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) \end{aligned} \tag{4.22}$$

for all $b_- \in \mathfrak{B}_-$. Since V is minimal dilation for T so is $V \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}$ for $T \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}$. We know then that there exists an isometry $W_{[1,n]} : \mathcal{K} \otimes \mathcal{H}_{[1,n]} \mapsto \mathcal{K}$ defined by

$$\begin{aligned} W_{[1,n]}(\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]})(V \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]})\psi_- \otimes \psi_{[1,n]} \\ := \pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]})V\psi_- \otimes \psi_{[1,n]} \end{aligned} \tag{4.23}$$

for all $b_- \in \mathfrak{B}_-$, $\psi_- \in \mathcal{H}_-$ and $\psi_{[1,n]} \in \mathcal{H}_{[1,n]}$ such that

$$\pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]})W_{[1,n]} = W_{[1,n]}(\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) \tag{4.24}$$

for all $b_- \in \mathfrak{B}_-$, and

$$W_{[1,n]}(V \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) = V. \tag{4.25}$$

Now let $\mathfrak{M} := \pi'(\mathfrak{B}_-)$ be the commutant of the observable algebra $\mathfrak{B}_-$ and let $S_{[1,n]} : \mathfrak{B}_{[1,n]} \otimes \mathfrak{M} \mapsto \mathcal{L}(\mathcal{K}) \otimes \mathfrak{A}_{[1,n]}$ be defined by

$$S_{[1,n]}(b_{[1,n]} \otimes m) := W^\dagger_{[1,n]} \pi(b_{[1,n]}) m W_{[1,n]} \tag{4.26}$$

for all $b_{[1,n]} \in \mathfrak{B}_{[1,n]}$ and $m \in \mathfrak{M}$. The memory initializing channel $R : \mathfrak{M} \mapsto \mathfrak{A}_-$ is then given by

$$R(m) := V^\dagger m V \quad \forall m \in \mathfrak{M}. \tag{4.27}$$

In order to justify this choice we need to show that

$$S_{[1,n]}(\mathfrak{B}_{[1,n]} \otimes \mathfrak{M}) \subset \mathfrak{M} \otimes \mathfrak{A}_{[1,n]}. \tag{4.28}$$

Since $\mathfrak{B}_{[1,n]}$ and $\mathfrak{M}$ commute with $\mathfrak{B}_-$ we see from (4.24) that

$$\begin{aligned}
& S_{[1,n]}(b_{[1,n]} \otimes m) \cdot (\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) \\
= \; & W^{\dagger}_{[1,n]} \pi(\mathbb{I}^{\mathcal{B}}_{-} \otimes b_{[1,n]}) m W_{[1,n]} (\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) \\
= \; & W^{\dagger}_{[1,n]} \pi(\mathbb{I}^{\mathcal{B}}_{-} \otimes a_{[1,n]}) m \pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]}) W_{[1,n]} \\
= \; & W^{\dagger}_{[1,n]} \pi(\mathbb{I}^{\mathcal{B}}_{-} \otimes a_{[1,n]}) \pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]}) m W_{[1,n]} \\
= \; & W^{\dagger}_{[1,n]} \pi(b_- \otimes \mathbb{I}^{\mathcal{B}}_{[1,n]}) \pi(\mathbb{I}^{\mathcal{B}}_{-} \otimes b_{[1,n]}) m W_{[1,n]} \\
= \; & (\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) W^{\dagger}_{[1,n]} \pi(\mathbb{I}^{\mathcal{B}}_{-} \otimes b_{[1,n]}) m W_{[1,n]} \\
= \; & (\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) \cdot S_{[1,n]}(b_{[1,n]} \otimes m) \qquad (4.29)
\end{aligned}$$

for all $b_{[1,n] \in \mathfrak{B}_{[1,n]}}$, $b_- \in \mathfrak{B}_-$ and $m \in \mathfrak{M}$, implying that $S_{[1,n]}(b_{[1,n]} \otimes m)$ commutes with $\pi(b_-) \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}$ from which (4.28) follows. The next thing we need to show is that $S_{[1,n]}$ has the right concatenation properties:

$$\begin{aligned}
R(m) &= (R \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) S_{[1,n]}(\mathbb{I}^{\mathcal{B}}_{[1,n]} \otimes m) \quad \text{and} \\
T(b_{[1,n]}) &= (R \otimes \mathbb{I}^{\mathcal{A}}_{[1,n]}) S_{[1,n]}(b_{[1,n]} \otimes \mathbb{I}^{\mathcal{M}}), \qquad (4.30)
\end{aligned}$$

however this directly follows from definitions of $S_{[1,n]}$, R and (4.25). For the translationally invariant case to obtain the result we set $S_i = S := S_{[1,1]}$. In the translationally non-invariant case we need to parcel the $S_{[1,n]}$ into separate S_i, what can be done via the weaker structure theorem presented in [21]. The proof again uses the uniqueness of minimal Stinespring's dilation and the same tools used in this proof. The structure theorem is visualized on figure 4.3. □

For every causal channel there exists a collision model $S_{[1,n]}$ and appropriate channel initializer R which has an equivalent input output behavior. This is an important result which proves that collision models are a highly general concept.

4.4 Quantum cellular automata

The idea of generalizing the notion of classical cellular automata to quantum settings, can be traced back to Feynman where he in his paper [22] argued that quantum computation might out power the classical one. Various approaches have been considered for generalizing cellular automata to quantum regime [66, 14, 65, 20, 26]. In this section we will introduce the approach of Werner et al. [54, 62, 1, 25] and in following chapter we will show that it bears an explicit connection to memory channels.

Cellular automata are transformations of certain cell structure by a set of local rules. In quantum setting a single cell will refer to single quantum system in finite dimensional Hilbert space $\mathcal{A}_k$ with $d(k) = \dim \mathcal{A}_k$ and the cell structure will then be formed by observable algebras $\mathcal{L}(\mathcal{A}_k) = \mathfrak{A}_k$ resulting into the quasi-local algebra as in Section 4.1. This also favors the Heisenberg picture as preferable when dealing with such transformations. In the whole work we will

deal only with one dimensional automata, also index of a cellular automata is defined only in 1D.

It is evident from the above that since cellular automata act on the same structure as memory channels, that they will be closely related. The relation will be made more evident throughout the rest of this chapter and in the following chapter also.

Definition 4.4.1 (Quantum cellular automaton)**.** Quantum cellular automata (QCAs) are local automorphisms of the quasi-local algebra $C : \mathfrak{Z} \mapsto \mathfrak{Z}$. By automorphism we mean that they preserve the algebraic structure of $\mathfrak{Z}$:

$$C(x \cdot y) = C(x) \cdot C(y), \tag{4.31}$$

for all $x, y \in \mathfrak{Z}$ and by locality[1] we understand that

$$C(\mathfrak{B}_{\Lambda_1}) \subset \mathfrak{A}_{\Lambda_2}, \tag{4.32}$$

where Λ_2 is finite for every finite Λ_1.

Due to the automorphic property if C is a QCA then also C^{-1} is also QCA. Hence quantum cellular automata are always reversible.

One example of a quantum cellular automaton we have already seen in the shift channel σ_x in Example 4.1.3. Since a QCA doesn't need to be causal σ_x is a QCA even for $k < 0$, when the causality condition is not met.

Example 4.4.2 (Clifford QCA)**.** A Clifford QCA is such QCA which maps tensor products of Pauli operators to tensor products of Pauli operators. As an example we take following (translationally invariant) qubit example defined by these local rules:

$$\begin{aligned} \alpha_C(X_j) &= X_{j-1}Z_jX_{j+1} \\ \alpha_C(Z_j) &= X_j. \end{aligned} \tag{4.33}$$

Note that $\alpha_C(Y_j)$ is already fixed by the automorphism property of α_C

$$\alpha_C(Y_j) = \alpha_C(iX_jZ_j) = i\alpha_C(X_j)\alpha_C(Z_j) = -X_{j-1}Y_jX_{j+1}, \tag{4.34}$$

It is also easy to check that $[\alpha_C(a_j), \alpha_C(a_k)] = 0$ for all $j \neq k$. α_C is not causal, however we can use the shift channel σ_1, and $\sigma_1 \circ \alpha_C$ will be a causal QCA. Since every QCA is local, this trick can be employed in every QCA to make it causal. Thus QCA's can be thought of as special case of memory channels.

QCA form a group in the sense that composition of two QCAs is a QCA and since QCA is an automorphism there exists an inverse which is also automorphism and is also local.

[1]The notion of locality is in some papers referred to as "causality" since it creates a cone where the information can propagate after several time steps. However since the term causality is already used for causal processes and moreover we would like to speak also about causal QCA we stick to the term local. This problem arises only because the natural time in QCA is usually the time step used to transform the whole cell structure. But in this work we are interested only in one such time step, and introduce the time parameter *within* the cell structure, in the ordering of cells. This is completely natural from the point of memory channels and causal processes. If this footnote just confuses you and otherwise the notion of causality and locality in this work are clear to you, just ignore this footnote.

4.4.1 Index of a QCA

As was shown in [25] for every QCA α we can calculate index of a QCA, $\mathrm{ind}(\alpha)$. This can be computed from a sufficiently large portion of the automaton (it can be computed locally) and is constant along the chain (it can be computed locally on any sufficiently large portion). This quantity also respects the group structure of QCAs, $\mathrm{ind}(\alpha_1 \circ \alpha_2) = \mathrm{ind}(\alpha_1)\mathrm{ind}(\alpha_2)$. In order to define and calculate the index we need to fix few things first.

Definition 4.4.3 (Nearest neighbor automaton)**.** Nearest neighbor automaton (nn-QCA) α is such that

$$\alpha(\mathfrak{B}_j) \in \mathfrak{A}_{[j-1,j+1]} \tag{4.35}$$

for every $j \in \mathbb{Z}$.

Any QCA can be made nn-QCA by suitably regrouping the cells into larger disjoint blocks.

Next we will define support algebras. Support algebras should somehow quantify how much of one algebra is contained in another. The definition is due to Zanardi [70], where it is called interaction algebra. However we will use the term used in [25], since it better reflects the purpose of the object in this setting.

Definition 4.4.4 (Supportalgebra)**.** Let $\mathfrak{A} \subset \mathfrak{X}_1 \otimes \mathfrak{X}_2$ be a subalgebra of finite C^*-algebras $\mathfrak{X}_1$ and $\mathfrak{X}_2$. The smallest C^*-subalgebra $\mathfrak{C}_1 \subset \mathfrak{X}_1$ such that $\mathfrak{A} \subset \mathfrak{C}_1 \otimes \mathfrak{X}_2$ is called a *support algebra* of $\mathfrak{A}$ on $\mathfrak{X}_1$ and denoted $\mathfrak{C}_1 = S(\mathfrak{A}, \mathfrak{X}_1)$.

Given a basis $\{x_\mu\} \subset \mathfrak{X}_2$, so that every $a \in \mathfrak{A}$ has a unique expansion $a = \sum_\mu a_\mu \otimes x_\mu$ with $a_\mu \in \mathfrak{X}_1$, the $S(\mathfrak{A}, \mathfrak{X}_1)$ is generated by all such elements a_μ. Now we can approach the definition of $\mathrm{ind}(\alpha)$ due to [25].

Definition 4.4.5 (Index of a QCA)**.** Let α be a nn-QCA on a cell structure $\mathfrak{A}_k$. Define

$$\mathfrak{R}_{2k} = S\big(\alpha(\mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}), \mathfrak{A}_{2k-1} \otimes \mathfrak{A}_{2k}\big) \tag{4.36}$$

$$\mathfrak{R}_{2k+1} = S\big(\alpha(\mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}), \mathfrak{A}_{2k+1} \otimes \mathfrak{A}_{2k+2}\big). \tag{4.37}$$

$\mathfrak{R}_k \equiv \mathcal{L}(\mathcal{R}_k)$ and $\dim \mathcal{R}_k = r(k)$. The index of α is the defined as

$$\mathrm{ind}(\alpha) = \frac{r(2k)}{d(2k)} = \frac{d(2k+1)}{r(2k+1)}, \tag{4.38}$$

where $d(k) = \dim \mathcal{A}_k$. Index of any QCA is hence a nonzero positive rational number.

From the definition it is not evident that $\mathfrak{R}_k \equiv \mathcal{L}(\mathcal{R}_k)$. The proof is taken form [25] and is included here to make the presentation self-contained. In following we will need also this lemma:

Lemma 4.4.6. *Let $\mathfrak{A}_1 \subset \mathfrak{X}_1 \otimes \mathfrak{X}_2$ be and $\mathfrak{A}_2 \subset \mathfrak{X}_2 \otimes \mathfrak{X}_3$ be subalgebras such that $\mathfrak{A}_1 \otimes \mathbb{I}_3$ and $\mathbb{I}_1 \otimes \mathfrak{A}_2$ commute in $\mathfrak{X}_1 \otimes \mathfrak{X}_2 \otimes \mathfrak{X}_3$. Then $S(\mathfrak{A}_1, \mathfrak{X}_2)$ and $S(\mathfrak{A}_2, \mathfrak{X}_2)$ commute in $\mathfrak{X}_2$.*

Proof. Pick bases $\{x_\mu\} \subset \mathfrak{X}_1$ and $\{b'_\nu\} \subset \mathfrak{X}_3$ and let $a \in \mathfrak{A}_1$ and $a' \in \mathfrak{A}_2$. We may uniquely expand $a = x_\mu \otimes a_\mu$ and $a' = a'_\nu \otimes x'_\nu$. By the assumption we have

$$0 = [a \otimes \mathbb{I}_3, \mathbb{I}_1 \otimes a'] = \sum_{\mu\nu} x_\mu \otimes [a_\mu, a'_\nu] \otimes x'_\nu. \tag{4.39}$$

Since $x_\mu \otimes x'_\nu$ create basis in $\mathfrak{X}_1 \otimes \mathfrak{X}_3$ the expansion is unique so we must have that $[a_\mu, a'_\nu] = 0$ $\forall$ μ,ν. This property transfers to the algebras generated by a_μ and a'_ν and hence also to aforementioned support algebras. □

The even support algebras $\mathfrak{R}_{2k}$ mark the information which flows to the left and the odd $\mathfrak{R}_{2k+1}$ the information which travels to the right. The $\mathfrak{R}_k$ commute because they are either localized in disjoint regions (as for example $\mathfrak{R}_{2k}$ and $\mathfrak{R}_{2k+1}$) or by the virtue of Lemma 4.4.6 and automorphism property of α. Clearly we know since α is nnQCA that

$$\alpha(\mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}) \subset \mathfrak{R}_{2k} \otimes \mathfrak{R}_{2k+1}. \tag{4.40}$$

Hence the $\mathfrak{R}_k$ together generate algebra which contains the whole algebra $\alpha(\mathfrak{Z})$ which is the same as $\mathfrak{Z}$ since α is automorphic. If any $\mathfrak{R}_k$ had a nontrivial center, i.e., if there would be an element $X \in \mathfrak{R}_k$ which commutes with the whole $\mathfrak{R}_k$ different from $\mathbb{I}$, this X would be also in the center of $\mathfrak{Z}$. However it is known that this center is trivial [5]. Therefore each $\mathfrak{R}_k$ has to have trivial center and is isomorphic to algebra of $r(k) \times r(k)$ complex matrices.

The full proof together with the definition of index and much more is in [25].

Example 4.4.7. Consider the channel σ_x defined in Example 4.1.3 which also happens to be a QCA. Since this is not a nnQCA we will need to regroup it first. Let $\mathfrak{B}^x_l = \mathfrak{B}_{l.x} \otimes \mathfrak{B}_{l.x+1} \otimes \ldots \otimes \mathfrak{B}_{[l.x+x-1]}$ be the new cell grouping ordered by $l \in \mathbb{Z}$. By definition of σ_x we have

$$\sigma_x(\lambda^{\mathcal{B}}_l) = \lambda^{\mathcal{A}}_{l-1}, \tag{4.41}$$

for a suitable basis $\{\lambda^{\mathcal{B}}_l\} \subset \mathfrak{B}^x_l$ and $\{\lambda^{\mathcal{A}}_l\} \subset \mathfrak{A}^x_l$ for all $l \in \mathbb{Z}$. This is now a nnQCA. We can proceed and calculate the index $\mathrm{ind}(\sigma_x)$. The support algebras are easy to calculate

$$\begin{aligned} \mathfrak{R}_{2l} = S\big(\sigma_x(\mathfrak{B}_{2l} \otimes \mathfrak{B}_{2l+1}), \mathfrak{A}_{2l-1} \otimes \mathfrak{A}_{2l}\big) &= \mathfrak{A}_{2l-1} \otimes \mathfrak{A}_{2l} \\ \mathfrak{R}_{2l+1} = S\big(\sigma_x(\mathfrak{B}_{2l} \otimes \mathfrak{B}_{2l+1}), \mathfrak{A}_{2l+1} \otimes \mathfrak{A}_{2l+2}\big) &= \mathbb{C}_1. \end{aligned} \tag{4.42}$$

From this follows that

$$\mathrm{ind}(\sigma_x) = \frac{d(2l-1)d(2l)}{d(2l)} = d(2l-1) \equiv d(l) = d^x, \tag{4.43}$$

where $d = \dim \mathcal{A}_k$ is the dimension of Hilbert space of the original cell and $x \in \mathbb{Z}$.

Chapter 5

Memory

In this chapter we will stress out the nature and strength of the memory effects in memory channels introduced in previous chapter. We will define forgetful memory channels and strictly forgetful memory channels [72], and connect them with quantum cellular automata. We will find out that the index of a QCA has the meaning of the minimal dimension of Hilbert space of memory system in appropriate collision model. This part of work was done in collaboration with group in Hannover under the supervision of Prof. Reinhard Werner. There have been several attempts to address this topic for the memory overhead in quantum convolutional codes [30] where the result is quite complicated and requires to find a longest path in a non-comutativity graph. In our case we just need to calculate the index of the QCA which is simpler to calculate. Finally we will address repeatability of channels in the quantum memory channel setting and stroboscopic simulation of evolution of open systems.

5.1 Forgetfulness

In the proof of structure theorem in Section 4.3 we needed to establish a channel initializer R which described the influence of the inputs from the remote past on the memory system. This influence is in generally nontrivial however for certain class of memory channels it becomes irrelevant in the long run. These memory channels should have the property that information that is localized far in the past will not affect the behavior of channel in present. Channels where the initializer R will become irrelevant are forgetful memory channels.

Definition 5.1.1 (Forgetful memory channel, Schrödinger picture)**.** Let $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ be a collision model of a quantum memory channel. Suppose $\rho_1, \rho_2 \in \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_{[1,n]})$ such that $\mathrm{Tr}_{\mathcal{M}}(\rho_1) = \mathrm{Tr}_{\mathcal{M}}(\rho_2)$. Then $S^*_{[1,n]}$ is forgetful iff

$$\lim_{n\to\infty} \| \mathrm{Tr}_{\mathcal{B}_{[1,n]}}[S^*_{[1,n]}(\rho_1 - \rho_2)] \|_{\mathrm{tr}} = 0, \tag{5.1}$$

for all ρ_1, ρ_2 such that $\mathrm{Tr}_{\mathcal{M}}(\rho_1) = \mathrm{Tr}_{\mathcal{M}}(\rho_2)$ and where $\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}$ is the trace-norm.

This only means that the state of memory after many collisions depends only on the state of inputs $\mathrm{Tr}_{\mathcal{M}}(\rho_i)$ and not on the initial state of the memory $\mathrm{Tr}_{\mathcal{B}_{[1,n]}}(\rho_i)$. The definition in Heisenberg picture makes this only more evident

Definition 5.1.2 (Forgetful memory channel, Heisenberg picture)**.** Let $S_k : \mathfrak{A}_k \otimes \mathfrak{M} \mapsto \mathfrak{M} \otimes \mathfrak{A}_k$ be a collision model of a quantum memory channel. Then let $\hat{S}_{[1,n]} : \mathfrak{M} \mapsto \mathfrak{M} \otimes \mathfrak{A}_{[1,n]}$ be its concatenation where the outputs are ignored $\hat{S}_{[1,n]}(m) := S_{[1,n]}(\mathbb{I}_{[1,n]} \otimes m)$ for all $m \in \mathfrak{M}$. Then S is forgetful iff there exists a sequence of quantum channels $\tilde{S}_{[1,n]} : \mathfrak{M} \mapsto \mathfrak{A}_{[1,n]}$ such that

$$\lim_{n\to\infty} \| \hat{S}_{[1,n]} - \mathbb{I}_{\mathcal{M}} \otimes \tilde{S}_{[1,n]} \|_{\mathrm{cb}} = 0. \tag{5.2}$$

It can be proven that forgetful channels are dense in the set of all memory channels, see [37]. The main idea is to add a infinitesimally small amount of white noise on the memory system and the channel will become forgetful. Such noise is present usually in real life applications hence generally the channels will be forgetful. For such channels coding theorems can be proven and channel capacities calculated. It is rather hard to prove that some memory channel is forgetful.

Channels for which the limit in (5.1) and (5.2) is attained for finite $n < \infty$ are called *strictly forgetful* memory channels or channels with finite depth of memory, where the depth of memory is the number of times you have to use it to completely forget the state of memory.

Remark 5.1.3. In the case of pure memory channels which are strictly forgetful with depth of memory δ, we can observe that the dimension of memory system and the dimension of the system of δ consequent input cells have to be co-divisible in order to be strictly forgetful. Let $m \in \mathcal{L}(\mathcal{M})$ have $\dim \mathcal{M} = d_{\mathcal{M}}$ distinct eigenvalues. Then the operator $\hat{S}_{[1,\delta]}(m) = S_{[1,\delta]}(m \otimes \mathbb{I}_{[1,\delta]})$ has the same eigenvalues with multiplicity $\dim \mathcal{A}_{[1,\delta]} = d_\delta$ because $S_{[1,\delta]}$ is unitary. Since we know that the channel has finite depth, we can write $S_{[1,\delta]}(m \otimes \mathbb{I}_{[1,\delta]}) = \mathbb{I}_{\mathcal{M}} \otimes \tilde{S}_{[1,n]}(m)$. The operator $\mathbb{I}_{\mathcal{M}} \otimes \tilde{S}_{[1,n]}(m)$ has d_δ eigenvalues with multiplicity $d_{\mathcal{M}}$ but also it has $d_{\mathcal{M}}$ eigenvalues with multiplicity d_δ. Hence $d_{\mathcal{M}}$ and d_δ have to be co-divisible.

Forgetfulness as defined, is very hard to prove. To illustrate some problems lets see next example.

Example 5.1.4 (Partial swap interaction)**.** Let us have a qubit-qubit translationally invariant pure collision model specified by unitary interaction $U_\phi : \mathcal{M} \otimes \mathcal{A} \mapsto \mathcal{M} \otimes \mathcal{B}$:

$$U_\phi = \cos\phi \mathbb{I} + i \sin\phi U_\sigma, \tag{5.3}$$

where U_σ is the same as in (4.8). For $\cos\phi = 0$, this is just a shift channel which is forgetful and for $\cos\phi = 1$ it is a non-interacting ideal channel (however not forgetful). It has been proven numerically in [59] for some interval of ϕ that partial swap is forgetful, the proof for the whole interval except $\phi = k\pi$, $k \in \mathbb{N}$ is not known. Intuition tells us that whenever $\sin\phi$ is nonzero a portion of information from memory is transferred to output and therefore eventually the memory will get reset after long enough time.

What we can say is that if we restrict our input signal to be factorized $\omega_{[1,n]} = \bigotimes_{k=1}^{n} \omega_k$ then the limit in equation (5.1) will be zero for any such $\omega_{[1,n]}$.

Let ξ_k be the state of memory before k-th collision. Then we have

$$\xi_{k+1} = C_k^*(\xi_k) = \mathrm{Tr}_k[U_\phi(\xi_k \otimes \omega_k)U_\phi^\dagger], \tag{5.4}$$

where $C^* : \mathcal{T}(\mathcal{M}) \mapsto \mathcal{T}(\mathcal{M})$ is the so called *concurrent* channel which describes the dynamics of memory

$$\xi_k = C^*_{k-1} \cdots C^*_1(\xi_1). \tag{5.5}$$

Limit in equation (5.1) now becomes

$$\lim_{n\to\infty} C^*_{n-1} \cdots C^*_1(\xi_1 - \xi'_1). \tag{5.6}$$

We would like to show that this limit is zero. First we will show that any C^*_k is a strict contraction of the Bloch ball. We will show this by contradiction. Since every channel is a continuous mapping of a convex body into itself, it has to have a fixed point. Due to Banach fixed point theorem, for contractive channels this fixed point is unique. Thus a mapping that is not contractive has to have at least two fixed points. Due to linearity of channels also any linear combination of these two points is a fixed point, making it a line of fixed points in the vector space of all states. Hence there have to exist at least two states from the border of physical state space which are fixed points of a not contractive channel. In qubit case this means that these two points are pure. Let denote them ξ^a and ξ^b.

Looking back on partial swap interaction we must find that there has to be such ω_k that

$$\begin{aligned} C^*_k(\xi^a) = \xi^a \Rightarrow U_\phi(\xi^a \otimes \omega_n)U^\dagger_\phi = \xi^a \otimes (W_a \omega_k W^\dagger_a) \\ C^*_k(\xi^b) = \xi^b \Rightarrow U_\phi(\xi^b \otimes \omega_n)U^\dagger_\phi = \xi^a \otimes (W_b \omega_k W^\dagger_b), \end{aligned} \tag{5.7}$$

for some qubit unitaries W_a, W_b. Since partial swap commutes with $V \otimes V$ where V is arbitrary qubit unitary we may fix $\xi^a = |0\rangle_\mathcal{M}$.

The input state ω_k can be also taken pure. If it would be mixed we could find at least two pure states (eigenstates of ω_k) with the needed property due to linearity. There is only one pure state $\omega_k = |0\rangle_k$ such that $C^*_k(|0\rangle_\mathcal{M}) = |0\rangle_\mathcal{M}$. However for this input state $|0\rangle_k$ the concurrent channel C^*_k has only one fixed point $|0\rangle_\mathcal{M}$, and therefore C^*_k is contractive for any ω_k.

Since the dynamics of memory is then a composition of contractive channels the limit in (5.6) is zero. Similarly we could do this proof for any input sequence $\omega_{[1,n]}$ such that it would be composed of finite factorized parts of length l: $\omega_{[1,n]} = \bigotimes_{k=0}^{n/l} \omega_{[k.l+1,k.l+l-1]}$. The concatenation of partial swaps then commutes with $V^{\otimes l+1}$ hence we can again fix $\xi^a = |0\rangle_\mathcal{M}$. Similarly we will come to the conclusion that $\omega_{[k.l+1,k.l+l-1]} = \bigotimes_{x=k.l+1}^{k.l+l-1} |0\rangle$. However again C^*_k induced by this input sequence is contractive and we come to the same conclusion that after infinitely long input sequence the memory will be set to a fixed state independent of the initial state, dependent only on the input sequence.

The conclusion we can draw from this is that the computationally challenging part in proving forgetfulness of partial swap lies in sequences with infinitely long correlations. We might be well satisfied with the constraints on input sequence being factorized, and we do not need this memory channel to be forgetful in order to reset memory system. The long correlations however do indeed change the dynamics of memory as can be seen on Figure 5.1.

Forgetfulness describes the ability of the memory channel to completely reset the state of memory. If we are interested only in the input-output relations, this might be a too strict or

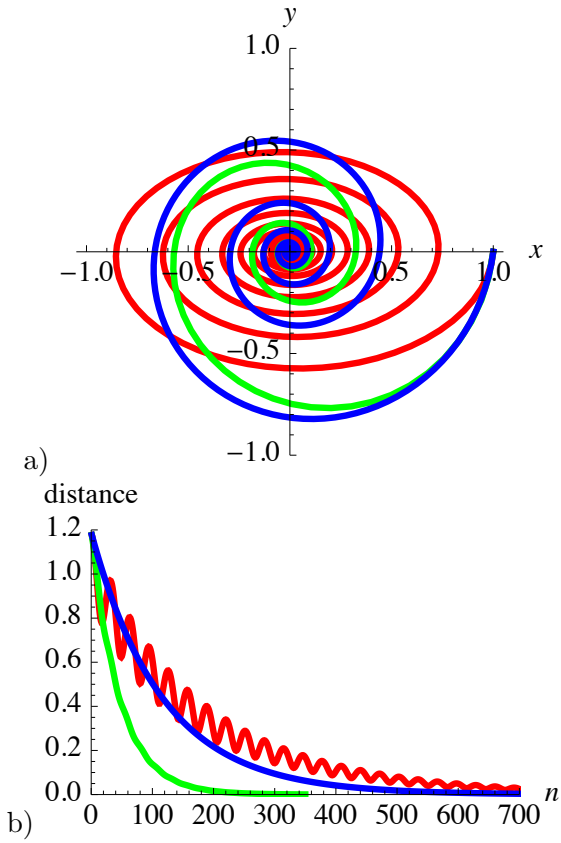


Figure 5.1: a) The dynamics of memory system of a memory channel with partial swap interaction $U_{\phi=0.1}$ for differently structured input sequences. The blue spiral corresponds to factorized inputs in state $\omega_k = 0.81|0\rangle\langle 0| + 0.19|1\rangle\langle 1|$. The green curve is corresponds to inputs entangled in pairs: $\omega_{[2k,2k+1]} = 0.9|00\rangle + \sqrt{0.19}|11\rangle$ and the red curve corresponds to the input sequence fully correlated across the chain: $\omega_{[1,n]} = 0.9|0\ldots 0\rangle + \sqrt{0.19}|1\ldots 1\rangle$. These sequences were chosen such that each local subsystem of input sequence is in the *same state*, hence the difference in dynamics is solely due to correlations. b) The distance from the final state where the dynamics is converging is plotted. We can see that the correlated sequence exhibits recurrences and converges more slowly compared to the factorized inputs.

counterintuitive approach. Memory channels where there is no interaction with memory, hence no memory effects are present, but the memory system is still included in the description of model, will be non-forgetful. Eventually any forgetful channel can become non-forgetful if we are "creative" enough to include some noninteracting memory into the model. We might relax the definition of forgetfulness in following way.

Definition 5.1.5 (I-O forgetfulness)**.** Let $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ be a collision model of a quantum memory channel. Suppose $\rho_1, \rho_2 \in \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_{[1,n+m]})$ such that $\mathrm{Tr}_{\mathcal{M}}(\rho_1) = \mathrm{Tr}_{\mathcal{M}}(\rho_2) = \omega_{[1,n+m]}$ for some finite $m \in \mathbb{N}$. Then $S^*_{[1,n+m]}$ is forgetful iff

$$\lim_{n\mapsto\infty} \| \, T_1^*(\omega_{[n,n+m]}) - T_2^*(\omega_{[n,n+m]}) \, \|_{\mathrm{tr}} = 0 \tag{5.8}$$

for every finite $m \in \mathbb{N}$ where $T_j^*(\omega_{[n,n+m]}) = \mathrm{Tr}_{\mathcal{M}\otimes\mathcal{B}_{[1,n]}}[S^*_{[1,n+m]}(\rho_j)]$. We can speak about *strict* I-O forgetfulness, if this limit is attained for every $n \geq \delta$, where $0 \leq \delta < \infty$.

In Heisenberg picture the definition reflects only how much the initial memory state affects the transformation in far future. I-O forgetfulness is weaker than forgetfulness in the sense that the set of all forgetful channels S_{ff} is a subset of I-O forgetful channels $S_{\mathrm{I-O}}$. Following example is here to show that the difference between I-O forgetfulness and normal forgetfulness is not only caused by the "noninteracting ancilla".

The memory channel from Example 4.2.2 is not I-O forgetful, since the transformation explicitly depends on diagonal elements of memory state. However we can add some noise to the memory system to make it I-O forgetful but not forgetful.

Example 5.1.6 (I-O Forgetful but not forgetful)**.** Let us have a translationally invariant qubit-qubit collision model with collision $S^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}) \mapsto \mathcal{T}(\mathcal{B} \otimes \mathcal{M})$

$$S^*(m \otimes \omega) = E^* \otimes \mathbb{I}\big(U_{\mathrm{cnot}}(m \otimes \omega){U_{\mathrm{cnot}}}^{\dagger}\big) \tag{5.9}$$

where E^* is the contraction to the x-axis from Example 2.3.8 and U_{cnot} is also defined there. It is easy to see that this channel is not forgetful. The states $|+\rangle_{\mathcal{M}} \otimes |+\rangle$ and $|-\rangle_{\mathcal{M}} \otimes |+\rangle$ are fixed points of channel S^*. Thus the x component of state in $\mathcal{M}$ will stay the same if the input sequence is $\omega_{[1,n]} = |+\ldots+\rangle$.

It is a bit tedious to prove that it is I-O forgetful, the idea is that U_{cnot} takes into account only the diagonal elements of memory and that E^* resets them to $\frac{1}{2}$ irrespective of the original state, hence the effect of initial memory state will be zero after first use and all subsequent uses will be independent of the initial state.

Here is the detailed proof. Lets write the joint state of memory and inputs $\rho \in \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_{1,n})$ as

$$\rho = |i\rangle_{\mathcal{M}}\langle j| \otimes \Omega_{ij}. \tag{5.10}$$

This way the state of input sequence is $\mathrm{Tr}_{\mathcal{M}}\rho = \Omega_{00} + \Omega_{11} = \omega_{[1,n]}$. The first collision will transform ρ into

$$\begin{aligned} S_1^*(\rho) &= E^* \otimes \mathbb{I}_{[1,n]}[(U_{\mathrm{cnot}} \otimes \mathbb{I}_{[2,n]})\rho({U_{\mathrm{cnot}}}^{\dagger} \otimes \mathbb{I}_{[2,n]})] \\ &= E^*(|0\rangle_{\mathcal{M}}\langle 0|) \otimes \Omega_{00} + E^*(|0\rangle_{\mathcal{M}}\langle 1|) \otimes (\Omega_{01}X_1) \\ &\quad + E^*(|1\rangle_{\mathcal{M}}\langle 0|) \otimes (X_1\Omega_{10}) + E^*(|1\rangle_{\mathcal{M}}\langle 1|) \otimes (X_1\Omega_{11}X_1) \\ &= \frac{1}{2}[\mathbb{I}_{\mathcal{M}} \otimes (\Omega_{00} + X_1\Omega_{11}X_1) + X_{\mathcal{M}} \otimes (\Omega_{01}X_1 + X_1\Omega_{10}), \end{aligned} \tag{5.11}$$

because $E^*(|i\rangle\langle i|) = \frac{1}{2}\mathbb{I}_{\mathcal{M}}$ and $E^*(|0\rangle\langle 1|) = E^*(|1\rangle\langle 0|) = \frac{1}{2}X_{\mathcal{M}}$. The output state clearly depends on the initial memory state. If the initial memory state was $|0\rangle_{\mathcal{M}}\langle 0|$, which would mean that $\Omega_{11} = 0$, the transformation would read

$$T^*(\omega_{[1,n]}) = \omega_{[1,n]}, \tag{5.12}$$

however if the initial state was $|1\rangle_{\mathcal{M}}\langle 1|$, then $\Omega_{00} = 0$, the transformation would be

$$T^*(\omega_{[1,n]}) = X_1\omega_{[1,n]}X_1. \tag{5.13}$$

As we can already see the Ω_{01} and Ω_{10} will never enter the channel T^* because they will be always attached to a traceless operator on memory system, we can ignore them ($\Omega_{01} = \Omega_{10} = 0$). Now if we look on the composite state of memory and inputs sequence except first input we find that

$$\begin{aligned}
\mathrm{Tr}_{\mathcal{B}_1}[S_1^*(\rho)] &= \frac{1}{2}\mathrm{Tr}_{\mathcal{B}_1}[\mathbb{I}_{\mathcal{M}} \otimes (\Omega_{00} + X_1\Omega_{11}X_1)] \\
&= \frac{1}{2}\mathbb{I}_{\mathcal{M}} \otimes (\mathrm{Tr}_{\mathcal{B}_1}[\Omega_{00}] + \mathrm{Tr}_{\mathcal{B}_1}[X_1\Omega_{11}X_1]) \\
&= \frac{1}{2}\mathbb{I}_{\mathcal{M}} \otimes (\mathrm{Tr}_{\mathcal{B}_1}[\Omega_{00} + \Omega_{11}]),
\end{aligned} \tag{5.14}$$

where the last equality is due to invariance of partial trace under unitary conjugation on subsystem which is being traced over, a consequence of invariance of trace under cyclic permutations. This state is independent of the initial memory and hence the initial memory state will have no effects after the first collision. Thus the whole memory channel is I-O forgetful, even strictly I-O forgetful with $\delta = 1$.

Lemma 5.1.7 (Equivalence class of collision models)**.** *Let $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ be a collision model of a quantum memory channel. Then memory channel with collisions*

$$\hat{S}_k^* = (\mathbb{I}_k \otimes W^\dagger)S_k^*(W \otimes \mathbb{I}_k), \tag{5.15}$$

where $W : \mathcal{M} \mapsto \mathcal{M}$ is an arbitrary unitary and $\mathbb{I}_k$ is the identity operation on k-th subsystem, has the same I-O relation as the former memory channel.

Proof. This directly follows from the concatenation properties of the collision model. The concatenation of $\hat{S}_k^*$ will be

$$\begin{aligned}
&\hat{S}_{[1,n]}^* = (\mathbb{I}_{[1,n]} \otimes W^\dagger)(\mathbb{I}_{[1,n-1]} \otimes S_n^*)(\mathbb{I}_{[1,n-1]} \otimes W \otimes \mathbb{I}_n) \\
&\cdots(\mathbb{I}_{[1,2]} \otimes W^\dagger \otimes \mathbb{I}_{[3,n]})(\mathbb{I}_1 \otimes S_2^* \otimes \mathbb{I}_{[3,n]})(\mathbb{I}_1 \otimes W \otimes \mathbb{I}_{[2,n]}) \\
&(\mathbb{I}_1 \otimes W^\dagger \otimes \mathbb{I}_{[2,n]})(S_1^* \otimes \mathbb{I}_{[2,n]})(W \otimes \mathbb{I}_{[1,n]}) = \\
&(\mathbb{I}_{[1,n]} \otimes W^\dagger)(\mathbb{I}_{[1,n-1]} \otimes S_n^*)\cdots(\mathbb{I}_1 \otimes S_2^* \otimes \mathbb{I}_{[3,n]})(S_1^* \otimes \mathbb{I}_{[2,n]})(W \otimes \mathbb{I}_{[1,n]}).
\end{aligned} \tag{5.16}$$

Since the partial trace is invariant under unitary conjugations of the system being traced over, the trace over memory system will be tha same as for plain S_k^* without the unitary W. □

5.2 Finite depth memory channels

We will add another slightly different notion of forgetfulness. The notion of memory depth was already used for strictly forgetful channels, where it marks the maximal number of uses of the memory channel needed to forget. Now we will formalize memory channels with finite depth as those possessing this property, but we will require only I-O strict forgetfulness and require it only for factorized input sequences. This will also implicitly bind us to the Schrödinger picture, where the factorization of input sequence is more natural. The reason is that when using such memory channel for communication or estimation, it is much easier to construct uncorrelated sequences.

Definition 5.2.1 (Finite depth)**.** Let $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ be a collision model of a quantum memory channel. Let the input sequence restrict to $\omega_{[1,n]} = \bigotimes_{k=1}^{n} \omega_k$. Assume two different initial states of memory $\xi_1, \xi_2 \in \mathcal{T}(\mathcal{M})$ Then $S_{[1,n]}^*$ has finite depth δ if

$$\| T_1^*(\omega_{[\delta+1,n]}) - T_2^*(\omega_{[\delta+1,n]}) \|_{\mathrm{tr}} = 0 \tag{5.17}$$

for every finite $n > \delta$ where $T_j^*(\omega_{[\delta+1,n]}) = \mathrm{Tr}_{\mathcal{M}\otimes\mathcal{B}_{[1,\delta]}}[S_{[1,n]}^*(\xi_j \otimes \omega_{[1,n]})]$.

Strictly I-O forgetful, hence also strictly forgetful, memory channels have automatically finite depth. Note that finite depth is defined only on factorized inputs. However, next lemma shows that this can be trivialy shifted to any state of inputs and memory, hence finite depth also implies strict I-O forgetfulness.

Lemma 5.2.2. *If a memory channel has finite depth, then it is strictly I-O forgetful.*

Proof. Let us have a collision model ${S^*}_{[1,n]} : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_{[1,n]}) \mapsto \mathcal{T}(\mathcal{B}_{[1,n]} \otimes \mathcal{M})$ such that it has finite depth δ. Then

$$\mathrm{Tr}[S_{[1,\delta+1+m]}^*\big((\xi_1 - \xi_2) \otimes \omega_1 \otimes \cdots \otimes \omega_{\delta+1+m}\big)(\mathbb{I}_{[1,n]} \otimes b_{[\delta+1,\delta+1+m]} \otimes \mathbb{I}_{\mathcal{M}})] = 0 \tag{5.18}$$

holds for every $\xi_1, \xi_2 \in \mathcal{S}(\mathcal{M})$, $\omega_k \in \mathcal{S}(\mathcal{A}_k)$, $b_{[\delta+1,\delta+1+m]} \in \mathfrak{B}_{[\delta+1,\delta+1+m]}$ and $m \in \mathbb{N}$. Any $\rho_1 \in \mathcal{S}(\mathcal{M} \otimes \mathcal{A}_{[1,\delta+1+m]})$ can be written as a (not necessary convex or positive) sum of factorized sequences $\xi_1 \otimes \omega_1 \otimes \cdots \otimes \omega_{\delta+1+m}$ since they provide an over-complete operator basis. Due to linearity it has to be than true that

$$\mathrm{Tr}[S_{[1,\delta+1+m]}^*\big(\rho_1 - \rho_2\big)(\mathbb{I}_{[1,n]} \otimes b_{[\delta+1,\delta+1+m]} \otimes \mathbb{I}_{\mathcal{M}})] = 0 \tag{5.19}$$

for any $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{M} \otimes \mathcal{A}_{[1,\delta+1+m]})$ such that $\mathrm{Tr}_{\mathcal{M}}\rho_1 = \mathrm{Tr}_{\mathcal{M}}\rho_2$. □

This is a very interesting result since it greatly simplifies the conditions for for strict I-O forgetfulness. It is enough to check only all factorized sequences.

As we have seen in previous section, the memory might be composed of some part which does not affect the transformation on inputs, we will call this the irrelevant part of memory. Let us fix a collision model $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ with the initial state of memory $\xi = 1/\sqrt{d_{\mathcal{M}}}(\tau_0 + \vec{m}\vec{\tau})$ and constrain the input sequence to factorized states, $\omega_{[1,n]} = \bigotimes_{k=1}^{n} \omega_k$. Then we can define the irrelevant subspace.

Definition 5.2.3 (Irrelevant degrees of freedom)**.** A traceless operator $\varsigma = \vec{s}\vec{\tau}$ is called *irrelevant* if for every $n \in \mathbb{N}$

$$\mathrm{Tr}_{\mathcal{M}}[S^*{}_{[1,n]}(\varsigma \otimes \omega_{[1,n]})] = O \tag{5.20}$$

holds. The space that all such ς span is called the *irrelevant subspace* of memory and we will denote it by I_{S^*}.

One can see that any two initial memory states ξ_1, ξ_2 such that $\xi_1 - \xi_2 \in I_{S^*}$ have the same I-O relation, ie. produce the same channel $T^*(\omega_{[1,n]})$. Note that the space I_{S^*} can be only spanned by traceless operators because for any operator with nonzero trace the equation (5.20) does not hold because of trace preserving property of S^*.

Since the input is factorized, the memory evolves through composition of series of concurrent channels (see Example 5.1.4) $C_k^* : \mathcal{T}(\mathcal{M}) \mapsto \mathcal{T}(\mathcal{M})$

$$\xi \mapsto C_n^* \cdots C_1^*(\xi). \tag{5.21}$$

The finite depth property can be then expressed as

$$C_\delta^* \cdots C_1^*(\xi_1 - \xi_2) \in I_{S^*} \tag{5.22}$$

for all $\xi_1, \xi_2 \in \mathcal{S}(\mathcal{M})$.

For every memory channel with finite depth δ, inputs separated by δ uses evolve under factorized transformation

$$T^*(\omega_k \otimes \omega_{k+\delta+1}) = T^*(\omega_k) \otimes T^*(\omega_{k+\delta+1}). \tag{5.23}$$

This has to be so because neither initial state of memory nor ω_k can have an effect on the transformation. Let us move to the easiest case study.

5.2.1 Qubit-qubit case study

In this case study, we will consider only pure, translationally invariant memory channels with qubit memory interacting with qubit subsystem, $\dim \mathcal{M} = \dim \mathcal{A} = \dim \mathcal{B} = 2$. The collision will be described by an unitary interaction FU, where F is the swap as usual. The U can be parametrized in following way [34]:

$$U = (W_2 \otimes V_2)D(W_1 \otimes V_1), \tag{5.24}$$

where $W_i : \mathcal{M} \mapsto \mathcal{M}$ and $V_i : \mathcal{A} \mapsto \mathcal{A}$ are unitaries on appropriate Hilbert spaces and D is 2-qubit unitary of a special form

$$D = e^{i\sum_{k=1}^{3} \frac{1}{2}\alpha_k \sigma_k \otimes \sigma_k}, \tag{5.25}$$

where $\sigma_1 = X$, $\sigma_2 = Y$ and $\sigma_3 = Z$ are the Pauli matrices and $\alpha_k \in \mathbb{R}$. Since $\sigma_k \otimes \sigma_k$ commute the D can be written as

$$D = \prod_{k=1}^{3} (\cos(\alpha_k/2)\mathbb{I} + i \sin(\alpha_k/2)\sigma_k \otimes \sigma_k). \tag{5.26}$$

This is a very nice parametrization for our task since it separates the local transformations from the interacting part. We can substitute W_1 for identity for our purposes, because of the I-O invariance of collision models from Lemma 5.1.7 by suitably choosing $W = W_1^\dagger$. Since W_2 can be any unitary this will not change the description.

Lets fix initial memory in state $\xi \in \mathcal{S}(\mathcal{M}) =: \xi_1$. We restrict our inputs to be of factorized form, $\omega_{[1,n]} = \bigotimes_{k=1}^n \omega_k$ uncorrelated to ξ. Firstly concentrate our attention to identifying the relevant and irrelevant subspace. The transformation on first input will be

$$\begin{aligned} T^*(\omega_1) &= \mathrm{Tr}_{\mathcal{M}}[U(\xi_1 \otimes \omega_1)U^\dagger] = V_2 \mathrm{Tr}_{\mathcal{M}}[D\big(\xi_1 \otimes (V_1\omega_1 V_1^\dagger)\big)D^\dagger]V_2^\dagger \\ &=: V_2 \hat{T}^*(V_1\omega_1 V_1^\dagger)D^\dagger]V_2^\dagger. \end{aligned} \tag{5.27}$$

The irrelevant subspace is fully determined by $\hat{T}^*$ since only there interaction with memory system occurs. Lets fix the operator basis for qubit as usual $\tau = 1/\sqrt{2}(\mathbb{I}, X, Y, Z)$. Assume that $\xi_1 = 1/\sqrt{2}(\tau_0 + \vec{m}\vec{\tau})$. Then channel $\hat{T}^*$ is

$$\hat{T}^*(\omega_1) = \mathrm{Tr}_{\mathcal{M}}[D(\xi_1 \otimes \omega_1)D^\dagger], \tag{5.28}$$

what in vector representation looks $A(\hat{T}^*)_{ij} = \mathrm{Tr}[D(\xi_1 \otimes \tau_j)D^\dagger(\mathbb{I} \otimes \tau_i)]$:

$$A(\hat{T}^*) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ m_1 s_2 s_3 & c_2 c_3 & m_3 c_2 s_3 & -m_2 s_2 c_3 \\ m_2 s_1 s_3 & -m_3 c_1 s_3 & c_1 c_3 & m_1 s_1 c_3 \\ m_3 s_1 s_2 & m_2 c_1 s_2 & -m_1 s_1 c_2 & c_1 c_2 \end{pmatrix}, \tag{5.29}$$

where $c_i = \cos\alpha_i$ and $s_i = \sin\alpha_i$. We see that for any nonzero operator $\varsigma = \vec{m}\vec{\tau}$ to be irrelevant we need $s_i = 0$ for all $i = 1, 2, 3$. This implies that $\alpha_i = k_i\pi$ where $k_i \in \mathbb{N}$. However if this holds then D is a product of two equal Pauli matrices, i.e. no interaction at all. We conclude that the irrelevant subspace is trivial unless the whole interaction is of factorized form. Then the irrelevant subspace is maximal, meaning that the relevant subspace is spanned only by identity.

Now we will move on to the concurrent channel. Lets denote

$$\xi_{k+1} := C_k^* \cdots C_1^*(\xi_1) = C_k^*(\xi_k) \tag{5.30}$$

the state of memory after k-th collision, where C_k^* are the concurrent channels

$$C_k^*(\xi_k) = \mathrm{Tr}_k[U(\xi_k \otimes \omega_k)U^\dagger]. \tag{5.31}$$

After applying the parametrization (5.24) we get that

$$C_k^*(\xi_k) = W_2 \mathrm{Tr}_k[D\big(\xi_k \otimes (V_1\omega_k V_1^\dagger)\big)D^\dagger]W_2^\dagger, \tag{5.32}$$

where we see that V_2 doesn't affect the transformation on memory. Since the irrelevant subspace is trivial we require

$$C_\delta^* \cdots C_1^*(\varsigma) = O \tag{5.33}$$

for any $\varsigma = \vec{m}\vec{\tau}$ a memory channel of depth δ. In vector representation this means that if we define

$$A(C_k^*) = \begin{pmatrix} 1 & 0 \\ \vec{c}_k & \widetilde{C}_k \end{pmatrix} \tag{5.34}$$

then

$$\vec{m} \mapsto \widetilde{C}_\delta \cdots \widetilde{C}_1 \vec{m} + \sum_{k=1}^{\delta-1} \widetilde{C}_\delta \cdots \widetilde{C}_{k+1} \vec{c}_k + \vec{c}_\delta. \tag{5.35}$$

If we need the equation (5.33) to hold we need that

$$\widetilde{C}_\delta \cdots \widetilde{C}_1 = O \tag{5.36}$$

because only then the right hand side of (5.35) will be independent of $\vec{m}$. A necessary condition for that is that any $\widetilde{C}_k$ has to be singular, since we can have input sequences $\omega_{[1,n]} = \omega^{\otimes n}$ with arbitrary ω. Any $\widetilde{C}_k$ looks in vector representation as

$$\widetilde{C}_k = R_2 \begin{pmatrix} c_2c_3 & w_{k,3}c_2s_3 & -w_{k,2}s_2c_3 \\ -w_{k,3}c_1s_3 & c_1c_3 & w_{k,1}s_1c_3 \\ w_{k,2}c_1s_2 & -w_{k,1}s_1c_2 & c_1c_2 \end{pmatrix}, \tag{5.37}$$

where we have used that $V_1\omega_k V_1^\dagger = 1/\sqrt{2}(\mathbb{I} + \vec{w}_k\vec{\tau})$ and $R_2 = A(W_2)$ is the rotation due to unitary W_2. Since $\det(\widetilde{C}_k) = (c_1c_2c_3)^2 + w_{k,1}(s_1c_2c_3)^2 + w_{k,2}(c_1s_2c_3)^2 + w_{k,3}(c_1c_2s_3)^2$, this matrix is singular for arbitrary $\vec{w}_k$ only when $c_j = 0$ for two different j. Lets say that $c_1 = c_2 = 0$ and $c_3 \neq 0$, then

$$\begin{aligned} \widetilde{C}_k &= R_2 \begin{pmatrix} 0 & 0 & \pm w_{k,2}c_3 \\ 0 & 0 & \pm w_{k,1}c_3 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \pm \begin{pmatrix} 0 & 0 & w_{k,1}c_3R_{2,12} \pm w_{k,2}c_3R_{2,11} \\ 0 & 0 & w_{k,1}c_3R_{2,22} \pm w_{k,2}c_3R_{2,21} \\ 0 & 0 & w_{k,1}c_3R_{2,32} \pm w_{k,2}c_3R_{2,31} \end{pmatrix}. \end{aligned} \tag{5.38}$$

Under further investigation we find that condition (5.36) can hold for arbitrary factorized input sequence only if $R_{2,32} = R_{2,31} = 0$. Since R_2 is rotation this also implies that $R_{2,33} = \pm 1$ and

$$R_2 = \begin{pmatrix} S & 0 \\ 0 & \pm 1 \end{pmatrix}, \tag{5.39}$$

where S is rotation in the xy-plane. Such unitaries commute with Z thus

$$W_2 = e^{\beta Z}. \tag{5.40}$$

Furthermore the depth of such memory channel is necessary $\delta = 2$ since we can check that

$$\widetilde{C}_{k+1}\widetilde{C}_k = O, \tag{5.41}$$

for any factorized input sequence. So if $c_1 = c_2 = 0$ and $c_3 \neq 0$ and $W_2^z = e^{\beta Z}$ then any $U_{\delta=2}^z$ of form

$$U_{\delta=2}^z = (W_2^z \otimes V_2)D(\mathbb{I}_{\mathcal{M}} \otimes V_1), \tag{5.42}$$

with arbitrary V_k defines a collision model with memory depth $\delta = 2$: $S_k^*(\xi \otimes \omega_k) = FU_{\delta=2}(\xi \otimes \omega_k)U_{\delta=2}^\dagger F$. Similarly you will get the same thing for $U_{\delta=2}^y$ with $c_1 = c_3 = 0$ and $c_2 \neq 0$ and $W_2^y = e^{\beta Y}$ and $U_{\delta=2}^x$ with $c_2 = c_3 = 0$ and $c_1 \neq 0$ and $W_2^x = e^{\beta X}$. Furthermore it is easy to check that if an two qubit unitary U corresponds to a memory channel with depth of memory 2 then also $U^\dagger$ has depth of memory 2. Lets say that U is of the $U_{\delta=2}^z$ type, then

$$U^\dagger = (\mathbb{I}_{\mathcal{M}} \otimes V_1^\dagger)D^\dagger(W_2^{z\dagger} \otimes V_2^\dagger). \tag{5.43}$$

Since the I-O relation is invariant under conjugation of memory, we are able to move the inverse of W_2 back to the left side with suitable conjugation. Also the inverses of V_i can be arbitrary and hence are of no interest now, so we can omit the daggers on them. Thus

$$U^\dagger = (W_2^{z\dagger} \otimes V_1)D^\dagger(\mathbb{I}_{\mathcal{M}} \otimes V_2). \tag{5.44}$$

We can also ignore the dagger on W_2^z since it does not change the type of the unitary, it only changes the sign in front of β. If D is such that $c_1 = c_2 = 0$ then also $D^\dagger$ will fulfill this condition, because the dagger operation only introduces sign changes in front of α_i and this won't affect the cosines in condition, thus

$$U^\dagger = (W_2^z \otimes V_1)D(\mathbb{I}_{\mathcal{M}} \otimes V_2), \tag{5.45}$$

which is of depth 2 again.

We will also get finite depth if we set $c_1 = c_2 = c_3 = 0$, then $\widetilde{C}_k = O$ automatically and $\delta = 1$. It follows then that

$$U_{\delta=1} = (W_2 \otimes V_2)F(W_1 \otimes V_1), \tag{5.46}$$

with arbitrary W_k and V_k where F is the two qubit swap unitary. Trivially since the local unitaries are arbitrary and F self-adjoint, $U_{\delta=1}^\dagger$ has also memory depth 1.

Interesting thing is that there is no translationally invariant pure memory channel in this situation with finite depth higher than 2. If we would drop the translational invariance this would stop to hold. Imagine a pure qubit-qubit memory channel $S_k^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}_k) \mapsto \mathcal{T}(\mathcal{B}_k \otimes \mathcal{M})$ such that every

$$S_{m\Delta}^* = \mathbb{I}, \tag{5.47}$$

for fixed Δ and $m \in \mathbb{N}$, meaning that every Δ step you would swap the entire memory with your system. Naturally, such memory channel has depth at least $\delta \leq \Delta$ for any $\Delta < \infty$ and is even strictly forgetful. In translationally invariant case the intuition tells that you have to erase at least one degree of freedom from the relevant subspace of memory at each collision. Thus the depth should be bounded by the number of linearly independent traceless operators

$d^2_{\mathcal{M}} - 1$, where $d_{\mathcal{M}} = \dim \mathcal{M}$. Taking this to consequences we should then ask why there is no $\delta = 3$ in this qubit case? The answer is that such maps, which erase only one degree of freedom are not completely positive for qubits, see Example 2.3.3. If we look onto classical case we have discrete set of possible interactions - permutations of a set of four elements, which can be represented by 24, 4×4 permutation matrices, and states are diagonal, hence described by only one traceless parameter $\xi = 1/2(\mathbb{I} + pZ)$. It can be easily checked that these interactions can be grouped to three groups. First group would consist of factorized interactions - independent evolution of memory and input. This group has $\delta = 0$. Second group would be the swap like interactions, with $\delta = 1$ and third group are control unitary interactions (see Example 4.2.2), where the memory is control and the input target or the other way round. Such interactions have not finite depth. Thus the quantum case is slightly richer.

5.3 QCA as a finite depth pure memory channel

The structure of memory channels and QCAs is deeply connected. With appropriate shift channel σ_x, any QCA can be made causal as was already noted in Example 4.4.2. Thus every QCA is a memory channel up to a shift. Because of locality, any QCA is also a memory channel with finite depth. However the structure of such memory channel is not obvious. Specifically one does not know the size of memory environment we would need to use for a memory channel implementation of QCA, which represents the resource overhead of the implementation. Even the Stinespring dilation cannot help us much even to found a bound, since the transformed system is infinite. In this section we will show that this memory system has dimensions equal to the index of a considered (causal) automaton. Hence as such is constant along the line of inputs and more importantly finite. This result was obtained during a scientific visit to a group in Hannover and is still in preparation, [27]. Without further hesitation let us continue with the theorem.

Theorem 5.3.1. *Let $\alpha : \mathfrak{B} \mapsto \mathfrak{A}$ be a causal QCA, then there exists a collision model $S_k : \mathfrak{B}_k \otimes \mathfrak{M} \mapsto \mathfrak{M} \otimes \mathfrak{A}_k$ with $\dim \mathcal{M} = \mathrm{ind}\alpha$ and $\dim \mathcal{M}$ is the smallest possible.*

Proof. Let us assume that α is a nearest neighbor causal automaton such that $\alpha(\mathfrak{B}_x) \subset \mathfrak{A}_{x-1} \otimes \mathfrak{A}_x$. The causality property simplifies the relations of support algebras used for the computation of index of a QCA. Following the definitions given in 4.4.5 we get

$$\mathfrak{R}_{2k} = S\big(\alpha(\mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}), \mathfrak{A}_{2k-1} \otimes \mathfrak{A}_{2k}\big) \tag{5.48}$$

$$\mathfrak{R}_{2k+1} = S\big(\alpha(\mathfrak{B}_{2k+1}), \mathfrak{A}_{2k+1}\big), \tag{5.49}$$

where $\mathfrak{R}_{2k}$ remained untouched. In $\mathfrak{R}_{2k+1}$ due to causality

$$\alpha(\mathfrak{B}_{2k}) \cup \mathfrak{A}_{2k+1} \otimes \mathfrak{A}_{2k+2} = \mathbb{I} \tag{5.50}$$

and

$$\alpha(\mathfrak{B}_{2k+1}) \cup \mathfrak{A}_{2k+1} \otimes \mathfrak{A}_{2k+2} \subset \mathfrak{A}_{2k+1}. \tag{5.51}$$

From subsection 4.4.1 we know that

$$\alpha(\mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}) \equiv \mathfrak{R}_{2k} \otimes \mathfrak{R}_{2k+1}. \tag{5.52}$$

Further we necessarily get that

$$\alpha(\mathfrak{B}_{2k}) \subset \mathfrak{R}_{2k}, \tag{5.53}$$

this means that $\alpha(\mathfrak{B}_{2k})$ is fully localized only in $\mathfrak{R}_{2k}$. Since $\mathfrak{R}_{2k}$ is a full matrix algebra $\mathfrak{B}_{2k}$ is also and α is automorphism, then $\mathfrak{R}_{2k}$ is isomorphic to

$$\mathfrak{R}_{2k} \equiv \mathfrak{B}_{2k} \otimes \mathfrak{M}, \tag{5.54}$$

for some algebra $\mathfrak{M}$ such that

$$[\alpha(\mathfrak{B}_{2k}), \mathfrak{M}] = O. \tag{5.55}$$

Thus $\mathfrak{M}$ is composed of all elements in $\mathfrak{R}_{2k}$ such that they commute with $\alpha(\mathfrak{B}_{2k})$ and secondly $\mathfrak{M}$ is isomorphic to $\mathcal{L}(\mathcal{M})$ of some Hilbert space $\mathcal{M}$. Furthermore the dimension of $\mathcal{M}$ is constant along the chain since

$$\mathrm{ind}\alpha = \frac{r_{2k}}{d_{2k}} = \frac{md_{2k}}{d_{2k}} = m := \dim \mathcal{M}, \tag{5.56}$$

it is equivalent to the index of α. Since

$$\alpha(\mathfrak{A}_{2k} \otimes \mathfrak{A}_{2k+1}) \equiv \mathfrak{R}_{2k} \otimes \mathfrak{R}_{2k+1}, \tag{5.57}$$

also

$$d_{2k}d_{2k+1} = r_{2k}r_{2k+1}, \tag{5.58}$$

because α is automorphic. However since $r_{2k} = md_{2k}$ we get that

$$d_{2k+1} = mr_{2k+1}, \tag{5.59}$$

thus

$$\alpha(\mathfrak{B}_{2k+1}) \equiv \mathfrak{M} \otimes \mathfrak{R}_{2k+1}. \tag{5.60}$$

As you might have already noticed due to causality the definition of $\mathfrak{R}_{2k+1}$ for every odd support algebra can be defined for every cell, without overlaps. To avoid confusion we will denote

$$\tilde{\mathfrak{R}}_k = S\big(\alpha(\mathfrak{B}_k), \mathfrak{A}_k\big). \tag{5.61}$$

Then

$$\alpha(\mathfrak{B}_k) \equiv \mathfrak{M} \otimes \tilde{\mathfrak{R}}_k. \tag{5.62}$$

Every single cell thus can be decomposed to a part whose localization will remain unchanged, the $\alpha(\tilde{\mathfrak{B}}_k) = \tilde{\mathfrak{R}}_k \subset \mathfrak{A}_k$ and a part which will be causally shifted to the preceding cell $\alpha(\hat{\mathfrak{B}}_k) \subset \mathfrak{A}_{k-1}$:

$$\mathfrak{B}_k \equiv \tilde{\mathfrak{B}}_k \otimes \hat{\mathfrak{B}}_k. \tag{5.63}$$

Note that the tensor product is not necessarily in the same basis as the tensor product of $\mathfrak{B}_k \otimes \mathfrak{B}_{k+1}$, since this relation emphasizes the isomorphic property. The important thing is that the parts have to commute, because α is automorphism:

$$\begin{aligned} &[\alpha(\tilde{\mathfrak{B}}_k), \alpha(\hat{\mathfrak{B}}_k)] = O \\ &\Rightarrow [\alpha(\tilde{\mathfrak{B}}_k), \alpha(\hat{\mathfrak{B}}_k)] = \alpha([\tilde{\mathfrak{B}}_k, \hat{\mathfrak{B}}_k]) = [\tilde{\mathfrak{B}}_k, \hat{\mathfrak{B}}_k] = O. \end{aligned} \tag{5.64}$$

The memory structure can be then uncovered by defining a unitary $U_k : \mathcal{M} \otimes \mathcal{A}_k \mapsto \mathcal{B}_k \otimes \mathcal{M}$ as

$$U_k^\dagger(\tilde{b}_k \otimes \mathbb{I}_{\mathcal{M}})U_k = \mathbb{I}_{\mathcal{M}} \otimes \tilde{a}_k \tag{5.65}$$

for every $\tilde{b}_k \in \tilde{\mathfrak{B}}_k$ such that $\alpha(\tilde{b}_k) = \tilde{a}_k \in \mathfrak{A}_k$ and

$$U_k^\dagger(\hat{b}_k \otimes \mathbb{I}_{\mathcal{M}})U_k = m(\hat{b}_k) \otimes \mathbb{I}_{\mathcal{A}_k} \tag{5.66}$$

for every $\hat{b}_k \in \hat{\mathfrak{B}}_k$ and such $m(\hat{b}_k) \in \mathfrak{M}$ that $\hat{b}_k \mapsto m(\hat{b}_k)$ is automorphic, and finally

$$U_k^\dagger(\mathbb{I}_{\mathcal{B}_k} \otimes m)U_k = \mathbb{I}_{\mathcal{M}} \otimes \hat{a}_k(m) \tag{5.67}$$

for all $m \in \mathfrak{M}$, where again the relation $m \mapsto \hat{a}_k(m)$ represents automorphism. To successfully construct the collision model structure we have to ensure that what was encoded by U_k inside the memory region has to be moved by U_{k-1} to $\mathfrak{A}_{k-1}$ to comply with the local rules of α such that $\hat{b}_k \mapsto m(\hat{b}_k) \mapsto \hat{a}_{k-1}(m) = \alpha(\hat{b}_k)$:

$$U_{k-1}^\dagger U_k^\dagger(\mathbb{I}_{\mathcal{B}_{k-1}}\hat{b}_k \otimes \mathbb{I}_{\mathcal{M}})U_k U_{k-1} = \mathbb{I}_{\mathcal{M}} \otimes \hat{a}_{k-1} \otimes \mathbb{I}_{\mathcal{A}_k}, \tag{5.68}$$

for every $\hat{b}_k \in \hat{\mathfrak{B}}_k$ and such $\hat{a}_{k-1} \in \mathfrak{A}_{k-1}$ that $\alpha(\hat{b}_k) = \hat{a}_{k-1}$. □

We have thus identified in each cell in nearest neighbor grouping a part of the cell whose localization will not change under the action of the automaton, the $\tilde{\mathfrak{B}}_k$ part, and the part which is causally shifted to the left, the $\hat{\mathfrak{B}}_k$ part and the causal shift can be viewed as a pure memory channel. The local rules fix the interaction completely up to the unitary freedom on memory system, which cannot be avoided, but doesn't matter at all. The whole idea is illustrated on figure 5.2. The memory channel structure can be also employed within the individual cells in nn-grouping due to the structure theorem. The nn-grouping scheme could be chosen arbitrarily, hence the memory requirement has to stay constant also inside the nn-grouping blocks.

Since QCAs are reversible, one can also construct a causal inverse of α, α', which is inverse up to shift,

$$\alpha \circ \alpha' = \sigma_k, \tag{5.69}$$

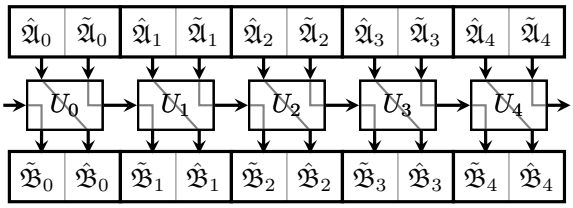


Figure 5.2: The flow of information in a causal QCA. Each cell can be split into a part whose localization will remain in the original cell, those parts are denoted by algebras with a hat, and a part whose localization is shifted by one cell to the left, those parts are denoted by algebras with tilde. The shifts are then employed as a collision model, where the size of memory algebra is equivalent to the size of the moving parts.

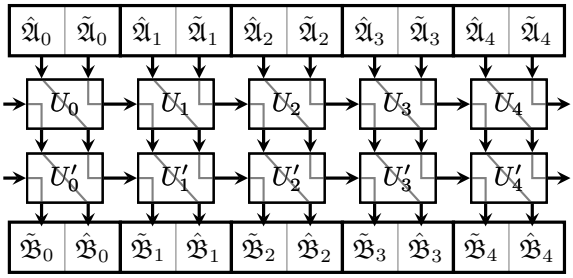


Figure 5.3: Causal inverse of automaton α such that $\alpha' \circ \alpha(\mathfrak{B}_k) = \mathfrak{A}_{k-1}$

where k is the size of the minimal nn-block (which is in turn determined by the depth of memory of the transformation). Note that we are stating things in Heisenberg picture and therefore decoding inverse α' comes before encoding transformation α and the input algebra $\mathfrak{A}$ of α' is the output algebra $\mathfrak{B}$ of α. This is however only matter of convention. The construction of such causal inverse is clarified on figure 5.3. As can be seen from the picture, for inverse the roles of tildes and hats is interchanged and the memory requirement is now determined by the tilded part which travels across the cells. If the nn-grouping is chosen to contain smallest possible blocks, then the memory requirement is exactly equal to the tilded part. Taking this back to collision models one can see that if we have a pure collision model with finite depth, which describes some communication between two parties, then it can be regarded as some encoding scheme. For this encoding there exists a perfect deterministic decoding scheme and what is more it can be used "on the fly", you do not have to wait until the communication has ended to decode. You can start the encoding after the first nn-block is already transferred or after δ uses where δ is the depth of the transformation α.

Example 5.3.2 (Nontrivial example)**.** Let us have a quantum cellular automaton α on a qubit chain with following local rules,

$$\begin{aligned}\alpha(X_k) &= c^2 X_{k-1} - s^2 Z_{k-2} X_{k-1} Z_k - cs(Z_{k-2} Y_{k-1} + Y_{k-1} Z_k)\\ \alpha(Z_k) &= Z_{k-1},\end{aligned} \tag{5.70}$$

where $c = \cos\phi$ and $s = \sin\phi$ for some angle ϕ. Since α is automorphism, $\alpha(Y_k) = -i\alpha(Z_k X_k)$ is fixed and the QCA is fully defined.

The automaton as defined is translationally invariant, but is not in nn-grouping so we have to regroup it into larger blocks, of two qubits. Let the new grouping be denoted as $\mathfrak{B}_{(k)} = \mathfrak{B}_{2k} \otimes \mathfrak{B}_{2k+1}$. First we will need to calculate the index. For this it is enough to calculate the $\tilde{\mathfrak{R}}_{(k)}$ from the equation (5.61)

$$\tilde{\mathfrak{R}}_{(k)} = S\big(\alpha(\mathfrak{B}_{(k)}), \mathfrak{A}_{(k)}\big). \tag{5.71}$$

Lets have an operator basis in $\mathfrak{B}_{(k)}$ as $\{\tau_{(k),i}\}$. Then we have to solve the system of linear equations

$$\sum_i t_i \alpha(\tau_{(k),i}) = \mathbb{I}_{(k-1)} \otimes \Omega_{(k)}, \tag{5.72}$$

where $\Omega_{(k)} \in \mathfrak{A}_{(k)}$ is some arbitrary operator. All such possible operators $\Omega_{(k)}$ span the algebra $\tilde{\mathfrak{R}}_{(k)}$ and all operators $\sum_i t_i \tau_{(k),i}$ span the algebra $\tilde{\mathfrak{B}}_{(k)}$ from (5.63). For this particular example we find that $\tilde{\mathfrak{B}}_{(k)}$ is isomorphic to a qubit algebra and is spanned by the operators

$$\begin{aligned}
\tilde{X}_{(k)} &= cX_{2k+1} + sY_{2k+1}Z_{2k} \\
\tilde{Y}_{(k)} &= cY_{2k+1} - sX_{2k+1}Z_{2k} \\
\tilde{Z}_{(k)} &= Z_{2k+1}
\end{aligned} \tag{5.73}$$

plus the identity. Hence The index of α is $\mathrm{ind}(\alpha) = d_{(k)}/2 = 2$ where $d_{(k)}$ is the dimension of Hilbert space $\mathcal{B}_{(k)}$. The traveling part of $\mathfrak{B}_{(k)}$, $\hat{\mathfrak{B}}_{(k)}$ is then the commutant of $\tilde{\mathfrak{B}}_{(k)}$ in $\mathfrak{B}_{(k)}$. Finding this commutant is again equivalent to solving a set of linear equations

$$\sum_i t_i \alpha(\tau_{(k),i}) = \Omega_{(k-1)} \otimes \mathbb{I}_{\mathcal{A}_{(k)}}, \tag{5.74}$$

for arbitrary $\Omega_{(k-1)} \in \mathfrak{A}_{(k-1)}$. Solving this we find that $\hat{\mathfrak{B}}_{(k)}$ is also isomorphic to a qubit algebra and is spanned by the operators

$$\begin{aligned}
\hat{X}_{(k)} &= cX_{2k} + sY_{2k}Z_{2k+1} \\
\hat{Y}_{(k)} &= cY_{2k} - sX_{2k}Z_{2k+1} \\
\hat{Z}_{(k)} &= Z_{2k}
\end{aligned} \tag{5.75}$$

plus the identity. The memory is of dimension 2 with operator algebra spanned by $\{\mathbb{I}_{\mathcal{M}}, X_{\mathcal{M}}, Y_{\mathcal{M}}, Z_{\mathcal{M}}\}$. Then we can reconstruct $U_{(k)} : \mathcal{M} \otimes \mathcal{A}_{(k)} \mapsto \mathcal{B}_{(k)} \otimes \mathcal{M}$ from equations (5.65)-(5.67) as

$$\begin{aligned}
U^\dagger\big(\tilde{(\cdot)}_{(k)} \otimes \mathbb{I}_{\mathcal{M}}\big)U_{(k)} &= \mathbb{I}_{\mathcal{M}} \otimes \alpha(\tilde{\cdot}) \\
U^\dagger\big(\hat{(\cdot)}_{(k)} \otimes \mathbb{I}_{\mathcal{M}}\big)U_{(k)} &= (\cdot)_{\mathcal{M}} \otimes \mathbb{I}_{\mathcal{A}_{(k)}} \\
U^\dagger\big(\mathbb{I}_{\mathcal{B}_{(k)}} \otimes (\cdot)_{\mathcal{M}}\big)U_{(k)} &= \mathbb{I}_{\mathcal{M}} \otimes \alpha(\hat{\cdot})_{(k)},
\end{aligned} \tag{5.76}$$

where $\cdot$ is to be replaced with one of $\{\mathbb{I}, X, Y, Z\}$. The solution to this is

$$U_{(k)} = \mathrm{diag}(1, e^{i\phi}, e^{i2\phi}, e^{i\phi}, e^{i\phi}, e^{i2\phi}, e^{i\phi}, 1), \tag{5.77}$$

what can be further decomposed as

$$U_{(k)} = \big(\mathbb{I}_{\mathcal{B}_{2k}} \otimes \mathrm{diag}(1, e^{i\phi}, e^{i\phi}, 1)\big)\big(\mathrm{diag}(1, e^{i\phi}, e^{i\phi}, 1) \otimes \mathbb{I}_{\mathcal{A}_{2k+1}}\big), \tag{5.78}$$

thus the memory channel corresponding to α employs the interaction $U_k : \mathcal{M} \otimes \mathcal{A}_k \mapsto \mathcal{B}_k \otimes \mathcal{M}$ where

$$U_k = \mathrm{diag}(1, e^{i\phi}, e^{i\phi}, 1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & e^{i\phi} & 0 \\ 0 & e^{i\phi} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.79}$$

This memory channel has depth of memory $\delta = 2$ and is among the solutions from previous section.

To construct the decoding inverse α' we have to find such operators $b_x, b_y, b_z \in \mathfrak{B}$ such that

$$\alpha(b_x) = X_k\,, \quad \alpha(b_y) = Y_k\,, \quad \alpha(b_z) = Z_k\,, \tag{5.80}$$

then we define α' as

$$\alpha'(X_{k+\delta}) = b_x\,, \quad \alpha'(Y_{k+\delta}) = b_y\,, \quad \alpha'(Z_{k+\delta}) = b_z. \tag{5.81}$$

The solution for this example is

$$\begin{aligned} b_x &= c^2 X_{k+1} - s^2 Z_k X_{k+1} Z_{k+2} + cs(Z_k Y_{k+1} + Y_{k+1} Z_{k+2}) \\ b_z &= Z_{k+1} \end{aligned} \tag{5.82}$$

and due to automorphic property of α' the b_y is already fixed at $b_y = -ib_z b_x$ hence the local rules for α' are

$$\begin{aligned} \alpha'(X_k) &= c^2 X_{k-1} - s^2 Z_{k-2} X_{k-1} Z_k + cs(Z_{k-2} Y_{k-1} + Y_{k-1} Z_k) \\ \alpha'(Z_k) &= Z_{k-1}, \end{aligned} \tag{5.83}$$

which is surprisingly almost the same as α but with $\phi \mapsto -\phi$. Hence the interaction for the inverse collision model is the inverse of interaction in original collision model $U'_k = U_k^\dagger$.

As a result in [37] proves that all pure memory channels with finite dimensional memory have ideal quantum capacity, however the construction of the inverse is not local. We have shown here that for strictly forgetful memory channels the causal inverse is also a strictly forgetful memory channel with finite dimensional memory. For translationally invariant maps the dimension of memory is δd where δ is the memory depth and d is the dimension of the primitive cell. This gives us hope that a similar scheme could work for forgetful channels, since such channels represent causal transformations which are not local however still could represent automorphism of quasi-local algebra. Intuition tells us that forgetful channels could achieve "approximate" locality and also an "approximate" inverse could exist where the approximation would get better as one would enlarge the "approximate" nn-grouping. Also some similar scheme could be used for memory channels which are not pure, but are still local, and this should be usually the case because entanglement tends to be hard to preserve and correlations should drop to effective zero in number of collisions.

5.4 Repeatable channels

In this section we will turn our attention in a slightly different direction. Until now we have discussed the effects of memory and how to counter them. Now we will try to find such memory channels which from some perspective do not exhibit any memory. Generally speaking a causal transformation T^* exhibits memory if

$$T^*(\omega_{[1,n]}) \neq T^* \otimes \cdots \otimes T^*(\omega_{[1,n]}). \tag{5.84}$$

Now let us have a pure memory channel with finite dimensional memory $\xi \in \mathcal{S}(\mathcal{M})$, with factorized input sequences $\omega_{[1,n]} = \omega_1 \otimes \cdots \omega_n \in \mathcal{S}(\mathcal{A}_1) \otimes \cdots \otimes \mathcal{S}(\mathcal{A}_n)$ and collision $U_k : \mathcal{M} \otimes \mathcal{A}_k \mapsto \mathcal{M} \otimes \mathcal{B}_k$. Since the input sequence is factorized, the transformation on any $\omega_{[a,b]}$, $T^*(\omega_{[a,b]})$ is a valid channel, ie. completely positive mapping. Let denote $T^*_{[a,b]}$ the channel T^* acting on $\omega_{[a,b]}$. Then we restate the relation 5.84 in a more readable way:

$$T^*_{[1,n]} \neq T^*_1 \otimes \cdots \otimes T^*_n \tag{5.85}$$

for a general memory channel. One can now be interested in the possibility of repeating some transformation G^* many times. The transformation may be a part of an experiment for example and experiments by their nature should be repeated many times. So there is a valid question which transformations G^* can be repeated in principle. From the example 4.2.3 of memoryless channels we have learned that if we need to repeat a channel G^* infinitely many times we need a memory system containing infinitely many particles at least of dimension of the minimal Stinespring's dilation for the channel G^*. This means that unless the dilation environment is trivial we need infinitely many particles and we consider this unphysical. However the dilation space is trivial only for unitary transformations. From this perspective the perfectly repeatable transformations G^* which require finite dimensional memory are only unitary.

There is not much we can do unless we lower the constraints on the repeatability of transformation. This is expressed in following definition:

Definition 5.4.1 (Repeatable transformation)**.** A channel G^* is *repeatable* if there exists a pure memory channel with finite dimensional memory, such that when constrained to factorized inputs $\omega_{[1,n]} = \omega_1 \otimes \cdots \omega_n$ we have

$$T^*_k = G^* \tag{5.86}$$

for all $i \in \mathbb{Z}$ where T^*_k is local transformation on k-th site.

Note that it stil holds that $T^*_{[a,b]} \neq G^* \otimes \cdots \otimes G^*$. We only require that *locally* the transformation on each site is equal G^*. As an example we can show that all random unitary channels are repeatable by explicitly constructing a corresponding memory channel. A random unitary channel $G^*_{\text{ru}} : \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B})$ is of form:

$$G^*_{\text{ru}}(\omega) = \sum_i p_i U_i \omega U_i^\dagger, \tag{5.87}$$

for any $p_i > 0$ such that $\sum_i p_i = 1$ and arbitrary set of unitaries $U_i : \mathcal{A} \mapsto \mathcal{B}$. The memory channel which employs these channels in repeatable fashion is the γ^* already used in Example

4.2.2. The memory channel exploited the unitary interaction $U_\gamma = F(\sum_{i=0}^{d_\mathcal{M}-1} |i\rangle\langle i| \otimes U_i)$ where we got

$$T^*{}_{\omega_{[1,n]}} = \sum_{i=0}^{d_\mathcal{M}-1} \langle i|\xi|i\rangle U_i^{\otimes n} \omega_{[1,n]} U_i^{\dagger\otimes n}, \tag{5.88}$$

where $\xi \in \mathcal{S}(\mathcal{M})$ was the initial state of memory. We compute $T_i^* = \mathrm{Tr}_{\forall l\neq k} T^*$ and get that

$$T_k^*(\omega_k) = \sum_i p_i U_i \omega_k U_i^\dagger, \tag{5.89}$$

where we assigned $p_i = \langle i|\xi|i\rangle$. Note that we do not even need the input sequence and the memory system to be factorized and this result will still hold true. Thus all random unitary channels are repeatable. Note that if a channel is repeatable it does not mean that it is always employed in repeatable fashion. The repeatability merely admits such option.

We can show that a necessary condition for a channel to be repeatable is that it has to be unital, ie. preserves the identity operator.

Theorem 5.4.2. *If a channel G^* is repeatable, then it is unital.*

Proof. Let us assume that we have a pure memory channel with finite dimensional memory system in initial state $\xi \in \mathcal{S}(\mathcal{M})$ with dimension $d_\mathcal{M} = \dim \mathcal{M}$ which employs a channel G^* in repeatable fashion. Assume factorized input sequence. Then

$$S(\xi) + \sum_{i=1}^n S(\omega_i) = S\big(U_{[1,n]}(\xi \otimes \omega_1 \otimes \cdots \omega_n) U_{[1,n]}^\dagger\big). \tag{5.90}$$

From equation 2.23 we know that the entropy of the whole is always less or equal to the entropy of the parts. Thus

$$S\big(U_{[1,n]}(\xi \otimes \omega_1 \otimes \cdots \omega_n) U_{[1,n]}^\dagger\big) \leq S\big(C^*_{[1,n]}(\xi)\big) + \sum_{i=1}^n S\big(G^*(\omega_i)\big), \tag{5.91}$$

where $C^*_{[1,n]}(\xi)$ is the final local state of memory after concurrent evolution and $G^*(\omega_i)$ is the local state of i-th particle after collision. Joining these two relations we get

$$\sum_{i=1}^n S(\omega_i) - S\big(G^*(\omega_i)\big) \leq S\big(C^*_{[1,n]}(\xi)\big) - S(\xi). \tag{5.92}$$

This means that the entropy *loss* on inputs cannot be greater than the entropy *gain* on memory system. The right hand side of (5.92) is bounded from above, because the dimension of memory is finite. The lowest entropy one can get on memory system is 0 and the highest entropy is $\log d_\mathcal{M}$ of a maximal mixture. Thus

$$\sum_{i=1}^n S(\omega_i) - S\big(G^*(\omega_i)\big) \leq \log d_\mathcal{M}. \tag{5.93}$$

Assume that all inputs are equal, $\omega_i = \omega$. Then

$$n\Big(S(\omega) - S\big(G^*(\omega)\big)\Big) \leq \log d_{\mathcal{M}}, \tag{5.94}$$

for every $n \in \mathbb{N}$. For $n \to \infty$ we have to get that $S(\omega) - S\big(G^*(\omega)\big) \leq 0$, what means that the transformation G^* cannot be entropy decreasing. This also means that such transformation has to preserve the complete mixture because complete mixture is the unique state which has the highest entropy. Since complete mixture is just scaled identity the map G^* has to preserve identity and therefore it has to be unital. It can be also shown that unital channels are entropy non-decreasing, see Appendix B.1. □

In qubit case all unital channels are also random unitary hence in qubit case unitality *implies* repeatability. Let us note that the concept of repeatability is similar to the concept of quantum cloning [60] in a sense that the channels (just like copies in quantum cloning) are not completely independent if measurements are taken into account. The impact of measurements on repeatability of quantum memory channels deserves further investigation, and will be partially addressed in the next chapter.

5.5 Simulation of indivisible qubit channels in collision models

Dynamics of open quantum systems is often modeled by the so-called master equations [38, 33]. For a comprehensive reference on evolution of open quantum systems see [17]. The idea is to get a time dependent channel describing the evolution, such that $\rho(t) = E_t^*(\rho)$, where ρ is the state of system at time $t = 0$ (thus $E_{t=0}^* = \mathbb{I}$). Often a Markovian approximation is made, where we assume that the environment is large and effectively doesn't change during the interaction with the system. In this approximation the one parametric class of channels E_t^* has a semigroup property: $E_s^* \circ E_t^* = E_{s+t}^*$. Such evolution can be stroboscopically simulated using a simple collision model.

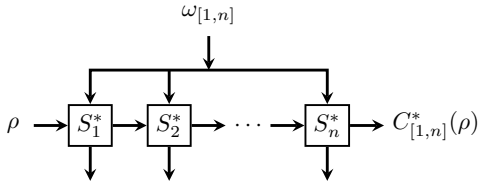


Figure 5.4: The system ρ evolves in discrete time steps in a collision model effectively simulating some evolution of an open system interacting with an environment.

Definition 5.5.1 (Stroboscopic simulation)**.** We say that a collision model $S^* : \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{A}) \mapsto \mathcal{T}(\mathcal{B}) \otimes \mathcal{T}(\mathcal{M})$ stroboscopically simulates time evolution E_t^* if

$$C_{[1,n]}^* = E_{n\Delta}^*, \tag{5.95}$$

for all $n \in \mathbb{N}$ and some $\Delta > 0$ and where $C^*_{[1,n]}$ is the concurrent channel on memory

$$C^*_{[1,n]}(\rho) = \mathrm{Tr}_{[1,n]}\big(S^*_{[1,n]}(\rho \otimes \omega[1,n])\big), \tag{5.96}$$

where $\mathrm{Tr}_{[1,n]}$ denotes partial trace over outputs $1, \ldots, n$. See for reference figure 5.4.

It is not important that the collision model is a memory channel. The inputs act only as an environment to the open system, which happens to be the memory system and we are not interested in the input-output relation of the memory channel. We focus only on the concurrent part. Stroboscopic simulation then simulates an evolution of open quantum system by discrete collisions with some structured environment, and approximates the continuous evolution with discrete time steps.

All Markovian evolutions are stroboscopically simulable. Let $E^*_t : \mathcal{T}(\mathcal{M}) \mapsto \mathcal{T}(\mathcal{M})$ be a Markovian evolution of some system with Hilbert space $\mathcal{M}$. Then let $U : \mathcal{M} \otimes \mathcal{A} \mapsto \mathcal{M} \otimes \mathcal{A}$ be a Stinespring's dilation of channel E^*_Δ:

$$E^*_\Delta(\rho) = \mathrm{Tr}_{\mathcal{A}}\big(U(\rho \otimes |0\rangle\langle 0|)U^\dagger\big), \tag{5.97}$$

for some $|0\rangle \in \mathcal{A}$. Then if we engineer environment in state $\omega_{[1,n]} = |0\rangle^{\otimes n}$ we get that

$$C^*_{[1,n]} = {E^*_\Delta}^{\otimes n} = E^*_{n\Delta}, \tag{5.98}$$

because of Markovianity of E^*_t. For any Markovian evolution we constructed a stroboscopic simulation with arbitrary small time steps. A natural property of Markovian evolutions which is also implicitly used in their stroboscopic simulation is their divisibility of any channel E^*_t (if it is not unitary evolution).

Definition 5.5.2 (Channel divisibility)**.** A channel E^* is called divisible if it can be written as a composition of two non-unitary channels:

$$E^* = E^*_1 \circ E^*_2. \tag{5.99}$$

It is important that the channels E^*_1 and E^*_2 are not unitary. Otherwise we would get that all channels are trivially divisible.

We say that a channel E^* is stroboscopically simulable if there exists a stroboscopic simulation of a continuous time evolution E^*_t such that $E^*_{t=1} = E^*$. It is an interesting question whether also indivisible channels are stroboscopically simulable.

As a partial result we are able to say that every random unitary channel is stroboscopically simulable. For this we have to fix a right interaction and an appropriate state of environment. Let us have a random unitary channel

$$E^*(\rho) = \sum_{k=0}^{d-1} p_k V_k \rho V_k^\dagger, \tag{5.100}$$

where V_k is unitary and $\sum_k p_k = 1$. Let H_k be a Hamiltonian of V_k such that $e^{iH_k} = V_k$. Fix then $\dim \mathcal{A} = d$ and interaction $U : \mathcal{M} \otimes \mathcal{A} \mapsto \mathcal{M} \otimes \mathcal{A}$

$$U = \sum_{k=0}^{d-1} e^{\frac{i}{n}H_k} \otimes |k\rangle\langle k|. \tag{5.101}$$

Finally set the state of n environmental particles to be $\omega_{[1,n]} = \sum_{k=0}^{d-1} \sqrt{p_k}|k\rangle^{\otimes n}$. This will yield a stroboscopic evolution after m collisions

$$C^*_{[1,m]}(\rho) = \sum_k p_k e^{\frac{im}{n}H_k} \rho e^{-\frac{im}{n}H_k}, \tag{5.102}$$

giving us thus that

$$C^*_{[1,n]} = \sum_k p_k e^{iH_k} \rho e^{-iH_k} = E^*. \tag{5.103}$$

One can replace this stroboscopic simulation by a continuous time evolution where

$$E^*_t(\rho) = \sum_k p_k e^{itH_k} \rho e^{-itH_k}, \tag{5.104}$$

thus the steps can be arbitrarily small.

As was first reported in [68], all indivisible qubit channels are of form:

$$E^*(\rho) = p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z, \tag{5.105}$$

$p_x p_y p_z > 0$, hence all indivisible qubit channels are stroboscopically simulable. On figure 5.5 stroboscopic evolution of universal NOT evolution $E^*_{\mathrm{NOT}}(\rho) = 1/3(X\rho X + Y\rho Y + Z\rho Z)$ is shown.

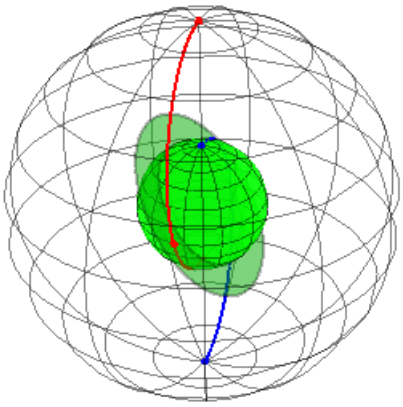


Figure 5.5: The collision model simulating the continuous time evolution towards the universal NOT gate (shrunk Bloch sphere inversion). In particular, the transformation of the Bloch sphere (lines capture the time evolution of eigenstates of Z operator) is depicted for the time interval $t \in [0, 1]$. For $t = 2/3$ the channel $E^*_{t=2/3}$ is not invertible ($\det E^*_{t=2/3} = 0$) and at this time the Bloch sphere is mapped onto a two-dimensional disk. Let us note that images of eigenstates of Z operator are internal points of the disk. In fact, the whole disk is the image of pure states only. An animation of this evolution can be found at [74].

Chapter 6

Estimation of memory channels

6.1 Process estimation in memory settings

Given a quantum box which accepts one quantum input and produces a quantum output we would like to estimate the action of such box. The box has some internal workings described by the memory degrees of freedom and the interaction between the memory an the input. We will concentrate on the most natural case when the interaction is fixed, hence the model is described by a translationally invariant memory channel. The initial state of the memory is unknown. After we use our box once, on the second time we face a *different* quantum box, with different internal state of memory, which affects the transformation. We cannot access the memory degrees of freedom at any time, making the uses explicitly dependent on the history. Thus when collecting events not only you need to store the number of occurrences, also the ordering of the events may contain important information. The question is how to approach such data.

To our best knowledge, there is no general way how to do parameter estimation in memory settings. There has been some interesting studies of hamiltonian estimations with restricted access [6, 7, 8, 18, 29, 67]. In these works the task was to estimate coupling strengths of interaction between a set of particles. This set was divided into two subsets, one of which had the experimenter full control and the rest, which was inaccessible to experimenter, and served as the memory system. If the subset under control of experimenter possessed a certain simple property, called infectivity, the experimenter was able to manipulate the inaccessible part, in order to supplement its relaxation to a desired state, and then obtain the coupling strengths of the particular model. The work [18] didn't require the relaxation of inaccessible part, due to the symmetry of interaction.

Other studies [12] focused on discrimination of combs, where combs are in principle memory channels with finite length input. They assumed that the experimenter is able to replicate the comb perfectly, thus this task is equivalent to discrimination of memoryless channels with causal structure. This structure gives additional resources to the experimenter. One can vary the input states according to outputs of previous inputs, and thus introducing a different distance measure on such processes. As a result this distance measure allows a larger set of channels to be perfectly distinguishable than the usual *cb*-norm.

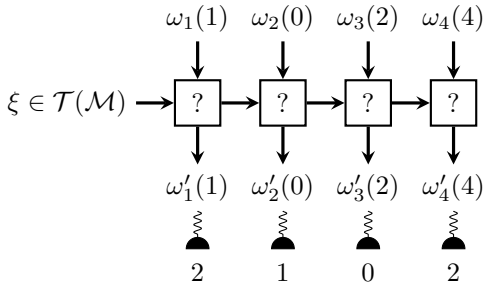


Figure 6.1: Estimation scheme of a memory channel. The inputs ω_i are drawn from a finite set of preparations, $\omega_i(k)$ means i-th input is preparation number k. Then the output $\omega'_i(k)$ is produced and measured with some outcome l. The result of estimation is then a string of events, where event is a pair (k, l), preparation - outcome.

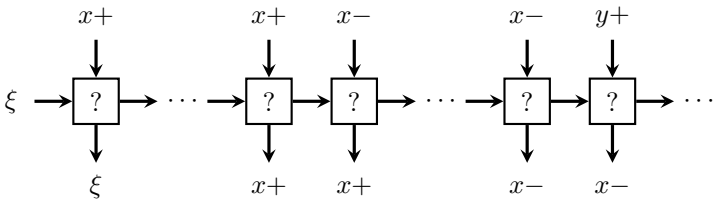


Figure 6.2: Inputs are grouped to sequences of equal preparations. Independent of measurement the result of estimation will be an almost perfect channel.

An estimation scheme of a memory channel is depicted on figure 6.1. The data from estimation are collected in the string of pairs preparation - outcome. Such estimation scheme can be interpreted as a single measurement of the memory system. This tells us that the string of observed events cannot contain more information about the initial state of the memory than one can obtain from a single measurement of the memory system. This is in contrast with memoryless channels where, depending on the interaction, some nontrivial information can be obtained about the initial state of memory. Thus the observed string of events only contains information about the inputs, measurements taken and the interaction.

To illustrate the difficulties with estimation caused by the memory effects, let us have look on following two examples. They serve as motivation as well as to show the complications we have to face, when we do not have the ability to repeat experiments independently.

Example 6.1.1 (Shift channel)**.** As was noted in previous chapters (see examples 4.1.3, 4.2.1), a simple shift channel σ_1^* on a qubit chain is modeled by a concatenated swap collision with a qubit memory. We will use an estimation procedure illustrated on figure 6.1. Let us have 6 preparations $k \in \{x+, x-, y+, y-, z+, z-,\}$ producing the eigenstates of respective Pauli matrices with ± 1 eigenvalues. Assume that we order the inputs so that first n inputs will be $\omega_i(x+), 0 < i \leq n$ then $\omega_i(x-), n < i \leq 2 * n$ and so on, see figure 6.2. Since $\omega'_i = \omega_{i-1}$ for all $i > 0$ and fix $\omega_0 \equiv \xi$, the resulting estimation will converge to an ideal channel. Because most of

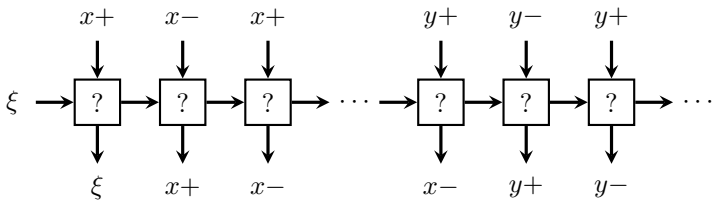


Figure 6.3: Inputs are grouped to sequences of alternating orthogonal preparations. Independent of measurement the result of estimation will be a perfect NOT, a not completely positive mapping, hence unphysical.

the time $\omega_i' = \omega_i$, only when $i = jn$ for some integer j it does not hold. As the statistics grows with n these cases become insignificant very fast. The result of estimation is independent of the details of measurement because of swap interaction.

We could choose a different strategy. We could first alternate $x+, x-$ inputs then alternate $y+, y-$ and so on. This strategy is depicted on figure 6.3. Surprisingly we will find that any input state goes to its orthogonal state. The result of estimation would be a non-completely positive mapping, a perfect NOT gate. We can see that the ordering of inputs can have significant impact on the result of estimation.

Yet, using a third strategy, we can choose the input states randomly, according to some discrete distribution with probabilities p_k of input being $\omega_i(k)$. This models a situation when the memory channel is used for communication, different preparations correspond to distinct "letters" and are distributed more or less randomly. In this case we find out that on average the output state of *any* input ω_i is $\overline{\rho} = \sum p_k \omega_{i-1}(k)$, the same as the average state of input. Thus the result of estimation would be a contraction to single state $\overline{\rho}$. This is a completely positive mapping.

For memoryless channels the ordering is not important, and you cannot possibly run into such problems. As we will see in next example not only ordering of inputs causes trouble.

Example 6.1.2 (Control not channel)**.** Let us have a memory channel with control not interaction U_{cnot}, initial qubit memory state ξ and first input $\omega_1 = 1/2(\mathbb{I} + Z)$ in the positive eigenstate of Pauli matrix Z. Let the first measurement by also a perfect S-G measurement along the z-axis. Let $p_\pm$ denote the probabilities of measuring the qubit aligned or anti-aligned with the z-axis. Then

$$\begin{aligned} p_+ &= \mathrm{Tr}(U_{\text{cnot}}(\xi \otimes \omega_1)U_{\text{cnot}}^\dagger(\mathbb{I}_\mathcal{M} \otimes E_+)) = \langle 0|\xi|0\rangle \\ p_- &= \mathrm{Tr}(U_{\text{cnot}}(\xi \otimes \omega_1)U_{\text{cnot}}^\dagger(\mathbb{I}_\mathcal{M} \otimes E_-)) = \langle 1|\xi|1\rangle \end{aligned} \tag{6.1}$$

where $E_\pm = 1/2(\mathbb{I} \pm Z)$ are the effects of measurement observable. If the positive outcome was measured then the state of memory after collision is $\xi = 1/2(\mathbb{I} + Z)$ and if the negative outcome occurred the state would be $\xi = 1/2(\mathbb{I} - Z)$. However if memory is in one of those two states, the channel will behave exactly as unitary one. For the positive result it will be ideal channel and for the negative result the inputs will experience rotation by Pauli matrix X. The state

of memory will remain untouched after the first collision. The first outcome will decide how the channel will behave afterwords. The result of estimation will be either an ideal channel or unitary rotation by X with probabilities depending on the diagonal of initial state of memory.

In the example we have observed the impact of ordering on the result of estimation. In the first case we used sequences of equal inputs. By the ordering we introduced a correlation. The state of memory before collision was in most cases the same as the input. In this sense input and state of memory were correlated. This happened also in the case of second ordering, when the state of memory before interaction was perpendicular to input. This relation of memory and colliding input was there put by hand, by means of the ordering. Thus it would be best to put the least "information" into the ordering. This can be achieved by the third option, when we have chosen the ordering to be random. In fact, if we choose random ordering the result of estimation will be always a completely positive map with a direct connection to interaction.

For further argumentation let us assume that we have a box with initial memory system $\mathcal{M}$ in state $\xi \in \mathcal{S}(\mathcal{M})$ and an interaction $S^* : \mathcal{T}(\mathcal{M} \otimes \mathcal{A}) \mapsto \mathcal{T}(\mathcal{A} \otimes \mathcal{M})$. Moreover we have at hand a set of distinct preparations $\{k \mapsto \rho_k \in \mathcal{S}(\mathcal{A})\}$ and a measurement with effects E_l. The idea is very simple. If the input states occur randomly during the process of estimation, the average state of memory entering the collision, is independent of the input state entering the collision. The state of memory conducts a classical random walk. At each step the memory gets kicked, and the kick depends on the chosen input state k and the outcome of measurement l. This kick is described by an instrument $\mathcal{I}_{k,l}$. The random walk is then just a series of these instruments. Lets denote

$$\xi_i = \frac{\mathcal{I}_{\{k,l\}_{j=1}^{i}}(\xi)}{\mathrm{Tr}\big(\mathcal{I}_{\{k,l\}_{j=1}^{i}}(\xi)\big)} \tag{6.2}$$

the state of memory after i-th collision, where $\mathcal{I}_{\{k,l\}_{j=1}^{i}}$ represents the concatenation of instruments of first i kicks. The set of ξ_i then represents a particular trajectory of evolution of memory state conditioned to the measured outcomes and particular sequence of input chosen. We can now define the mean state of memory as

$$\overline{\xi} = \sum_{i=1}^{n} \frac{1}{n} \xi_i, \tag{6.3}$$

where n is the size of statistics. Since we throw inside the testing states randomly, the average state of memory entering a collision with a fixed testing state is for $n \to \infty$ also $\overline{\xi}$. This is the same as sampling a random sample from the set of all ξ_i. For illustration see simulations of such estimation procedure on figures 6.4 and 6.5.

This leads to

$$p(k,l) = \mathrm{Tr}\big(S^*(\overline{\xi} \otimes \rho_k)(E_l \otimes \mathbb{I}_{\mathcal{M}})\big), \tag{6.4}$$

where $p(k,l)$ is the probability of measuring event (ρ_k, E_l). All measured probabilities of events agree with a memoryless channel:

$$\mathcal{E}_1^*(\rho) = \mathrm{Tr}_{\mathcal{M}}\big(S^*(\overline{\xi} \otimes \rho)\big), \tag{6.5}$$

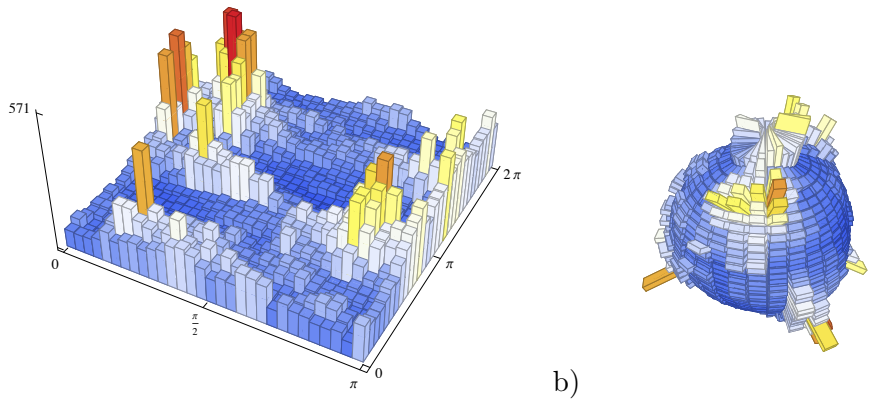


Figure 6.4: Simulation of distribution of the states of memory during a process estimation. Memory system is a qubit, interacting with chain of qubits. The interaction is a randomly chosen two qubit unitary. Testing states are 4 pure states corresponding to effects of a qubit SIC measurement chosen with equal probabilities and SIC measurement was also used to measure the outputs. a) Occurrence of memory states ξ_i in a solid angle on Bloch sphere, b) the same shown on an actual Bloch sphere. Size of statistics: 100000 events.

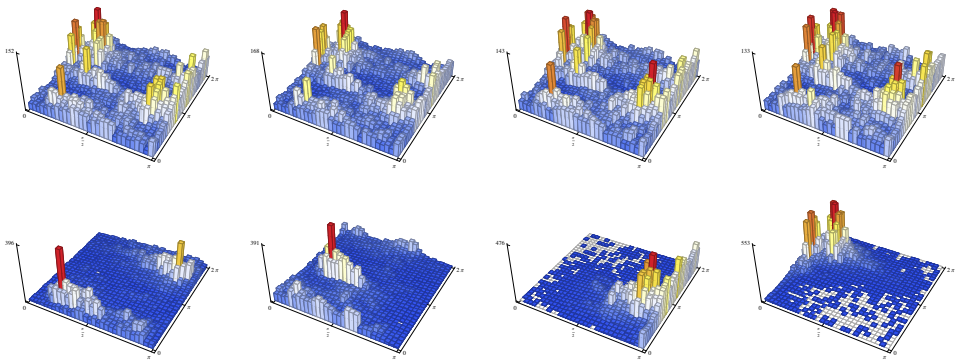


Figure 6.5: The same experiment as in figure 6.4. First row shows occurrences of memory states ξ_i entering a collision with a fixed SIC testing state. We can see that this distribution is very similar for all four SIC testing states. Second row shows the same occurrences of memory states ξ_i *after* the collision with the fixed input state. These distributions are different for every testing state, and when added together they will reconstruct the distribution of entering states.

for all $\rho \in \mathcal{S}(\mathcal{A})$. From the measured data we can actually obtain a series of maps

$$\mathcal{E}^*_{[1,m]}(\rho) = \mathrm{Tr}_{\mathcal{M}}\big(S^*_{[1,m]}(\overline{\xi} \otimes \rho_{[1,m]})\big), \tag{6.6}$$

just by grouping the inputs to larger clusters, and redefining m uses of the memory channel to one use. These maps possibly hold more information about S^* as only $\mathcal{E}^*_1$, because they reflect how the inputs get correlated. As was already observed in Example 6.1.2 not all information about the interaction can be always obtained.

For finite statistics, the accuracy of the maps $\mathcal{E}^*_{[1,m]}$ deteriorates very fast, since the number of possible events grows exponentially with m. Another issue is that though we know that state $\overline{\xi}$ exists, we do not know it. If at least one of the concurrent mappings induced by testing states,

$$C^*[k](\xi) = \mathrm{Tr}_{\mathcal{A}}\big(S^*(\xi \otimes \rho_k)\big), \tag{6.7}$$

is contractive, then the composite concurrent channel induced by an infinite input sequence projects memory to exactly one state, irrespective of initial state and thus events separated by infinite number of uses become uncorrelated. Due to contractiveness, the state space of memory is compressed exponentially, thus events separated by finite but large enough sequences will be negligibly correlated. The average memory state will be then the fixed point of the average concurrent mapping

$$\overline{C}^*(\xi) = \mathrm{Tr}_{\mathcal{A}}\big(S^*(\xi \otimes \overline{\rho})\big). \tag{6.8}$$

In the case of a pure memory channel with finite dimensional memory and additionally a distribution of testing states such that $\sum_k p_k\rho_k = 1/\dim \mathcal{A}\mathbb{I}_{\mathcal{A}}$, i.e. the average input state is a maximal mixture, then also the average output state has to be maximal mixture. The argumentation is exactly the same as in theorem 5.4.2, the entropy drop of the output is bounded by the dimension of memory and since this is finite, it has to be zero for infinitely many uses. Thus the average state of memory has to be such that the channel $\mathcal{E}^*_1$ is unital. In many cases of pure memory channels it means that the average memory is also a maximal mixture, because maximal mixture always induces a unital channel, see 2.3.9. However as we will se in the next section, the converse does not hold. For control unitary interactions, arbitrary state of memory induces a unital channel.

6.2 Control unitary interaction

In example 6.1.2 we have seen that the result of estimation was a unitary channel, in spite of that the interaction of memory and input was not factorized. In this section we will prove that this will be true whenever the interaction is a control unitary interaction, i.e. whenever

$$U = \sum_{i=0}^{d_{\mathcal{M}}-1} |i\rangle\langle i| \otimes U_i, \tag{6.9}$$

where U_i are unitaries on $\mathcal{A}$. And this result will be irrespective of details of the estimation procedure. The proof is as follows. Fix any estimation procedure. Given a memory channel, we

can also define a probability distribution on all possible outcomes that could come out of the estimation procedure. If we would repeat the same experiment, same initial memory state and same input sequence, infinitely many times the outcomes of the estimation will be distributed according to this probability distribution.

The only relevant parameters of memory are the diagonal elements of memory state ξ in the $\{|i\rangle\}$ basis as was already observed in (4.13). If, say $\langle 0|\xi|0\rangle = 1$ and all others are zero the probability distribution over possible outcomes, P_0, is as if the box was memoryless unitary evolution U_0 and similarly for other diagonal elements with distributions P_i. Due to linearity of the whole procedure, for a general state ξ the probability distribution over possible outcomes, P_ξ, will be the convex combination of P_i

$$P_\xi = \sum_{i=0}^{d_\mathcal{M}-1} \langle i|\xi|i\rangle P_i. \tag{6.10}$$

Thus only outcomes belonging to unitary channels U_i have solid probabilities. If the estimation procedure was informationally complete and statistics infinite, the result of estimation will be a memoryless unitary channel U_i with probability $\langle i|\xi|i\rangle$. For any finite statistics such estimation cannot disprove that the channel is not one of the U_i thus asymptotically will converge to the anticipated result.

In terms of the one shot measurement of the memory, the estimation is equal to measurement of the memory in the basis $\{|i\rangle\}$, assuming all U_i are different. If the input is an eigenstate of U_i, call it $|\psi\rangle$ and you measure on the output a perpendicular state $|\psi_\perp\rangle$. Then the result of estimation cannot be U_i, because it prohibits such events. The diagonal term ξ_{ii} will be set to zero after such event occurs. This has an interesting application. Given imperfect measurements on $\mathcal{A}$ one can attain asymptotically a projective measurement on $\mathcal{M}$ in aforementioned basis, by doing tomography of the memory channel. The off-diagonal terms of initial memory state go to zero exponentially and the diagonal terms ξ_{ii} are proportional to the probability of measuring the particular chain of events in unitary channel U_i. In the limit of many uses, the memory will be projected to one pure diagonal state.

Note that a special case of a control unitary interaction is factorized interaction

$$U = W \otimes V = \sum_{i=0}^{d_\mathcal{M}-1} |\psi_i\rangle\langle\psi_i| \otimes e^{i\phi_i} V, \tag{6.11}$$

where $|\psi_i\rangle$ and $e^{i\phi_i}$ are the eigenvectors and eigenvalues of W. For these interactions the result is completely obvious and natural.

We can use similar argumentation for a larger class of interactions of following form

$$U = \sum_{i=0}^{d_\mathcal{M}-1} |\pi(i)\rangle\langle i| \otimes U_i, \tag{6.12}$$

where $\pi(i)$ gives another member of the basis of memory. As in the previous case, the only relevant parameters of memory are the diagonal elements in the $\{|i\rangle\}$ basis. Now the pure diagonal state of memory is not stationary but constantly cycles through the basis states. Thus

the sequence of events will correspond to a cyclical change of unitaries. In fact this is again a control unitary channel if we group the inputs to larger sequences, this is when $\pi \cdots \pi(i) = i$ again, i.e. when the memory makes full cycle and returns to the original state. Then

$$U_{[1,n]} = \sum_{i=0}^{d_{\mathcal{M}}-1} |i\rangle\langle i| \otimes U_i \otimes U_{\pi(i)} \otimes U_{\pi(\pi(i))} \otimes \ldots, \tag{6.13}$$

where n is the length of the cycle.

6.3 2D case study

We are going to thoroughly examine the simplest example when we have a two dimensional memory system and sequence of two dimensional inputs combined with a unitary interaction. Let us first start with unitary interaction of form

$$U = D = e^{i\frac{1}{2}(\alpha_x X\otimes X + \alpha_y Y\otimes Y + \alpha_z Z\otimes Z)}, \tag{6.14}$$

with $-\pi \leq \alpha_i \leq \pi$. Given this U, our task is to estimate the angles α_i. Due to the symmetries of the problem, not all information can be obtained. Since we restrict ourselves to have access only to the input - output part, and have no access to memory degrees of freedom, the class of all interaction that will yield the same I-O relation is (see 5.1.7):

$$U = e^{i\frac{1}{2}(\alpha_x X'\otimes X + \alpha_y Y'\otimes Y + \alpha_z Z'\otimes Z)} \equiv D', \tag{6.15}$$

where $X' = \gamma X \gamma^\dagger$ and others are just unitarily conjugated Pauli matrices with arbitrary unitary γ. This implies that any two signs of the angles α_i can be flipped simultaneously by choosing one of the Pauli matrices as the unitary γ. Thus we can only estimate the sign of the product of all three angles, $\alpha_x\alpha_y\alpha_z$, which is invariant under such conjugation. Furthermore the shift by π of all three angles simultaneously introduces only a global phase on the unitary U, which is undetectable.

We will use estimation scheme described in previous section 6.1, randomizing set of preparations and measurements. We have a set of testing preparations which prepare testing states $\{\rho_k\}$ that we input randomly with probabilities p_k and a set of measurement observables. Fix the average state of memory as

$$\overline{\xi} = \frac{1}{2}(\mathbb{I} + \overline{m}_x X + \overline{m}_y Y + \overline{m}_z Z), \tag{6.16}$$

for suitable $\overline{m}_x, \overline{m}_y, \overline{m}_z$. Then any probability $p(k,l)$ of some event (k,l) is consistent with a channel in vector representation

$$A(\mathcal{E}_1^*) = A^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \overline{m}_x s_y s_z & c_y c_z & \overline{m}_z c_y s_z & -\overline{m}_y s_y c_z \\ \overline{m}_y s_x s_z & -\overline{m}_z c_x s_z & c_x c_z & \overline{m}_x s_x c_z \\ \overline{m}_z s_x s_y & \overline{m}_y c_x s_y & -\overline{m}_x s_x c_y & c_x c_y \end{pmatrix}, \tag{6.17}$$

where the matrix is written with respect to the traceless operator basis of Pauli matrices and $c_i = \cos\alpha_i$ and $s_i = \sin\alpha_i$.

The diagonal elements of the vector representation of channel $\mathcal{E}_1^*$ are independent of the average state of memory. Thus we can prepare the two testing states $\rho_{i,\pm} = 1/2(\mathbb{I} \pm \sigma_i)$ with random probability $p_{i,+} = p_{i,-} = 1/2$ and a S-G experiment along the same axis with effects $E_{i,\pm} = 1/2(\mathbb{I} \pm \sigma_i)$ where σ_i is one of the Pauli matrices. The appropriate diagonal element is then

$$\begin{aligned} A^1_{ii} &= 1/2\mathrm{Tr}(\mathcal{E}_1^*(\sigma_i)\sigma_i) = 1/4\big(\mathrm{Tr}(\mathcal{E}_1^*(\rho_{i,+})E_{i,+}) + \mathrm{Tr}(\mathcal{E}_1^*(\rho_{i,-})E_{i,-}) \\ &\quad -\mathrm{Tr}(\mathcal{E}_1^*(\rho_{i,+})E_{i,-}) - \mathrm{Tr}(\mathcal{E}_1^*(\rho_{i,-})E_{i,+})\big) \\ &= 1/2(p_{(i+,i+)} + p_{(i-,i-)} - p_{(i+,i-)} - p_{(i-,i+)}), \end{aligned} \tag{6.18}$$

where $p_{(\cdot,\cdot)}$ is the probability of measured event. If all A^1_{ii}, then

$$\begin{aligned} \cos\alpha_x &= \sqrt{\frac{A^1_{yy}A^1_{zz}}{A^1_{xx}}} \\ \cos\alpha_y &= \mathrm{sgn}(A^1_{zz})\sqrt{\frac{A^1_{xx}A^1_{zz}}{A^1_{yy}}} \\ \cos\alpha_z &= \mathrm{sgn}(A^1_{yy})\sqrt{\frac{A^1_{xx}A^1_{yy}}{A^1_{zz}}}, \end{aligned} \tag{6.19}$$

where $\cos\alpha_x$ can be always taken positive because of the π-shift symmetry mentioned earlier, The $\mathrm{sgn}(A^1_{ii})$ is the sign of the diagonal element A^1_{ii}. The last missing piece of information, the sign of $\alpha_x\alpha_y\alpha_z$ cannot be read out from the local channel. It can be obtained from the map on two subsequent inputs, $\mathcal{E}^*_{[1,2]}$. We do not need the whole $\mathcal{E}^*_{[1,2]}$. It is enough to look at an *posterior mapping*, the channel after fixed input $\rho_{i,\pm}$:

$$\mathcal{E}_2^*(\omega|\rho_{i,\pm}) = \mathrm{Tr}_1\big(\mathcal{E}^*_{[1,2]}(\rho_{i,\pm}\otimes\omega)\big). \tag{6.20}$$

For example the vector representation of channel after the input $\rho_{x,+}$ is

$$\begin{aligned} A\big(\mathcal{E}_2^*(\cdot|\rho_{x,+})\big) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ \overline{m}'_x s_y s_z & c_y c_z & \overline{m}'_z c_y s_z & -\overline{m}'_y s_y c_z \\ \overline{m}'_y s_x s_z & -\overline{m}'_z c_x s_z & c_x c_z & \overline{m}'_x s_x c_z \\ \overline{m}'_z s_x s_y & \overline{m}'_y c_x s_y & -\overline{m}'_x s_x c_y & c_x c_y \end{pmatrix} \\ &=: A^{2|x+}, \end{aligned} \tag{6.21}$$

where $\overline{m}'$ describes the posterior average state of memory after the input ρ_{x+}:

$$\begin{aligned} \overline{m}'_x &= c_y c_z \overline{m}_x + s_y s_z \\ \overline{m}'_y &= c_x c_z \overline{m}_x + s_x c_z \\ \overline{m}'_z &= c_x c_y \overline{m}_x + s_x c_y. \end{aligned} \tag{6.22}$$

If we also calculate $A^{2|x-}$ and subtract it from $A^{2|x+}$ we will get

$$\frac{1}{2}(A^{2|x+} - A^{2|x-}) =: A^{2|X}, \tag{6.23}$$

where

$$A^{2|X} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ s_y^2 s_z^2 & 0 & -c_y^2 s_z s_x \overline{m}_y & -c_z^2 s_x s_y \overline{m}_z \\ s_x^2 c_z s_z \overline{m}_z & c_x c_y s_z s_x \overline{m}_y & 0 & c_z s_x s_y s_z \\ -c_y s_x^2 s_y \overline{m}_y) & c_x c_z s_y s_x \overline{m}_z & c_y s_x s_y s_z & 0 \end{pmatrix}. \tag{6.24}$$

The sign of $\alpha_x \alpha_y \alpha_z$ is then obtained from the sign of $A_{yz}^{2|X}$ or $A_{zy}^{2|X}$ irrespective of the average state of memory. The probabilities we need to measure are

$$\begin{aligned} A_{yz}^{2|x+} &= 1/2\mathrm{Tr}(\mathcal{E}_2^*(Z|\rho_{x,+})Y) = 1/4\big(\mathrm{Tr}(\mathcal{E}_2^*(\rho_{z,+}|\rho_{x,+})E_{y,+}) + \mathrm{Tr}(\mathcal{E}_2^*(\rho_{z,-}|\rho_{x,+})E_{y,-}) \\ &\quad -\mathrm{Tr}(\mathcal{E}_2^*(\rho_{z,+}|\rho_{x,+})E_{y,-}) - \mathrm{Tr}(\mathcal{E}_2^*(\rho_{z,-}|\rho_{x,+})E_{y,+})\big) \\ &= 1/2(p_{(z+,y+|\rho_{x,+})} + p_{(z-,y-|\rho_{x,+})} - p_{(z+,y-|\rho_{x,+})} - p_{(z-,y+|\rho_{x,+})}), \end{aligned} \tag{6.25}$$

and analogously for

$$A_{yz}^{2|x-} = 1/2(p_{(z+,y+|\rho_{x,-})} + p_{(z-,y-|\rho_{x,-})} - p_{(z+,y-|\rho_{x,-})} - p_{(z-,y+|\rho_{x,-})}), \tag{6.26}$$

where $p(a,b|c)$ is the posterior probability of event (a,b) right after c occurred. This requires that the average state of memory has to be the same for both maps during the estimation.

If A^1 has zero elements on diagonal then at least two of them have to be zero. This is when at least one of the $\alpha_i = \pm\pi/2$. One can get from (5.90) the only remaining nonzero product $A_{ii}^1 = c_k c_l$. To obtain c_k and c_l we have to look in the posterior part of $\mathcal{E}_{[1,2]}^*$:

$$\begin{aligned} A_{l,0}^{2|l\pm} &= \frac{1}{2}\Big(\mathrm{Tr}\big(\mathcal{E}_2^*(1/2\mathbb{I}|\rho_{l,\pm})\rho_{l,+}\big) - \mathrm{Tr}\big(\mathcal{E}_2^*(1/2\mathbb{I}|\rho_{l,\pm})\rho_{l,-}\big)\Big) = \\ &p(l+,l+|\rho_{l,\pm}) + p(l-,l+|\rho_{l,\pm}) - p(l+,l-|\rho_{l,\pm}) - p(l-,l-|\rho_{l,\pm}) \\ &= s_i s_k(c_i c_k \overline{m}_l \pm s_i s_k) \qquad (6.27) \\ &= \pm s_k^2, \qquad (6.28) \end{aligned}$$

where the last equality is because $c_i = 0$ and $s_i^2 = 1$. In case $c_i = c_k = 0$ we will get a swap of classical information, a memory channel with depth $\delta = 2$ similar to the one in example 5.3.2, where the l-component is swapped to the l-component of the subsequent input. And finally if all $c_i = 0$ we have a memory channel with depth $\delta = 1$, what is a swap-like interaction, which can be easily checked on the posterior maps. In this case also the sign of $\alpha_x \alpha_y \alpha_z$ cannot be measured, because it only introduces a global phase shift on the swap interaction.

Thus, a good strategy is following. For every i, input randomly states $\rho_{i,\pm}$ with equal probabilities $1/2$[1] and do a measurement along the same axis with effects $E_{i,\pm}$, in order to

[1]The probabilities can be arbitrary, however it is good to keep them equal to have comparable statistics of events.

obtain the diagonal elements of A^1 from (6.18) and the elements of posterior map using (6.27):

$$\frac{1}{2}(A^{2|l+}_{l,0} - A^{2|l-}_{l,0}) = s_i^2 s_k^2. \tag{6.29}$$

Then, when lets say $c_z \neq 0$, we randomly input four 2-qubit sequences $\rho_{x,+} \otimes \rho_{z,+}$, $\rho_{x,-} \otimes \rho_{z,+}$, $\rho_{x,+} \otimes \rho_{z,-}$ and $\rho_{x,-} \otimes \rho_{z,-}$ with equal probabilities $1/4$ and measure the second output along the y- axis to measure the probabilities in (6.25) and (6.26) for estimating the $c_z s_x s_y s_z$ and subsequently the sign of $\alpha_x \alpha_y \alpha_z$. Note that this sign is unobservable if at least one of the angles is a multiple of π.

6.3.1 Adding local unitaries

Let

$$U = (\mathbb{I} \otimes V_2) D'(W_1 \otimes V_1), \tag{6.30}$$

with unknown unitaries V_i. The task is again to estimate the angles α_i and the unitaries V_i and W_1.

The vector representation of $\mathcal{E}_1^*$ or of any posterior channel is:

$$A^1 = v_2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ \overline{m}'_x s_y s_z & c_y c_z & \overline{m}'_z c_y s_z & -\overline{m}'_y s_y c_z \\ \overline{m}'_y s_x s_z & -\overline{m}'_z c_x s_z & c_x c_z & \overline{m}'_x s_x c_z \\ \overline{m}'_z s_x s_y & \overline{m}'_y c_x s_y & -\overline{m}'_x s_x c_y & c_x c_y \end{pmatrix} v_1, \tag{6.31}$$

where v_i are the three dimensional rotations of operator space corresponding to V_i and $\overline{m}' = w_1 \overline{m}$ specifies the average state of memory entering the collision rotated by the unitary W_1.

We divide the problem into two parts when the average concurrent mapping $\overline{C}^*$ is contractive and when it is not. The concurrent mapping with input state $\omega = 1/2(\mathbb{I} + r_x X + r_y Y + r_z Z)$ is

$$A(\overline{C}^*) \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ r'_x s_y s_z & c_y c_z & r'_z c_y s_z & -r'_y s_y c_z \\ r'_y s_x s_z & -r'_z c_x s_z & c_x c_z & r'_x s_x c_z \\ r'_z s_x s_y & r'_y c_x s_y & -r'_x s_x c_y & c_x c_y \end{pmatrix} w_1 =: \overline{C}, \tag{6.32}$$

where $\vec{r}' = v_1 \vec{r}$ and w_1 is the rotation corresponding to W_1. If $\overline{C}$ is contractive for all input states, then it is contractive also for $\vec{r} = \vec{0}$ and this is contractive only if all three $c_k c_l$ are less then one in absolute value. This means that at least two α_i are *not* $\pm k_i \pi$ for $k_i \in \mathbb{N}$. If at least one $c_k c_l$ is 1, then $\overline{C}$ is not contractive.

Case A: $\overline{C}$ is contractive for all input states

Let us assume that the average input state is complete mixture, i.e. that $\vec{r} = \vec{0}$. Then $\overline{C}^*$ is unital, see Remark 2.3.9. Thus the unique fixed point of $\overline{C}^*$ is maximal mixture and so is the

average state of memory. Then $\overline{m_i} = 0$ and (6.31) becomes

$$A^1 = v_2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_y c_z & 0 & 0 \\ 0 & 0 & c_x c_z & 0 \\ 0 & 0 & 0 & c_x c_y \end{pmatrix} v_1. \tag{6.33}$$

The elements $|c_i c_k|$ are singular values of the map A^1. However the singular value decomposition has to be done in such way that v_i are real and $\det v_i = 1$, because V_i are unitary. This way the singular values can't always be all positive. Note that the way how we choose the signs and ordering of singular values and the v_i is not unique, however (6.33) will hold for any chosen option, and the estimated U is unique up to known symmetries. For example if two angles are equal in absolute value, then the rotations v_i are fixed up to rotations on this subspace. Having obtained v_i, and thus also V_i one can do the estimation in the same manner as without the local unitaries, just by canceling them out with appropriate unitary rotations on inputs and outputs: $U \to (\mathbb{I} \otimes V_2^\dagger) U (\mathbb{I} \otimes V_1^\dagger)$.

The unitary W_1 can be then extracted from posterior mappings, where the average posterior state of memory entering the collision is $\overline{m}'$:

$$\overline{m}' = w_1 \Delta v_1 \vec{r}, \tag{6.34}$$

where

$$\Delta = \begin{pmatrix} s_y s_z & 0 & 0 \\ 0 & s_x s_z & 0 \\ 0 & 0 & s_x s_y \end{pmatrix}, \tag{6.35}$$

and $\vec{r}$ is the Bloch vector of input after which the posterior mapping is taken. The mapping then looks exactly as in (6.31). We need three linearly independent vectors $\vec{r}$ and respective posterior mappings to calculate w_1 from them.

Note that if $c_x = c_y$, then by symmetry you can adjust $s_x = s_y$ by suitably changing the sign of s_z. Thus you can pose $\alpha_x = \alpha_y$. Then D commutes with rotations on xy subspace and only the products of $V_1 V_2$, $W_1 V_1$ and $V_2 W_1$ can be fixed on this subspace.

Problems arise when two or three of the singular values are zero. Then we cannot determine fully the v_i from $\mathcal{E}_1^*$. Lets say that only $\alpha_x = \pm\pi/2$, then A^1_{11} is the only nonzero singular value. Then v_i are know up to arbitrary rotation in yz-plane s_i, such that $v_1 \mapsto s_1 v_1$ and $v_2 \mapsto v_2 s_2$. The vector representation of a posterior map after input $\omega = 1/2(\mathbb{I} + r_x X + r_y Y + r_z Z)$ and average memory in maximal mixture is

$$A^{2|\omega} = v_2 s_2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ \overline{m}'_x s_y s_z & c_y c_z & \overline{m}'_z c_y s_z & -\overline{m}'_y s_y c_z \\ \overline{m}'_y s_x s_z & 0 & 0 & \overline{m}'_x s_x c_z \\ \overline{m}'_z s_x s_y & 0 & -\overline{m}'_x s_x c_y & 0 \end{pmatrix} s_1 v_1, \tag{6.36}$$

where $\overline{m}'$ is the average posterior state of memory, same as in (6.34) and $s_x = \pm 1$. Let us write $A^{2|\omega}$ as 2×2 block matrix:

$$A^{2|\omega} =: v_2 \begin{pmatrix} B & D \\ E & C \end{pmatrix} v_1, \tag{6.37}$$

where particularly

$$C = s_2 \begin{pmatrix} 0 & \overline{m}'_x s_x c_z \\ -\overline{m}'_x s_x c_y & 0 \end{pmatrix} s_1. \tag{6.38}$$

Then the singular values of C are $|\overline{m}'_x s_x c_z|$ and $|\overline{m}'_x s_x c_y|$. The rotations s_i can be fixed by the singular value decomposition of C in such a way that $\det s_i = 1$ plus adding a X flip to S_1 or S_2 so that C is anti-diagonal. The ratio of the singular values is $|c_y/c_z|$, together with $c_y c_z$ this gives you the separate values for $|c_y|$ and $|c_z|$. Note that the signs can be taken arbitrary. The parameters of w_1 can be calculated from the posterior mapping for three linearly independent $\vec{r}$ as well.

If two angles are $\pm\pi/2$ and average memory state is complete mixture then $\mathcal{E}_1^*$ has all singular values zero and we have to look solely on the posterior mappings. Assuming $c_x = c_y = 0$, we will get that

$$A^{2|\omega} = v_2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ \overline{m}'_x s_y s_z & 0 & 0 & -\overline{m}'_y s_y c_z \\ \overline{m}'_y s_x s_z & 0 & 0 & \overline{m}'_x s_x c_z \\ \overline{m}'_z s_x s_y & 0 & 0 & 0 \end{pmatrix} v_1 \tag{6.39}$$

$$= \begin{pmatrix} 1 & 0 \\ \vec{t} & T \end{pmatrix}, \tag{6.40}$$

where $s_x = \pm 1$ and $s_y = \pm 1$. From this we can get one row of v_1 since now every row of T is a scaled (in this case third) row of v_1. If the scale is 0 for any ω then also the third angle is $\pm\pi/2$ and we have a swap gate, this will be dealt with later.

Let the known row of v_1 be $(\vec{v_1})_z$. We set $\vec{r} = (\vec{v_1})_z$ in ω and we will get that $\overline{m}' = w_1.(0,0,1)$. The norm of vector $\vec{t}$ is independent of v_2 and is

$$\| \vec{t} \|^2 = \| v_2 \Delta w_1.(0,0,1) \|^2 = (w_1)^2_{zz} c_z^2 + s_z^2. \tag{6.41}$$

On the other hand the squared Hilbert-Schmidt norm of T is

$$\| T \|^2_{\mathrm{HS}} = (w_1)^2_{zz} c_z^2. \tag{6.42}$$

Thus

$$s_z^2 = \| \vec{t} \|^2 - \| T \|^2_{\mathrm{HS}} . \tag{6.43}$$

The rotations v_1 and v_2 can be further adjusted by decomposing T into the form of (6.39). The rotation w_1 can be then adjusted from $\vec{t}$ again. However since $c_x = c_y$, only the products of the matrices can be fixed on the xy subspace.

In case all $c_i = 0$ we have a swap interaction with additional rotation on output, and this rotation can be checked on the posterior maps where $\vec{t} = v_2 \Delta w_1 \Delta v_1 \vec{r}$.

Adding the last local unitary W_1 from 5.24 to interaction can only change the average state of memory and average posterior state of memory, thus the rotation can be extracted from this information, in the same manner as in previous cases. We will get that $\vec{r}' = w_1 v_1 \vec{r}$, where w_1 is

the 3d rotation corresponding to W_1. The vector $\vec{r}'$ is obtained from any posterior map, if we know the s_k. Additional freedom is again introduced when some $c_i = 0$, as in previous cases.

Since we have to estimate fully the map A^1, the best strategy is to use as few as possible input states and effects and maintain the average input state in complete mixture. Posterior maps after any operator are then calculated as linear combination of measured posterior maps.

Case B: $\overline{C}$ is not contractive

The concurrent mapping $\overline{C}^*$ is not contractive when at least one of the $c_k c_l = \pm 1$. Lets say that $\alpha_x = k_x\pi$ and $\alpha_y = k_y\pi$, then we have

$$\overline{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & q(k_y)c_z & r'_z q(k_y)s_z & 0 \\ 0 & -r'_z q(k_x)s_z & q(k_x)c_z & 0 \\ 0 & 0 & 0 & q(k_x)q(k_y) \end{pmatrix} w_1, \tag{6.44}$$

where $q(k_i)$ is the sign of c_i. This is not a contractive mapping for any input state ω since it has one singular value equal to 1.

The non-local part of U can be written as

$$e^{i\frac{1}{2}(\alpha_x X\otimes X+\alpha_y Y\otimes Y+\alpha_z Z\otimes Z)} = e^{i\frac{1}{2}\alpha_z Z\otimes Z}\Gamma(k_x,k_y), \tag{6.45}$$

where $\Gamma(k_x,k_y) = e^{i\frac{1}{2}(k_x\pi X\otimes X+k_y\pi Y\otimes Y)}$ is up to global phase either $\mathbb{I}\otimes\mathbb{I}$, $X\otimes X$, $Y\otimes Y$ or $Z\otimes Z$ depending on the values of k_x, k_y. Furthermore, $e^{i\frac{1}{2}\alpha_z Z\otimes Z}$ is a control unitary interaction:

$$\begin{aligned} e^{i\frac{1}{2}\alpha_z Z\otimes Z} &= \cos\frac{\alpha_z}{2}\mathbb{I}\otimes\mathbb{I} + i\sin\frac{\alpha_z}{2}Z\otimes Z = \\ &= |z+\rangle\langle z+|\otimes e^{i\frac{1}{2}\alpha_z Z} + |z-\rangle\langle z-|\otimes e^{-i\frac{1}{2}\alpha_z Z}. \end{aligned} \tag{6.46}$$

Thus any U can be written as

$$U = (|z+\rangle\langle z+|\otimes V_+ + |z-\rangle\langle z-|\otimes V_-)(W'_1\otimes\mathbb{I}), \tag{6.47}$$

where

$$\begin{aligned} W'_1 &= \gamma(k_x,k_y)W_1, \\ V_\pm &= V_2 e^{\pm i\frac{1}{2}\alpha_z Z}\gamma(k_x,k_y)V_1 \end{aligned} \tag{6.48}$$

and

$$\gamma(k_x,k_y)\otimes\gamma(k_x,k_y) := \Gamma(k_x,k_y). \tag{6.49}$$

If W_1 is X, Y or $e^{i\theta Z}$ then U is either a control unitary interaction or a generalized control unitary interaction discussed in Section 6.2. Depending on the values of k_x, k_y and W_1 we will either estimate V_+ or V_- when U is control unitary or $V_\pm\otimes V_\mp$ on two subsequent inputs if U

is the generalized control unitary. In the latter case we are able to extract the angle α_z from eigenvalues of operator

$$\begin{aligned} V_+V_-^\dagger &= V_2 e^{i\frac{1}{2}\alpha_z Z}\gamma(k_x,k_y)V_1V_1^\dagger\gamma(k_x,k_y)^\dagger e^{i\frac{1}{2}\alpha_z Z}V_2^\dagger \\ &= V_2 e^{i\alpha_z Z}V_2^\dagger. \end{aligned} \tag{6.50}$$

If W_1 is an arbitrary rotation, the mapping $\overline{C}^*$ is not contractive but still has a unique fixed point, the complete mixture, which is the same for any input state distribution. To obtain V_1, V_2 and α_z we just need to do the singular value decomposition as in Case A. The posterior mappings $A^{2|\omega}$ are equal to A^1 because concurrent channel $\overline{C}^*$ is unital for any ω, and thus the posterior state of memory is again complete mixture.

To obtain the three independent parameters of W_1 we will have to look at individual probabilities of joint events. It turns out that the probabilities depend only on two parameters. Let

$$W_1 = \begin{pmatrix} e^{ia}\cos c & -e^{-ib}\sin c \\ e^{ib}\sin c & e^{-ia}\cos c \end{pmatrix}, \tag{6.51}$$

where $a, b, c \in \mathbb{R}$. Then every probability of any sequence of events is independent of b. This is because, when ignoring W_1 the probability of any event depends only on the diagonal elements of memory. W_1 just shuffles the diagonal and off-diagonal elements. Everything what exits the diagonal due to action of W_1 will gain a phase e^{ib} and everything what enters the diagonal will gain an opposite phase e^{-ib}. If we start in complete mixture, then the number of "exits" and "enters" has to be the same, and the phase will cancel. The remaining parameters can be obtained from the probabilities of double and triple events, for example

$$p(y+, y+|x+, x+) = \frac{1}{4}(1+\cos 2c \sin^2\alpha_z) \tag{6.52}$$

$$\begin{aligned} p(y+, y+, y+|x+, x+, x+) = \frac{1}{32}\big(&5-\cos 2\alpha_z + 2\sin^2\alpha_z(4\cos 2c \\ &+\cos 4c - 2\cos 2a\cos\alpha_z\sin^2 2c)\big), \end{aligned} \tag{6.53}$$

where $p(y+, y+|x+, x+)/p(y+, y+, y+|x+, x+, x+)$ is the probability of measuring two/three times consecutively the event $(x+, y+)$.

6.3.2 Algorithm in a nutshell

The algorithm can be summarized as following cooking receipt:
Task:
Given a pure quantum memory channel with unitary 2 qubit interaction U, qubit input, output and memory system, estimate U.
Ingredients:

- A black box with quantum input and quantum output
- Set of qubit testing states (preparations) $\{\rho_k\}_{k=1}^N$ such that ρ_k linearly span the whole operator space

- Informationally complete set of measurement observables $\{A_k\}_{k=1}^N$
- Generator of N random numbers with known probability distribution $\{p_k\}_{k=1}^N$ such that $\sum_k p_k \rho_k = 1/2\mathbb{I}$

Directions

1. Generate a random number $x \in \{1, \ldots, N\}$;
2. prepare state ρ_x, input it into the black box;
3. measure observable A_x with outcome y on the output and record measured event (x, y);
4. repeat steps (1)-(3) and keep the time-ordered list of events.

Serving instructions: The measurement observables A_k can be the same for every testing state or different. They should be such that one could use them for estimation of memoryless channels. The only difference from estimation of memoryless channels is that you need to keep track of the ordering of events.

In the two qubit unitary case it is sufficient to keep track of frequencies of at most three subsequent events. From the frequencies of single events calculate, using the equation (3.22), the channels $\mathcal{E}_1^*$ vector representation A^1. Using the same equation, but post selected data, calculate the posterior channels $\mathcal{E}_2^*(\cdot|\rho_k)$ (see (6.20)) and their vector representation $A^{2|\rho_k}$. The data for posterior channel $A^{2|\rho_k}$ consists of frequencies of single events such that the preceding event had a specific fixed input ρ_k. Since the average input state is complete mixture the map A^1 has to be unital. This means that the matrix elements A^1_{i0} should be zero for $i > 0$. However due to finite statistics the calculated mapping will be only close to unital. The numbers A^1_{i0} can serve as a rough estimate on the error of the matrix elements. Two situations emerge:
Case A:

Posterior maps $A^{2|\rho_k}$ are not unital. In this case the concurrent mapping on memory is contractive and the average memory state is complete mixture. Next step is to decompose A^1 as in equation (6.33), what is in essence a singular value decomposition, using only real rotations v_1 and v_2. Then from singular values one can calculate the parameters α_i via the equation (6.19). If some of the singular values are zero one has to calculate also the posterior mapping in (6.36), for suitable conditional input state ω. This can be achieved by linearity because ρ_k span the whole operator space.
Case B:

Posterior maps $A^{2|\rho_k}$ are very close to unital. In this case the interaction U is a control unitary with additional unitary rotation on memory W_1 as in (6.47). If A^1 is unitary then W_1 can be set to identity and we know that U is a control unitary, though we only know V_+ or V_-. If A_1 is not unitary, but the channel on every even or odd input is unitary then W_1 is a unitary flip with arbitrary phases and V_+ and V_- can be extracted from the even / odd event statistics. If the even / odd channels are equal V_+ and V_- can be obtained from the singular value decomposition (6.33) because the average state of memory is again complete mixture. The parameters of W_1 can be then calculated from probabilities of triple events using equations (6.52) and (6.53).

Chapter 7

Summary

Memory channels present a compelling model capable to describe the laboratory experiments in their full generality. They naturally describe the causal effects of how experiments in past can possibly affect the experiments in future. This demands novel protocols for interpretation of data obtained from experiments, where the memory plays an important role. The structural properties of memory channels and memory effects also deserve thorough investigation and are subject of this thesis together with the estimation.

Structural properties

We have shown that 1D quantum cellular automata can be translated into strictly forgetful pure memory channels, up to a causal shift. This is not surprising, since memory channels capture all causal processes and quantum cellular automata can be made causal with appropriate shift. However we have also found a very natural interpretation of the index of QCA in the language of memory channels. The index was originally defined in [25] as a locally computable invariant of a QCA. We have shown in [27] that this local property of any causal QCA is equal to a minimal dimension of memory system of a memory channel required to implement the automaton.

Memory effects

We have defined and investigated the depth of a memory channel. We have shown that it is equivalent to strict I-O forgetfulness, thus simplifying the criteria for strict I-O forgetfulness. In [72] we have identified all pure qubit memory channels with qubit memory system that have finite depth. In this qubit case we have shown that if a memory channel has finite depth, then the depth is less or equal to 2.

We have also defined and studied the concept of repeatable maps in memory channel setting, see ref [58]. We have shown that if a map has a repeatable implementation, then it is necessary unital. In case of qubit maps, unitality implies also existence of a repeatable implementation of the map. Furthermore, any random unitary map has repeatable implementation via control unitary interaction with memory system. It remains a question whether there exists any other repeatable implementation of unital channels which would implement also other unital channels, not only random unitary.

Thirdly, in [71], we have examined the possibility of simulating open quantum system dynamics via collision models. We have observed that correlations in environment allow us to model certain non-Markovian dynamics resulting into indivisible channels. Again in qubit case, we have shown that all indivisible channels can be a result of such dynamics, as well as any random unitary, indivisible or divisible, channel in arbitrary dimensions.

Estimation of memory channels

In [73], we have described a method to determine the interaction parameters of a quantum process with memory. The estimation procedure is the same as in the memoryless case, but with randomizing over all possible estimation events during the estimation procedure. No a priori knowledge about the initial state of memory is required. In memoryless case the order of events does not play a role and one needs to record only the statistics for single events of estimation procedure. On the other side, for memory channels, the ordering plays a vital role and additional information is stored in the statistics of consequent events.

We have thoroughly examined the 2 qubit case where we have explicitly identified the set of all parameters that can be obtained from the input-output relation and described how to obtain these parameters. We have shown that in this case it is enough to gather the statistics of at most three consequent events.

The field is still young and promising. There are many interesting ideas to pursue and topics to explore within quantum cellular automata and memory channels.

Appendix A

Hilbert space refresher

A.1 Hilbert space

Let $\mathcal{H}$ be a complex vector space.

Definition A.1.1 (Inner Product)**.** A complex valued function $\langle\cdot|\cdot\rangle$ on $\mathcal{H}\times\mathcal{H}$ is called an *inner product* on $\mathcal{H}$ if it satisfies following three conditions for all vectors $\phi,\psi,\theta\in\mathcal{H}$ and $c\in\mathbb{C}$:

- $\langle\phi|\phi\rangle\geq 0$ if $\phi\neq 0$ - positive definiteness
- $\langle\phi|\psi+c\theta\rangle=\langle\phi|\psi\rangle+c\langle\phi|\theta\rangle$ - linearity in second argument
- $\overline{\langle\phi|\psi\rangle}=\langle\psi|\phi\rangle$ - conjugate symetricity

We say that two vectors $\phi,\psi\neq 0$ of an inner product space $\mathcal{H}$ are orthogonal iff $\langle\phi|\psi\rangle=0$. A set $X\subset\mathcal{H}$ is orthogonal set if any pair of vectors from X is orthogonal.

Definition A.1.2 (Finite Dimensional space)**.** The inner product space $\mathcal{H}$ is of dimension $d\in\mathbb{N}, d<\infty$ if there does not exist a orthogonal set of k vectors such that $k>d$ but a orthogonal set of d vectors exists. If no such d exists, the space $\mathcal{H}$ is *infinitely dimensional*. Any set of d orthogonal vectors forms a (unnormed) basis in d-dimensional Hilbert space.

A complex vector space $\mathcal{H}$ equipped with inner product is a normed space with a norm defined as:

Definition A.1.3 (Canonical Norm)**.** The canonical norm of a vector $\psi\in\mathcal{H}$ with respect to inner product on $\mathcal{H}$ is

$$\parallel\psi\parallel\equiv\langle\psi|\psi\rangle^{\frac{1}{2}}. \tag{A.1}$$

Normed space is *complete* if every Cauchy sequence is convergent and *separable* if it has a countable dense subset.

Definition A.1.4 (Hilbert Space)**.** Any complete separable inner product space with respect to the norm A.1 is a *Hilbert space*.

Every finite dimensional inner product space is separable and complete and thus a Hilbert space. In this work all Hilbert spaces will be mostly finite dimensional unless explicitly stated.

Definition A.1.5 (Basis)**.** Every set of d orthonormal vectors $\{\psi_i\}$ in Hilbert space $\mathcal{H}$, where d is the dimension of this Hilbert space is a basis of this Hilbert space. Any vector ϕ can be then expressed as

$$\phi = \sum_{i=1}^{d} c_i \psi_i, \tag{A.2}$$

where $c_i = \langle \psi_i | \phi \rangle$.

Remark A.1.6 (Cauchy-Schwartz Inequality)**.** For every inner product space following inequality holds: if $\psi, \phi \in \mathcal{H}$ then

$$|\langle \phi | \psi \rangle|^2 \leq \langle \phi | \phi \rangle \langle \psi | \psi \rangle. \tag{A.3}$$

The equality happens only if ϕ and ψ are linearly dependent, ie. $\phi = c\psi$ for some $c \in \mathbb{C}$.

Definition A.1.7 (Linear Functional, Dual Space)**.** A linear mapping f from a complex vector space V to the field of complex numbers is called a linear functional. If this V has norm defined we can construct the set of all continuous linear functionals on V called the dual space V^*. This is also a vector space where the linear structure can be defined pointwise: $(f_1 + cf_2)(v) = f_1(v) + cf_2(v)$ for all $v \in V$ and normed with norm

$$\| f \| := \sup_{v} \frac{|f(v)|}{\| v \|}. \tag{A.4}$$

In Hilbert space $\mathcal{H}$ every vector ϕ defines such linear functional by the formula

$$f_\phi = \langle \phi | \psi \rangle \tag{A.5}$$

for every $\psi \in \mathcal{H}$.

Lemma A.1.8 (Riesz)**.** *Let $f \in \mathcal{H}^*$, then there exists a unique vector $\phi \in \mathcal{H}$ such that*

$$f(\psi) = \langle \phi | \psi \rangle =: f_\phi(\psi) \tag{A.6}$$

for every $\psi \in \mathcal{H}$. Moreover $\| f_\phi \| = \| \phi \|$.

Remark A.1.9 (Dirac Notation)**.** A single vector $\psi \in \mathcal{H}$ can be written as $|\psi\rangle$ and is called as *ket vector*. Symbol $\langle \phi |$ will denote a linear functional:

$$\psi \mapsto \langle \phi | \psi \rangle, \tag{A.7}$$

and is called *bra vector*. The inner product $\langle \phi | \psi \rangle$ will be then called a *bra(c)ket*.

A.2 Linear operators on Hilbert spaces

Definition A.2.1 (Linear operator)**.** We call a mapping $A : \mathcal{H} \mapsto \mathcal{H}$ *linear* if

$$A(p\psi + q\phi) = pA\psi + qA\phi, \tag{A.8}$$

for every $\psi, \phi \in \mathcal{H}$ and every $p, q \in \mathbb{C}$.

Definition A.2.2 (Bounded operator)**.** We call a linear mapping $A : \mathcal{H} \mapsto \mathcal{H}$ an *operator*. An operator A is *bounded* if there exists such number $t < \infty$ that

$$\| A\psi \| \leq t \| \psi \| \tag{A.9}$$

for all $\psi \in \mathcal{H}$. The set of all bounded operators on $\mathcal{H}$ is $\mathcal{L}(\mathcal{H})$. This set has a structure of a complex vector space. In finite dimensional Hilbert space the elements of $\mathcal{L}(\mathcal{H})$ can be represented by square matrices with finite matrix elements $\mathcal{M}_d(\mathbb{C})$.

Definition A.2.3 (Operator Norm)**.** The norm of bounded operator A is

$$\| A \|_\infty = \sup_\psi \frac{\| A\psi \|}{\| \psi \|}. \tag{A.10}$$

Definition A.2.4 (Adjoint operator)**.** For every operator $A \in \mathcal{L}(\mathcal{H})$ we can define the *adjoint operator* $A^\dagger$ as

$$\langle \phi | A^\dagger \psi \rangle = \langle A\phi | \psi \rangle \tag{A.11}$$

for all $\psi, \phi \in \mathcal{H}$. For every A, $(A^\dagger)^\dagger = A$.

Definition A.2.5 (C^*-algebra)**.** A C^*-algebra $\mathfrak{A}$ is an associative algebra over complete normed complex vector space, equipped with a $\dagger$ involution such that

$$\begin{aligned}
(A + cB)^\dagger &= A^\dagger + \overline{c} B^\dagger && \text{(A.12)} \\
(AB)^\dagger &= B^\dagger A^\dagger && \text{(A.13)} \\
(A^\dagger)^\dagger &= A && \text{(A.14)} \\
\| A^\dagger A \| &= \| A \| \| A^\dagger \|, && \text{(A.15)}
\end{aligned}$$

for every $c \in \mathbb{C}$ and $A, B \in \mathfrak{A}$.

The algebra of bounded operators $\mathcal{L}(\mathcal{H})$ over Hilbert space $\mathcal{H}$ together with the $\dagger$ operation and operator norm $\| \cdot \|_\infty$ is a C^*-algebra and conversely every C^*-algebra can be viewed as subalgebra of operators over some suitable Hilbert space. We call a C^*-algebra $\mathfrak{A}$ *unital* if it has identity.

Definition A.2.6 (Self-adjoint operator)**.** Let $A \in \mathcal{L}(\mathcal{H})$. If $A^\dagger = A$ we call such operator *self-adjoint*. The set of all self-adjoint operators is $\mathcal{L}_S(\mathcal{H})$.

For every self-adjoint operator A we know that $\langle\psi|A\psi\rangle$ is real:

$$\overline{\langle\psi|A\psi\rangle} = \langle A\psi|\psi\rangle = \langle\psi|A^\dagger\psi\rangle = \langle\psi|A\psi\rangle. \tag{A.16}$$

Definition A.2.7 (Positive operator)**.** We call an operator $A \in \mathcal{L}(\mathcal{H})$ positive if for every $|\psi\rangle \in \mathcal{H}$

$$\langle\psi|A\psi\rangle \geq 0. \tag{A.17}$$

If A is an positive operator, we write $A \geq O$.

It directly follows that positive operators are self-adjoint. It also follows that $A^\dagger A$ is a positive operator since

$$\langle\psi|A^\dagger A\psi\rangle = \langle A\psi|A\psi\rangle = \| A\psi \|^2 \geq 0 \tag{A.18}$$

Lemma A.2.8 (Square root lemma)**.** *Let $A \in \mathcal{L}_S(\mathcal{H})$ such that $A \geq O$. Then there is a unique positive operator $A^{1/2}$ such that $A^{1/2}A^{1/2} = A$. The operator $A^{1/2}$ is called a square root of A.*

Definition A.2.9 (Absolute value)**.** Absolute value of an operator $A \in \mathcal{L}(\mathcal{H})$ is $|A| := (A^\dagger A)^{1/2}$.

Definition A.2.10 (Trace of a bounded operator)**.** We define a *trace* of a bounded operator $A \in \mathcal{L}(\mathcal{H})$ as

$$\mathrm{Tr}(A) := \sum_i \langle\psi_i|A\psi_i\rangle, \tag{A.19}$$

where $\{|\psi_i\rangle\}$ forms an orthonormal basis on $\mathcal{H}$.

Definition A.2.11. The set of all operators $A \in \mathcal{L}(\mathcal{H})$ for which $\mathrm{Tr}(|A|) < \infty$ is called a *trace class* and is denoted as $\mathcal{T}(H)$.

The reason why we insist that the trace of *absolute value* of operator exists is that in infinite dimensions this ensures that the trace is unitarily invariant.

Definition A.2.12 (Trace norm)**.** Trace norm of an operator $A \in \mathcal{T}(\mathcal{H})$ is

$$\| A \|_{\mathrm{tr}} \equiv \mathrm{Tr}|A| = \mathrm{Tr}\sqrt{AA^\dagger}. \tag{A.20}$$

Definition A.2.13 (Hilbert-Schmidt product)**.** We can define an inner product on $\mathcal{T}(\mathcal{H})$:

$$\mathrm{Tr}[A^\dagger B] =: \langle A|B\rangle_{\mathrm{HS}}. \tag{A.21}$$

Hence $\mathcal{T}(\mathcal{H})$ is also a Hilbert space in finite dimension. Given that $\{|\psi_i\rangle\}$ is an orthonormal basis in d-dimensional $\mathcal{H}$ we can define a $d \times d$ dimensional orthonormal basis on $\mathcal{T}(\mathcal{H})$ as $\{|\psi_i\rangle\langle\psi_j|\}$ and every trace class operator can be expresed as

$$A = \sum_{ij} a_{ij}|\psi_i\rangle\langle\psi_j|, \tag{A.22}$$

where $a_{ij} = \mathrm{Tr}[(|\psi_i\rangle\langle\psi_j|)^\dagger A] = \langle\psi_i|A\psi_j\rangle$. Thus $\mathcal{T}(\mathcal{H})$ can be identified with the set of $d \times d$ complex matrices $\mathcal{M}_d(\mathbb{C})$. Let a_{jk} be the matrix entries of an operator $A \in \mathcal{T}(\mathcal{H})$. Arbitrary linear functional $f : \mathcal{T}(\mathcal{H}) \mapsto \mathbb{C}$ can be written as

$$\begin{aligned} f(A) &= f(\sum_{ij} a_{ij}|\psi_i\rangle\langle\psi_j|) = \sum_{ij} a_{ij} f(|\psi_i\rangle\langle\psi_j|) \\ &= \sum_{ij} a_{ij} s_{ij} = \mathrm{Tr}(AS). \end{aligned} \tag{A.23}$$

Every linear functional on $\mathcal{T}(\mathcal{H})$ is thus represented by matrix $S \in \mathcal{M}_d(\mathbb{C})$ via the trace formula A.23. This is consistent with the Riesz lemma A.1.8 if we equip $\mathcal{M}_d(\mathbb{C})$ with the Hilbert-Schmidt inner product. Let us note that in finite dimensions $\mathcal{L}(\mathcal{H}) \equiv \mathcal{L}(\mathcal{H})^* \equiv \mathcal{T}(\mathcal{H}) \equiv \mathcal{T}(\mathcal{H})^*$.

When we move into infinite dimensional case $\mathcal{L}(\mathcal{H})$ and $\mathcal{T}(\mathcal{H})$ are no longer Hilbert spaces. However $\mathcal{T}(\mathcal{H})$ is still a normed vector space and is an ideal in $\mathcal{L}(\mathcal{H})$ so that $\mathrm{Tr}(BA) < \infty$ for every $A \in \mathcal{T}(\mathcal{H})$ and $B \in \mathcal{L}(\mathcal{H})$. For each $B \in \mathcal{L}(\mathcal{H})$ we define a linear functional f_B on $\mathcal{T}(\mathcal{H})$ with

$$f_B(A) = \mathrm{Tr}(BA), \tag{A.24}$$

for every $A \in \mathcal{T}(\mathcal{H})$. Thus $\mathcal{T}(\mathcal{H})^* = \mathcal{L}(\mathcal{H})$ for infinitely dimensional $\mathcal{H}$.

The reason why $\mathcal{T}(\mathcal{H})$ is not a Hilbert space in infinite dimensions is that in fact it is too small. It turns out that the Hilbert-Schmidt inner product makes sense for a larger class of operators than trace class operators. Operators which satisfy the Hilbert-Schmidt norm $\| A \|_{\mathrm{HS}} := \sqrt{\mathrm{Tr}(A^\dagger A)} \leq \infty$ form a Hilbert space.

Definition A.2.14 (Unitary operator)**.** We call an operator U unitary if

$$UU^\dagger = U^\dagger U = \mathbb{I}. \tag{A.25}$$

The inner product in $\mathcal{H}$ is invariant under unitary change of vectors:

$$\langle U\psi|U\phi\rangle = \langle U^\dagger U\psi|\phi\rangle = \langle\psi|\phi\rangle. \tag{A.26}$$

The trace of an operator is invariant under unitary change of basis in finite dimensions, due to rotational symmetry of trace on matrices.

$$\begin{aligned} \mathrm{Tr}(A) &= \sum_i \langle U\psi_i|AU\psi_i\rangle = \mathrm{Tr}(U^\dagger AU) \\ &= \mathrm{Tr}(UU^\dagger A) = \mathrm{Tr}(A) \end{aligned} \tag{A.27}$$

Definition A.2.15 (Eigenvalues and eigenvectors)**.** A complex number $\lambda \in \mathbb{C}$ is an *eigenvalue* of a bounded operator $A \in \mathcal{L}(\mathcal{H})$ if there exists a vector $|\lambda\rangle \in \mathcal{H}, |\lambda\rangle \neq 0$ such that $A|\lambda\rangle = \lambda|\lambda\rangle$. The vector $|\lambda\rangle$ is then the *eigenvector* of A associated with the eigenvalue λ.

Definition A.2.16 (Spectrum of bounded operators)**.** *Spectrum* of an operator $A \in \mathcal{L}(\mathcal{H})$ is the set of all $\lambda \in \mathbb{C}$ such that the operator

$$R(\lambda) = (A - \lambda\mathbb{I})^{-1} \tag{A.28}$$

is not a bounded operator in $\mathcal{L}(\mathcal{H})$.

In finite dimensional case, multiplicity of eigenvalue λ is the dimension of subspace spanned by all eigenvectors associated with this eigenvalue. All eigenvalues of A are in spectrum of A. For finite dimensional $\mathcal{H}$ also the converse holds. If multiplicity of eigenvalue λ is greater than 1, we call this eigenvalue degenerate.

Unitary conjugation preserves the eigenvalues of operator $A \in \mathcal{L}(\mathcal{H})$. Let λ be an eigenvalue of A associated with vector $|\lambda\rangle$. Then $U|\lambda\rangle$ is eigenvector of $UAU^\dagger$ associated with λ:

$$UAU^\dagger U|\lambda\rangle = UA|\lambda\rangle = \lambda U|\lambda\rangle. \tag{A.29}$$

Let $|\lambda_1\rangle, |\lambda_2\rangle$ be eigenvectors of A with associated eigenvalues. Then $\langle\lambda_2|\lambda_1\rangle = c$ and we can write

$$|\lambda_1\rangle = c|\lambda_2\rangle + |\lambda_{2\perp}\rangle, \tag{A.30}$$

where $|\lambda_{2\perp}\rangle$ is orthogonal to $|\lambda_2\rangle$. Since A is self-adjoint we have

$$\langle\lambda_2|A\lambda_{2\perp}\rangle = \lambda_2\langle\lambda_2|\lambda_{2\perp}\rangle = 0. \tag{A.31}$$

Thus $|A\lambda_{2\perp}\rangle = |\lambda'_{2\perp}\rangle$ is orthogonal to $|\lambda_2\rangle$. This leads to

$$\begin{aligned} A|\lambda_1\rangle &= cA|\lambda_2\rangle + A|\lambda_{2\perp}\rangle = c\lambda_2|\lambda_2\rangle + |\lambda'_{2\perp}\rangle, \\ A|\lambda_1\rangle &= \lambda_1|\lambda_1\rangle = c\lambda_1|\lambda_2\rangle + \lambda_1|\lambda_{2\perp}\rangle. \end{aligned} \tag{A.32}$$

Both equalities can be true at the same time only if $c = 0$, states $|\lambda_1\rangle$ and $|\lambda_2\rangle$ are orthogonal, or $\lambda_1 = \lambda_2$. In finite dimensional case eigenvectors of self-adjoint operators associated with the same eigenvalue λ span a linear subspace whose dimension is equal to multiplicity of λ. You can then find a basis in this subspace and members of this basis will be again eigenvectors. Every self-adjoint operator on finite d-dimensional Hilbert space has d real eigenvalues (counting multiplicity) and therefore also d linearly independent eigenvectors which define some basis $\{|\lambda_i\rangle\}$ on this Hilbert space. Every self-adjoint operator then can be written in this basis as

$$A = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|. \tag{A.33}$$

The trace of a self-adjoint operator is the only the sum of its eigenvalues counting multiplicities.

Example A.2.17. Let us have operator A

$$A = |0\rangle\langle 0| + |0\rangle\langle 1| + 2|1\rangle\langle 1|. \tag{A.34}$$

This operator is not self-adjoint and has two eigenvectors $|\lambda_1\rangle = |0\rangle$, $|\lambda_2\rangle = |0\rangle + |1\rangle$ associated with eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$. We see that eigenvectors corresponding to different eigenvalues are not orthogonal for operator which is not self-adjoint.

Example A.2.18 (Identity)**.** Identity operator is self-adjoint and has only one eigenvalue 1 with multiplicity d, where d is the dimension of underlying Hilbert space $\mathcal{H}$. Naturally we can choose any basis $\{|\psi_i\rangle\}$ in this space and it will be automaticaly basis of eigenvectors and thus

$$\mathbb{I} = \sum_i |\psi_i\rangle\langle\psi_i| \tag{A.35}$$

for any basis in $\mathcal{H}$.

Definition A.2.19 (Projectors)**.** A projector P is a self-adjoint operator for which $P^2 = P$. Such operator can have only eigenvalues 0 or 1 and any self-adjoint operator with such eigenvalues is a projector. If the multiplicity of eigenvalue 1 is 1 then we call such projector 1-dimensional. Every one dimensional projector can then be written as $|\psi\rangle\langle\psi| =: P_\psi$ for some normed vector $\psi, \| \psi \| = 1$. And every projector is a sum of one dimensional projectors. We call projectors P_1, P_2 orthogonal if $\langle P_1|P_2\rangle_{\mathrm{HS}} = 0 \Leftrightarrow P_1P_2 = O$

Unitary operators can be also decomposed in a nice way. Every unitary operator U can be written as

$$U = \sum_k e^{i\alpha_k}|\psi_k\rangle\langle\psi_k|, \tag{A.36}$$

where $\{|\psi_k\rangle\}$ form orthonormal basis in $\mathcal{H}$ and $e^{i\alpha_k}$ are eigenvalues of U.

Appendix B

Various

B.1 Monoticity of von Neumann entropy under unital channels

Lemma B.1.1. *If G^* is a unital channel, then $S\big(G^*(\varrho)\big) \geq S(\varrho)$ for all states ϱ.*

Proof. The proof of entropy monoticity for unital channels is a consequence of the monoticity of the relative entropy [**?**]. In particular, for arbitrary quantum channel G^*

$$S\big(G^*(\varrho)||G^*(\omega)\big) \leq S(\varrho||\omega)\,, \tag{B.1}$$

where $S(\varrho||\omega) = \mathrm{Tr}\big(\varrho(\log\varrho - \log\omega)\big)$ is the quantum relative entropy. Setting $\omega = \frac{1}{d}\mathbb{I}$ we get $S(\varrho||1/d\mathbb{I}) = -S(\varrho) + \log d$. Using this fact and assuming that G^* is unital the above inequality can be rewritten as

$$\begin{aligned} S\big(G^*(\varrho)||1/d\mathbb{I})\big) &\leq S(\varrho||1/d\mathbb{I}) \\ -S\big(G^*(\varrho)\big) &\leq -S(\varrho)\,, \end{aligned}$$

from which the lemma follows.

□

Bibliography

[1] P. Arrighi, V. Nesme, and R. F. Werner, *One-dimensional quantum cellular automata over finite, unbounded configurations*, Lecture Notes in Computer Science, vol. 5196, Springer Berlin / Heidelberg, 2008.

[2] J. Ball, A. Dragan, and K. Banaszek, *Exploiting entanglement in communication channels with correlated noise*, Phys. Rev. A (2004).

[3] A. Bayat, D. Burgarth, S. Mancini, and S. Bose, *Memory effects in spin chain channels for information transmission*, Phys. Rev. A (2008).

[4] G. Bowen, I. Devetak, and S. Mancini, *Bounds on classical information capacities for a class of quantum memory*, Phys. Rev. A (2005).

[5] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics. i+ii.*, Springer, 1979, 1997.

[6] D. Burgarth and K. Maruyama, *Indirect hamiltonian identification through a small gateway*, New Journal of Physics **11** (2009), no. 10, 103019.

[7] D. Burgarth, K. Maruyama, and F. Nori, *Coupling strength estimation for spin chains despite restricted access*, Phys. Rev. A **79** (2009), 020305.

[8] ———, *Indirect quantum tomography of quadratic hamiltonians*, New Journal of Physics **13** (2011), no. 1, 013019.

[9] P. Busch, P. Lahti, and P. Mittelstaedt, *The quantum theory of measurement*, Springer, Berlin; New York, 1996.

[10] F. Caruso, V. Giovannetti, C. Macchiavello, and M. B. Ruskai, *Qubit channels with small correlations*, Phys. Rev. A (2008).

[11] N. J. Cerf, J. Clavareau, C. Macchiavello, and J. Roland, *Quantum entanglement enhances the capacity of bosonic channels with memory*, Phys. Rev. A (2005).

[12] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Memory effects in quantum channel discrimination*, Phys. Rev. Lett. **101** (2008), 180501.

[13] D. Daems, *Entanglement-enhanced classical capacity of two-qubit quantum channels with memory: the exact solution*, Phys. Rev. A (2007).

[14] G. M. D'Ariano, *Physics as quantum information processing: quantum fields as quantum automata*, arxiv:1110.6725v4.

[15] A. D'Arrigo, G. Benenti, and G. Falci, *Quantum capacity of a dephasing channel with memory*, New J. Phys. 9 (2007).

[16] N. Datta and T. Dorlas, *The coding theorem for a class of quantum channels with long-term memory*, J. Phys. A: Math. Theor. (2007).

[17] E. B. Davies, *Quantum theory of open sysytems*, Academic Press Inc, London, 1977.

[18] C. Di Franco, M. Paternostro, and M. S. Kim, *Bypassing state initialization in hamiltonian tomography on spin-chains*, SUPP. **1** (2011), 181.

[19] T. Dorlas and C. Morgan, *The classical capacity of quantum channels with memory*, Phys. Rev. A (2009).

[20] Ch. Durr and M. Santha, *A decision procedure for unitary linear quantum cellular automata*, Proceeding of the 37th IEEE Symposium on Foundations of Computer Science, 1996, pp. 38–45.

[21] T. Eggeling, D. Schlingemann, and R. F. Werner, *Semicausal operations are semilocalizable*, Europhys. Lett. **57** (2002), no. 6, 782–788.

[22] R. P. Feynman, *Simulating physics with computers*, Int. J. Theor. Phys. (1982), 467–488, Reprinted in A.J.G. Hey (ed.), Feynman and Computation, Perseus Books 1999.

[23] V. Giovannetti, *A dynamical model for quantum memory channels*, J. Phys. A: Math. Gen. (2005).

[24] V. Giovannetti and S. Mancini, *Bosonic memory channels*, Phys. Rev. A (2005).

[25] D. Gross, V. Nesme, H. Vogts, and R. F. Werner, *Index theory of one dimensional quantum walks and cellular automata*, Communications in Mathematical Physics (2009), 38.

[26] J. Gruska, *Quantum computing*, John Wiley & Sons, Inc., 2007, Section 4.3 treates QCA.

[27] J. Gütschow, T. Rybár, and R. F. Werner, *Memory requirements for general reversible qubit stream processors*, in preparation.

[28] T. Heinosaari and M. Ziman, *The mathematical language of quantum theory: From uncertainty to entanglement*, Cambridge University Press, 2011.

[29] E. L. Hosseini, K. Maruyama, D. Burgarth, T. Takui, Y. Kondo, and M. Nakahara, *Estimation of coupling constants of a three-spin chain: a case study of hamiltonian tomography with nuclear magnetic resonance*, New Journal of Physics **14** (2012), no. 1, 013043.

[30] M. Houshmand, S. Hosseini-Khayat, and M. M. Wilde, *Minimal memory requirements for pearl-necklace encoders of quantum convolutional codes*, IEEE Transactions on Computers (2010), 17–20.

[31] V. Karimipour and L. Memarzadeh, *Entanglement and optimal strings of qubits for memory channels*, Phys. Rev. A (2006).

[32] E. Karpov, D. Daems, and N. J. Cerf, *Entanglement enhanced classical capacity of quantum communication channels with correlated noise in arbitrary dimensions*, Phys. Rev. A (2006).

[33] A. Kossakowski, *On quantum statistical mechanics of non-hamiltonian systems*, Reports on Mathematical Physics **3** (1972), no. 4, 247 – 274.

[34] B. Kraus and J. I. Cirac, *Optimal creation of entanglement using a two-qubit gate*, Phys. Rev. A **63** (2001), 062309.

[35] K. Kraus, *General state changes in quantum theory*, Annals of Physics **64** (1971), no. 2, 311 – 335.

[36] K. Kraus, A. Bohm, J. D. Dollard, and W. H. Wootters, *States, effects, and operations : fundamental notions of quantum theory : lectures in mathematical physics at the university of texas at austin / karl kraus ; edited by a. bohm, j.d. dollard, and w.h. wootters*, Springer-Verlag, Berlin ; New York :, 1983 (English).

[37] D. Kretschmann and R. F. Werner, *Quantum channels with memory*, Phys. Rev. A **72** (2005), no. 6, 062323.

[38] G. Lindblad, *On the generators of quantum dynamical semigroups*, Communications in Mathematical Physics **48** (1976), 119–130, 10.1007/BF01608499.

[39] C. Lupo, V. Giovannetti, and S. Mancini, *Capacities of lossy bosonic memory channels*, Phys. Rev. Lett. (2010).

[40] C. Lupo, L. Memarzadeh, and S. Mancini, *Forgetfulness of continuous markovian quantum channels*, Phys. Rev. A (2009).

[41] C. Lupo, O. V. Pilyavets, and S. Mancini, *Capacities of lossy bosonic channel with correlated noise*, New J. Phys. (2009).

[42] C. Macchiavello and G. M. Palma, *Entanglement-enhanced information transmission over a quantum channel with correlated noise*, Phys. Rev. A (2002).

[43] C. Macchiavello, G. M. Palma, and S. Virmani, *Transition behavior in the channel capacity of two-qubit channels with memory*, Phys. Rev. A (2004).

[44] M. Nielsen and I. Chuang, *Quantum computation and quantum information*, Cambridge University Press, 2000.

[45] M. Ozawa, *Conditional expectation and repeated measurements of continuous quantum observables*, Probability Theory and Mathematical Statistics, Lecture Notes in Mathematics, vol. 1021, Springer Berlin / Heidelberg, 1983, 10.1007/BFb0072947, pp. 518–525.

[46] ———, *Quantum measuring processes of continuous observables*, Journal of Mathematical Physics **25** (1984), no. 1, 79–87.

[47] M. Paris and J. Řeháček, *Quantum state estimation*, Lecture Notes in Physics, Springer, 2004.

[48] V. I. Paulsen, *Completely bounded maps and dilations*, Cambridge University Press, Cambridge, 2002.

[49] A. Peres, *Quantum theory: Concepts and methods*, Kluwer Academic Publishers, 1995.

[50] O. V. Pilyavets, V. G. Zborovskii, and S. Mancini, *A lossy bosonic quantum channel with non-markovian memory*, Phys. Rev. A (2008).

[51] M. B. Plenio and S. Virmani, *Spin chains and channels with memory*, Phys. Rev. Lett. (2007).

[52] ———, *Many body physics and the capacity of quantum channels with memory*, New J. Phys. (2008).

[53] E. Prugovečki, *Information-theoretical aspects of quantum measurement*, International Journal of Theoretical Physics **16** (1977), 321–331, 10.1007/BF01807146.

[54] S. Richter and R. Werner, *Ergodicity of quantum cellular automata*, Journal of Statistical Physics **82** (1996), 963–998, 10.1007/BF02179798.

[55] D. Rossini, V. Giovannetti, and S. Montangero, *Spin chain model for correlated quantum channels*, New Journal of Physics **10** (2008), no. 11, 115009.

[56] G. Ruggeri and S. Mancini, *Privacy of a lossy bosonic memory channel*, Phys. Lett. A (2007).

[57] G. Ruggeri, G. Soliani, V. Giovannetti, and S. Mancini, *Information transmission through lossy bosonic memory channels*, Europhys. Lett. (2005).

[58] T. Rybár and M. Ziman, *Repeatable quantum memory channels*, Phys. Rev. A (2008).

[59] T. Salfeld, *Vergesslichkeitskriterien für quantengedächtniskanäle*, 2010.

[60] V. Scarani, S. Iblisdir, N. Gisin, and A. Acin, *Quantum cloning*, REV.MOD.PHYS. **77** (2005), 1225.

[61] J. Schäfer, D. Daems, E. Karpov, and N. J. Cerf, *Capacity of a bosonic memory channel with gauss-markov noise*, Phys. Rev. A (2009).

[62] B. Schumacher and R. F. Werner, *Reversible quantum cellular automata*, quant-ph/0405174, 2005.

[63] A. J. Scott and M. Grassl, *Symmetric informationally complete positive-operator-valued measures: A new computer study*, Journal of Mathematical Physics **51** (2010), no. 4, 042203.

[64] W. F. Stinespring, *Positive functions on c^* algebras*, Proc. Amer. Math. Soc. **6** (1955), no. 2, 211–216.

[65] W. van Dam, *Quantum cellular automata*, 1996.

[66] J. Watrous, *On one-dimensional quantum cellular automata*, Proceedings of the 36th Annual Symposium on Foundations of Computer Science (Washington, DC, USA), FOCS '95, IEEE Computer Society, 1995, pp. 528–.

[67] M. Wieśniak and M. Markiewicz, *Finding traps in nonlinear spin arrays*, Phys. Rev. A **81** (2010), 032340.

[68] M. Wolf and J. Cirac, *Dividing quantum channels*, Communications in Mathematical Physics **279** (2008), 147–168, 10.1007/s00220-008-0411-y.

[69] J. Wouters, M. Fannes, I. Akhalwaya, and F. Petruccione, *Classical capacity of a qubit depolarizing channel with memory*, Phys. Rev. A (2009).

[70] P. Zanardi, *Stabilizing quantum information*, Phys. Rev. A **63** (2000), 012301.

[71] M. Ziman, S. Filippov, T. Rybár, and V. Bužek, *Simulation of indivisible qubit channels in collision models*, accepted to JPB.

[72] M. Ziman and T. Rybár, *Quantum finite-depth memory channels: Case study*, Phys. Rev. A **80** (2009), no. 4, 7.

[73] M. Ziman, T. Rybár, and V. Bužek, *Estimation of memory channels*, in preparation.

[74] http://youtu.be/C55F4FKysh0.

Acknowledgements

This thesis would never be possible without fruitful discussions and enthusiastic supervision of doc. Mgr. Mário Ziman, PhD. and Prof. RNDr. Vladimír Bužek, DrSc. Thank you!

I would also like to thank to Prof. Reinhard F. Werner for a very warm welcome into his group in Hannover during my short visit and for introducing me to the field of quantum cellular automata.

Thank you Dano Nagaj, the distinguished citizen of office 108, for keeping an eye on me and helping me relax by killing some innocent virtual worms.

I would like to thank also to the Hedgehog and the team Obrnení SAVci, Mário Ziman, Daniel "The Autist" Reitzner, Matej "Pivo" Pivoluska, Libor "Liborek" Caha, Honza Bouda, Martin Žonda and Martin "The Scrabbler" Rejdi for many sleepless nights, where we uncovered the mysteries of the dark, cold and wet and learned to think outside the box.

Thank you Mišo Sedlák for your interest and heated discussions and for the extra bit of calories you allowed me to hunt from your plate.

Thank you, my fellow Peter "Sharky" Rapčan for sharing the wisdom of science, life, women, politics and universe in the air on the ground and under the water surface but always with a glass of a fine red wine. True gentleman.

Thank you Martina for chaining the bureaucratic beast and keeping us all well and healthy fed.

Thank you, poets of life, Jano Smrek and Juro Tekel, without you I would possibly never graduate (and would probably have a proper job by now). And thank you Sergey Nikolajevič Fillipov, Marco Genoni, Matteo Bina, Mehdi Ahmady, sir Hammed Mohamady, Razieh Annabestani, Pinja Haika, Adetunmise Charles Dada, Vedran Dunjko, Houman Mouradpour, Laleh Mamerzadeh, Christoph Spengler, Kavan Modi, Sina Salek for giving me the privilege of calling you my friends.

Thank you my parents Viera and Martin and my wife Ela for making me who I am and teaching me that education and knowledge is the only thing, nobody will ever take from me (except Alzheimer).

I would like to acknowledge the generous financial support of the European Union projects 2010-248095 (Q-ESSENCE), HIP FP7-ICT-2007-C-221889, QAP and also the projects COST action, MP1006, CE SAS QUTE, APVV DO7RP-0002-10, VEGA 2/0092/09 (QWAEN), OP CE QUTE ITMS NFP 262401022, APVV project QIAM and especially the APVV LPP-0264-07 (QWOSSI). I hope that you all feel that the money was not spent in vain. Thank you very much.