Università
degli Studi
di Pavia

Dipartimento di
Fisica
“A. Volta”

DOTTORATO DI RICERCA IN FISICA – XXIII CICLO

Higher Order Quantum Maps

dissertation submitted by

Stefano Facchini

to obtain the degree of

DOTTORE DI RICERCA IN FISICA

Supervisor: Prof. G. M. D'Ariano

Referee: Prof. J. Watrous

Higher Order Quantum Maps
Stefano Facchini
PhD thesis – University of Pavia
Printed in Pavia, Italy, December 2010
ISBN 978-88-95767-41-3

Contents

Introduction **1**

1 Mathematical preliminaries **3**
1.1 Hilbert spaces 3
1.1.1 The link product 5
1.2 Extremal operators 7

2 Higher order transformations **9**
2.1 Constructive approach 10
2.1.1 Deterministic Choi-Jamiołkowski operators 10
2.1.2 Probabilistic Choi-Jamiołkowski operators 11
2.1.3 Memory channels 12
2.2 Axiomatic approach to quantum maps 14
2.2.1 Maps from N-combs to M-combs 20
2.2.2 No-signaling channels 22
2.3 No-switch theorem 30

3 Applications of higher order quantum maps **33**
3.1 Quantum learning of unitary transformations 33
3.2 Optimal quantum tomography 42
3.3 Localizability and entanglement-breaking 50

4 Quantum lambda calculus **55**
4.1 Untyped Lambda calculus 56
4.1.1 Syntax 56
4.1.2 α-equivalence 57
4.1.3 Operational semantics 57
4.2 Simple typed lambda calculus 59

4.2.1 Syntax 59
4.2.2 Operational semantics 61
4.3 Quantum lambda calculus 62
4.3.1 Syntax 62
4.3.2 Operational semantics 64
4.4 Categorical aspects 65

Bibliography **69**

List of publications **73**

Introduction

The clarification of quantum theory has undergone a great development in the past decades. We have a deeper understanding of admissible physical transformations according to quantum mechanics. The unified treatment of every quantum object in term of the Choi-Jamiołkowski operator led to Comb Theory, which proved to be very useful both in theoretical analysis of quantum mechanics, and in practical applications, allowing the optimization of many different tasks.

This approach to quantum mechanics also suggested the possibility to introduce higher order quantum transformations, i.e. transformations operating on ordinary quantum channels. We discuss the relation between the formal aspects of higher order quantum maps and lambda calculus.

We also give contributions to the characterization of no-signaling channels, providing a new structure theorem and showing how a long-standing conjecture about non trivial no-signaling channels is actually false.

Chapter 1

Mathematical preliminaries

1.1 Hilbert spaces

Finite dimensional complex Hilbert spaces are denoted by $\mathcal{H}$, with a label when we need to distinguish them, as

$$\mathcal{H}_0, \mathcal{H}_1, \dots \tag{1.1}$$

A vector ψ belonging to a Hilbert space $\mathcal{H}_i$ will be indicated with the "ket" notation $|\psi\rangle_i$. We will denote with $\mathcal{L}(\mathcal{H})$ the space of linear operator on Hilbert space $\mathcal{H}$. Linear operators from $\mathcal{H}_0$ to $\mathcal{H}_1$ will be denoted by $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$. Sometimes, especially in the applications, this notation will be slightly modified, in particular it is convenient to indicate Hilbert spaces with a roman letter (such as A,B, . . .), in order to avoid the notational overburden of many numerical indices.

In the following we will always assume that any d-dimensional Hilbert space $\mathcal{H}$ is given with some fixed orthonormal basis $|n\rangle, n = 0, \dots, d-1$, such that we can identify

$$\mathcal{H} \cong \mathbb{C}^d, \tag{1.2}$$

Moreover, we can identify an operator $A \in \mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ with a complex matrix

$$A_{nm} := {}_1\langle n| A |m\rangle_0 . \tag{1.3}$$

To express the well-known isomorphism

$$\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1) \cong \mathcal{H}_1 \otimes \mathcal{H}_0 \tag{1.4}$$

we will use the following explicit "double ket" correspondence

$$\begin{aligned} A \in \mathcal{L}(\mathcal{H}_0, \mathcal{H}_1) \longleftrightarrow |A\rangle\!\rangle_{0,1} &:= (A \otimes I_{\mathcal{H}_0})|I_{\mathcal{H}_0}\rangle\!\rangle \\ &= (I_{\mathcal{H}_1} \otimes A^T)|I_{\mathcal{H}_1}\rangle\!\rangle \end{aligned} \tag{1.5}$$

where $|I_{\mathcal{H}}\rangle\rangle := \sum_{n=0}^{\dim(\mathcal{H})} |n\rangle |n\rangle$, and transposition is made with respect to the fixed orthonormal bases.

Combining this isomorphism with the isomorphism $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0) \cong \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_0)$ we also obtain a third fundamental isomorphism between the space of linear maps from $\mathcal{L}(\mathcal{H}_0)$ to $\mathcal{L}(\mathcal{H}_1)$, and linear operators $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$:

$$\mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1)) \cong \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0). \tag{1.6}$$

The explicit correspondence is given by the following

Definition 1 (Choi-Jamiołkowski isomorphism) *The* Choi-Jamiołkowski isomorphism *is a bijection*

$$\mathfrak{C} : \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1)) \longrightarrow \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0) \tag{1.7}$$

which, for every map $\mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1))$ *gives the following operator* $M \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$

$$M = \mathfrak{C}(\mathscr{M}) := \mathscr{M} \otimes \mathscr{I}_{\mathcal{L}(\mathcal{H}_0)}(|I_{\mathcal{H}_0}\rangle\rangle\langle\langle I_{\mathcal{H}_0}|). \tag{1.8}$$

The inverse transformation $\mathfrak{C}^{-1}$ *defines a map* $\mathfrak{C}^{-1}(M)$ *acting on* $\mathcal{L}(\mathcal{H}_0)$ *as follows*

$$\mathscr{M}(X) = \mathfrak{C}^{-1}(M)(X) = \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^T)M] \tag{1.9}$$

Lemma 1 *A linear map* $\mathscr{M}$ *is trace-preserving if and only if its Choi-Jamiołkowski operator enjoys the property*

$$\mathrm{Tr}_{\mathcal{H}_1}[M] = I_{\mathcal{H}_0}. \tag{1.10}$$

Proof. The trace preserving condition is $\mathrm{Tr}[\mathscr{M}(X)] = \mathrm{Tr}[X]$. Since

$$\mathrm{Tr}[\mathscr{M}(X)] = \mathrm{Tr}[(I_{\mathcal{H}_1} \otimes X^T)M] = \mathrm{Tr}_{\mathcal{H}_0}[X^T \,\mathrm{Tr}_{\mathcal{H}_1}[M]], \tag{1.11}$$

and $\mathrm{Tr}[X] = \mathrm{Tr}[X^T]$, the trace-preserving condition is satisfied for arbitrary X if and only if $\mathrm{Tr}_{\mathcal{H}_1}[M] = I_{\mathcal{H}_0}$. ■

Lemma 2 *A linear map* $\mathscr{M}$ *is Hermitian preserving if and only if its Choi-Jamiołkowski operator* M *is Hermitian.*

Proof. A map $\mathscr{M}$ is Hermitian preserving if $\mathscr{M}(H)^\dagger = \mathscr{M}(H)$ for any Hermitian operator H. Equivalently, if $\mathscr{M}(X^\dagger) = \mathscr{M}(X)^\dagger$ for any operator X. We have that

$$\mathscr{M}(X)^\dagger = \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^*)M^\dagger] = \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^{\dagger T})M^\dagger]. \tag{1.12}$$

Clearly, if $M^\dagger = M$ one has $\mathscr{M}(X)^\dagger = \mathscr{M}(X^\dagger)$. On the other hand, if

$$\mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^{\dagger T})M^\dagger] = \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^{\dagger T})M] \tag{1.13}$$

for all X, then $M^\dagger = M$, due to the Choi-Jamiołkowski isomorphism. ■

Lemma 3 *A linear map $\mathscr{M}$ is completely positive (CP) if and only if its Choi-Jamiołkowski operator M is positive semidefinite.*

Proof. Clearly, if $\mathscr{M}$ is CP, by Eq. (1.8) $M \geqslant 0$. On the other hand, if $M \geqslant 0$, it can be diagonalized as follows

$$M = \sum_j |K_j\rangle\!\rangle\langle\!\langle K_j|, \tag{1.14}$$

and consequently, exploiting Eqs. (1.9) and (1.5), we can write its action in the Kraus form [1]

$$\mathscr{M}(X) = \sum_j K_j X K_j^\dagger. \tag{1.15}$$

The Kraus form coming from diagonalization of M is called *canonical*. On the other hand, since the same reasoning holds for any decomposition $M = \sum_k |F_k\rangle\!\rangle\langle\!\langle F_k|$, there exist infinitely many possible Kraus forms. The Kraus form implies complete positivity: indeed, the extended map $\mathscr{M} \otimes \mathscr{I}_{\mathcal{L}\mathcal{H}_A}$ transforms any positive operator $P \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_A)$ into a positive operator, as follows

$$\mathscr{M} \otimes \mathscr{I}_{\mathcal{L}(\mathcal{H}_A)}(P) = \sum_j (K_j \otimes I_{\mathcal{H}_A}) P (K_j^\dagger \otimes I_{\mathcal{H}_A}) \geqslant 0. \tag{1.16}$$

■

1.1.1 The link product

The Choi-Jamiołkowski isomorphism poses the natural question on how the composition of linear maps is translated to a corresponding composition between the respective Choi-Jamiołkowski operators.

Consider two linear maps $\mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1))$ and $\mathscr{N} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ with Choi-Jamiołkowski operators $M \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$ and $N \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_1)$, respectively. The two maps are composed to give the linear map $\mathscr{C} = \mathscr{N} \circ \mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_2))$. This can be easily obtained upon considering the action of $\mathscr{C}$ on an operator $X \in \mathcal{L}(\mathcal{H}_0)$ written in terms of the Choi-Jamiołkowski operators of the composing maps

$$\begin{aligned} \mathscr{C}(X) &= \mathrm{Tr}_{\mathcal{H}_1}[(I_{\mathcal{H}_2} \otimes \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_1} \otimes X^T) M]^T) N] \\ &= \mathrm{Tr}_{\mathcal{H}_1, \mathcal{H}_0}[(I_{\mathcal{H}_2} \otimes I_{\mathcal{H}_1} \otimes X^T)(I_{\mathcal{H}_2} \otimes M^{T_1})(N \otimes I_{\mathcal{H}_0})]. \end{aligned} \tag{1.17}$$

Upon comparing the above identity with the Eq. (1.9) for the map $\mathscr{C}$, namely $\mathscr{C}(X) = \mathrm{Tr}_{\mathcal{H}_0}[(I_{\mathcal{H}_2} \otimes X^T) C]$, one obtains

$$C = \mathrm{Tr}_{\mathcal{H}_1}[(I_{\mathcal{H}_2} \otimes M^{T_1})(N \otimes I_{\mathcal{H}_0})], \tag{1.18}$$

where M^{T_i} denotes the partial transpose of M on the space $\mathcal{H}_i$. The above result can be expressed in a compendious way by introducing the notation

$$N * M := \mathrm{Tr}_{\mathcal{H}_1}[(I_{\mathcal{H}_2} \otimes M^{T_1})(N \otimes I_{\mathcal{H}_0})], \tag{1.19}$$

which we call *link product* of the operators $M \in \mathcal{L}\mathcal{H}_1 \otimes \mathcal{H}_0$ and $N \in \mathcal{L}\mathcal{H}_2 \otimes \mathcal{H}_1$. The above result can be synthesized in the following statement.

Theorem 1 (Composition rules) *Consider two linear maps*

$$\mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_1)) \tag{1.20}$$

and

$$\mathscr{N} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2)) \tag{1.21}$$

with Choi-Jamiołkowski operators $M \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$ and $N \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_1)$, respectively. Then, the Choi-Jamiołkowski operator $M \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_0)$ of the composition $\mathscr{C} = \mathscr{N} \circ \mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_2))$ is given by the link product of the Choi-Jamiołkowski operators $C = N * M$.

In the following we will consider more generally maps with input and output spaces that are tensor products of Hilbert spaces, and which will be composed only through some of these spaces, e.g. for quantum circuits which are composed only through some wires. For describing these compositions of maps we will need a more general definition of link product. For such purpose, consider now a couple of operators $M \in \mathcal{L}(\bigotimes_{m\in \mathsf{M}} \mathcal{H}_m)$ and $N \in \mathcal{L}(\bigotimes_{n\in \mathsf{N}} \mathcal{H}_n)$, where M and N describe set of indices for the Hilbert spaces, which generally have nonempty intersection.

The general definition of link product then reads:

Definition 2 (General link product) *The* link product *of two operators $M \in \mathcal{L}(\bigotimes_{m\in \mathsf{M}} \mathcal{H}_m)$ and $N \in \mathcal{L}(\bigotimes_{n\in \mathsf{N}} \mathcal{H}_n)$ is the operator $M*N \in \mathcal{L}(\mathcal{H}_{\mathsf{N}\backslash\mathsf{M}} \otimes \mathcal{H}_{\mathsf{M}\backslash\mathsf{N}})$ given by*

$$N * M := \mathrm{Tr}_{\mathsf{M}\cap\mathsf{N}}[(I_{\mathsf{N}\backslash\mathsf{M}} \otimes M^{T_{\mathsf{M}\cap\mathsf{N}}})(N \otimes I_{\mathsf{M}\backslash\mathsf{N}})], \tag{1.22}$$

where the set-subscript X is a shorthand for $\bigotimes_{i\in\mathsf{X}} \mathcal{H}_i$, and $\mathsf{A}\backslash\mathsf{B} := \{i \in \mathsf{A}, i \notin \mathsf{B}\}$ for two sets A and B.

Examples. For $\mathsf{M} \cap \mathsf{N} = \emptyset$, e.g. for two operators M and N acting on different Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, their link product is the tensor product:

$$N * M = N \otimes M \in \mathcal{L}\mathcal{H}_1 \otimes \mathcal{H}_0. \tag{1.23}$$

For $\mathsf{N} = \mathsf{M}$, i.e. when the two operators M and N act on the same Hilbert space, the link product becomes the trace

$$A * B = \mathrm{Tr}[A^T B]. \tag{1.24}$$

Theorem 2 (Properties of the link product) *The operation of link product has the following properties:*

1. $M * N = E(N * M)E$, *where* E *is the unitary swap on* $\mathcal{H}_{\mathsf{N}\setminus\mathsf{M}} \otimes \mathcal{H}_{\mathsf{M}\setminus\mathsf{N}}$.
2. *If* M_1, M_2, M_3 *act on Hilbert spaces labeled by the sets* $\mathsf{I}_1, \mathsf{I}_2, \mathsf{I}_3$, *respectively, and* $\mathsf{I}_1 \cap \mathsf{I}_2 \cap \mathsf{I}_3 = \emptyset$, *then* $M_1 * (M_2 * M_3) = (M_1 * M_2) * M_3$.
3. *If* M *and* N *are Hermitian, then* $M * N$ *is Hermitian.*
4. *If* M *and* N *are positive semidefinite, then* $M * N$ *is positive semidefinite.*

Proof. Properties 1, 2, and 3 are immediate from the definition. For property 4, consider the two maps $\mathscr{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_{\mathsf{M}\setminus\mathsf{N}}), \mathcal{L}(\mathcal{H}_{\mathsf{M}\cap\mathsf{N}}))$ and $\mathscr{N} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_{\mathsf{M}\cap\mathsf{N}}), \mathcal{L}(\mathcal{H}_{\mathsf{N}\setminus\mathsf{M}}))$, associated to M and N by equation Eq. (1.9). Due to Lemma 3, the maps $\mathscr{M}, \mathscr{N}$ are both CP. Moreover, due to Theorem 1 the link product $C = N * M$ is the Choi-Jamiołkowski operator of the composition $\mathscr{C} = \mathscr{N} \circ \mathscr{M}$. Since the composition of two CP maps is CP, the Choi-Jamiołkowski operator $C = N * M$ must be positive semidefinite. ■

As it should be clear to the reader, the advantage in using multipartite operators instead of maps is that we can associate many different kinds of maps to the same operator $M \in \mathcal{L}(\bigotimes_{i\in I} \mathcal{H}_i)$, depending on how we group the Hilbert spaces in the tensor product. Indeed, any partition of the set I into two disjoint sets I_0 and I_1 defines a different linear map from $\mathcal{L}(\bigotimes_{i\in I_0} \mathcal{H}_i)$ to $\mathcal{L}(\bigotimes_{i\in I_1} \mathcal{H}_i)$ via Eq. (1.9).

1.2 Extremal operators

Consider Hermitian operators $R_E, R_P \in \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1)$, with $R_E \geqslant 0$ and $\mathrm{Tr}_1[R_E] = I_0$. If we want

$$R_E + \epsilon R_P \geqslant 0 \tag{1.25}$$

for sufficiently small ϵ, then $\mathsf{Supp}(R_P) \subseteq \mathsf{Supp}(R_E)$. In fact, if R_P is supported on $\ker(R_E)$, for every $|\psi\rangle \in \ker(R_E)$ we have that

$$\langle\psi|R_E + \epsilon R_P|\psi\rangle = 0 + \epsilon\langle\psi|R_P|\psi\rangle < 0 \tag{1.26}$$

for every positive (negative) ϵ if $\langle\psi|R_P|\psi\rangle$ is negative (positive). Moreover, diagonalizing R_E we have

$$R_E = \sum_i e_i |E_i\rangle\rangle\langle\langle E_i| \tag{1.27}$$

so that $R_P = \sum_{ij} p_{ij} |E_i\rangle\!\rangle\langle\!\langle E_j|$ we have that the condition $\mathrm{Tr}_1[R_P] = 0$ becomes

$$\mathrm{Tr}_1[R_P] = \mathrm{Tr}_1[\sum_{ij} p_{ij} |E_i\rangle\!\rangle\langle\!\langle E_j|] = \tag{1.28}$$

$$= \sum_{ij} p_{ij} \,\mathrm{Tr}_1[(I_1 \otimes E_i^T)|I\rangle\!\rangle\langle\!\langle I|_{1,1'}(I_1 \otimes E_j^*)] = \sum_{ij} p_{ij} E_i^T E_j^* = 0 \tag{1.29}$$

or equivalently

$$\sum_{ij} p_{ij}^* E_i^\dagger E_j = 0. \tag{1.30}$$

This proves the following [48]

Theorem 3 (Choi) *The map represented by R_E is extremal if and only if the set $\{E_i^\dagger E_j\}$ is linearly independent.*

Chapter 2

Higher order transformations

In past decades, the development of the mathematical formalism of quantum mechanics led to a deeper understanding of which physical transformations are admissible *in principle* according quantum theory. Thanks to the works of Kraus, Davies and Lewis [2], Ozawa [4], Choi, and others, we now have a complete knowledge and characterization of the transformations involving physical systems, in terms of complete positivity and trace preservation. This research has clarified that the essence of quantum mechanics lies in the probabilistic structure of the theory, and that the mathematical constraints on quantum maps are exactly those required in order to allow a consistent probabilistic interpretation.

Recently, however, thanks to developments in the fields of quantum games and new multiparty protocols in quantum information, the scientific community spotted that the analysis of quantum transformations needed to be extended to scenarios with multiple interactive agents. More generally, the attention focused on *higher order maps*, i.e. maps which transform other maps compatibly with probability. This analysis shows the emergence of a new rich structure for quantum channels, where the pivotal role is played by *causality* constraints. The no-signaling properties expressing the causal structure of a quantum channel define a hierarchy of maps where the channels themselves are the first level. The discovery of this hierarchical structure leads us to investigate to connection with the typed lambda calculus and the recursive nature of lambda terms.

In this chapter we discuss the mathematical theory of higher order quantum theory, investigating various possibility for defining a type theory for quantum maps.

We give an example of genuine higher order map, the quantum switch, which is not realizable linking together first-level objects.

2.1 Constructive approach

Although our aim is to develop the theory of higher order maps from a truly axiomatic standpoint, we begin by discussing the constructive approach to quantum networks [6, 43]. This is useful to settle notation and to give a first introduction to the physical interpretation of the mathematical formalism presented in chapter 1.

2.1.1 Deterministic Choi-Jamiołkowski operators

In the general description of quantum mechanics, quantum states are density matrices on Hilbert space $\mathcal{H}$ of the system, i.e. positive semidefinite operators $\rho \in \mathcal{L}(\mathcal{H})$ with $\mathrm{Tr}[\rho] = 1$. Deterministic transformations of quantum states are the so-called quantum channels, a quantum channel $\mathscr{C}$ from states on $\mathcal{H}_0$ to states on $\mathcal{H}_1$ being a trace-preserving completely positive map, with diagrammatic representation

$$\overset{0}{-}\boxed{\mathscr{C}}\overset{1}{-}. \tag{2.1}$$

According to Lemmas 1, 2, 3, the Choi-Jamiołkowski operator corresponding to $\mathscr{C}$ is a positive semidefinite operator $C \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$ satisfying $\mathrm{Tr}_{\mathcal{H}_1}[C] = I_{\mathcal{H}_0}$.

It is immediate to see that a density matrix is a particular case of Choi-Jamiołkowski operator of a channel, namely a Choi-Jamiołkowski operator with one-dimensional input space $\mathcal{H}_0$: in this case the condition $\mathrm{Tr}_{\mathcal{H}_1}[C] = I_{\mathcal{H}_0}$ becomes indeed $\mathrm{Tr}[C] = 1$. This reflects the fact that having a quantum state is equivalent to having at disposal one use of a suitable preparation device. A state is represented by

$$\left(\rho\right|\overset{1}{-}. \tag{2.2}$$

The application of the channel $\mathscr{C}$ to the state ρ is equivalent to the composition of two channels, and is indeed given by the link product of the corresponding Choi-Jamiołkowski operators

$$\mathscr{C}(\rho) = C * \rho, \tag{2.3}$$

which agrees both with Eq. (1.9) and Theorem 1.

The opposite example is the completely demolishing "trace channel" $\mathscr{T}(\rho) = \mathrm{Tr}[\rho]$, which transforms quantum states into their probabilities (of course, normalized density matrices give unit probabilities): this channel has one-dimensional output space $\mathcal{H}_1$, and, accordingly its Choi-Jamiołkowski operator is $T = I_{\mathcal{H}_0}$. We picture this channel as

$$\overset{0}{-}\left|I\right). \tag{2.4}$$

Notice that the normalization of the Choi-Jamiołkowski operator $C \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_0)$ of a channel $\mathscr{C}$ can be also written in terms of concatenation with the trace channel as

$$C * I_{\mathcal{H}_1} = I_{\mathcal{H}_0}. \tag{2.5}$$

2.1.2 Probabilistic Choi-Jamiołkowski operators

In addition to the Choi-Jamiołkowski operators of deterministic quantum devices, one can consider their probabilistic versions. A complete family of probabilistic transformations from states on $\mathcal{H}_0$ to states on $\mathcal{H}_1$, known as *quantum instrument*, is a set of CP maps $\{\mathscr{C}_i \mid i \in I\}$ summing up to a trace-preserving CP map $\mathscr{C} = \sum_{i\in I} \mathscr{C}_i$. The corresponding Choi-Jamiołkowski operators $\{C_i \mid i \in I\}$ are positive semidefinite operators summing up to a deterministic Choi-Jamiołkowski operator $C = \sum_{i\in I} C_i$ with $C * I_{\mathcal{H}_1} = I_{\mathcal{H}_0}$. For families of probabilistic transformations, the index i has always to be intended as a classical outcome, that is known to the experimenter, and heralds the occurrence of different random transformations.

For one-dimensional input space $\mathcal{H}_0$, a complete family of probabilistic Choi-Jamiołkowski operators $\{\rho_i \mid i \in I\}$ with $\sum_i \rho_i = \rho, \mathrm{Tr}[\rho] = 1$ describes a *random source of quantum states*. Applying the trace channel $\mathscr{T}$ after the source gives the probability of the source emitting the i-th state: $p_i = \mathrm{Tr}[\rho_i] = \rho_i * I_{\mathcal{H}_1}$ (of course $p_i \geq 0$ and $\sum_i p_i = 1$).

For one-dimensional output space $\mathcal{H}_1$, a complete family of probabilistic Choi-Jamiołkowski operators is instead a POVM $\{P_i \mid i \in I\}$, $\sum_i P_i = I_{\mathcal{H}_0}$. The diagrammatic representation of a POVM is

$$\overset{0}{-}\!\boxed{P_i}\,. \tag{2.6}$$

Measuring the POVM on the state ρ is equivalent to applying the random device described by $\{P_i\}$ after the preparation device for state ρ, producing as the outcome the probabilities

$$p(i|\rho) = \rho * P_i = \mathrm{Tr}[\rho P_i^T]. \tag{2.7}$$

Apart from the transpose, which can be absorbed in the definition of the POVM, this is nothing but the Born rule for probabilities, obtained here from the composition of a preparation channel with a random transformation with one-dimensional output space.

In conclusion, states, channels, random sources, instruments, and POVMs can be treated on the same footing as deterministic and probabilistic transformations, which in turn can be described using only Choi-Jamiołkowski operators and link product.

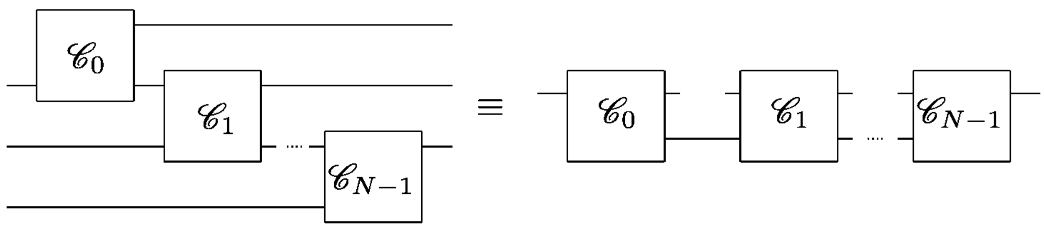


Figure 2.1

2.1.3 Memory channels

A quantum network is obtained by assembling a number of elementary circuits, each of them represented by its Choi-Jamiołkowski operator.

To build up a particular quantum network one needs to have at disposal the whole list of elementary circuits and a list of instructions about how to connect them. In connecting circuits there are clearly two restrictions: *i)* one can only connect the output of a circuit with the input of another circuit, and *ii)* there cannot be cycles. These restrictions ensure causality, namely the fact that quantum information in the network flows from input to output without loops. This implies that the connections in the quantum network can be represented in a directed acyclic graph (DAG), where each vertex represents a quantum circuit, and each arrow represents a quantum system traveling from one circuit to another. Notice that such a graph represents only the internal connections of the networks, while to have a complete graphical representation one should also append to the vertices a number of free incoming and outgoing arrows representing quantum systems that enter or exit the network. In other words, the graphical representation of a quantum network is provided by a DAG where some sources (vertices without incoming arrows) and some sinks (vertices without outgoing arrows) have been removed. The free arrows remaining after removing a source represent input systems entering the network, while the free arrows remaining after removing a sink represent output systems exiting the network.

The flow of quantum information along the arrows of the graph induces a partial ordering of the vertices: we say that the circuit in vertex v_1 *causally precedes* the circuit in vertex v_2 ($v_1 \preceq v_2$) if there is a directed path from v_1 to v_2. A well known theorem in graph theory states that for a directed acyclic graph there always exists a way to extend the partial ordering $\preceq$ to a total ordering $\leq$ of the vertices. Intuitively speaking, the relation $\leq$ fixes a schedule for the order in which the circuits in the network can be run, compatibly with the causal ordering of input-output relations. In general, the total ordering $\leq$ is not uniquely determined by the partial ordering $\preceq$:

the same quantum network can be used in different ways, corresponding to different orders in which the elementary circuits are run.

A quantum network with a given sequential ordering of the vertices becomes a compound quantum circuit, in which different operations are performed according to a precise schedule. Totally ordered quantum networks have a large number of applications in quantum information, and, accordingly, they have been given different names, depending on the context. For example, they are referred to as *quantum strategies* in quantum game theoretical and cryptographic applications [6]. Moreover, a totally ordered quantum network is equivalent to a sequence of channels with memory, as illustrated in Fig. 2.1.

We have the following characterization theorem for memory channels

Theorem 4 *Let $R^{(N)}$ be a positive operator satisfying relations*

$$\begin{aligned} &R^{(j)} * I_{2j-1} = R^{(j-1)} * I_{2j-2}, \quad 2 \leqslant j \leqslant N \\ &R^{(1)} * I_1 = I_0. \end{aligned} \tag{2.8}$$

for suitable operators $R^{(j)}$. Then, $R^{(N)}$ is the Choi-Jamiołkowski operator of a memory channel.

Proof. We want to prove that $\mathscr{R}^{(N)} = \mathfrak{C}(R^{(N)})$ is a memory channel. In particular, we want to show that $\mathscr{R}^{(N)}$ is obtained as a concatenation of N isometries.

The proof is by induction. For $N = 1$ the statement is equivalent to Stinespring's dilation of channels [3]: there is an isometry $W^{(1)}$ with ancillary space A such that $\mathscr{R}^{(1)}(\rho) = \mathrm{Tr}_A[W^{(1)}\rho W^{(1)\dagger}]$.

We now suppose that the minimal isometry $W^{(N)}$ dilating $\mathscr{R}^{(N)}$ arises from the concatenation of N isometries, and show that also the isometry $W^{(N+1)}$ is the concatenation of $N+1$ isometries.

$$I_{2N+1} * R^{(N+1)} = I_{2N} \otimes R^{(N)} \tag{2.9}$$

implies that

$$I_{2N+1} * R^{(N+1)} * \rho = I_{2N} \otimes R^{(N)} * \rho \tag{2.10}$$

for any state ρ on $\bigotimes_{j=0}^{N} \mathcal{H}_j$. Therefore this channel has two isometric dilations, namely

$$V_1 = V(I_{2N+1})V(R^{(N+1)}) \tag{2.11}$$

and

$$V_2 = V(I_{2N}) \otimes V(R^{(N)}) \tag{2.12}$$

where we have indicate with $V(R)$ the minimal Stinespring dilation of the channel whose Choi-Jamiołkowski operator is R. Isometry V_2 is minimal because it is the tensor product of two minimal dilations, in fact

$$\mathrm{Tr}_{out}[V_2 V_2^\dagger] = I_A \tag{2.13}$$

where "*out*" and "A" are shorthand labels to indicate all the output spaces and the ancillary spaces, respectively. Since any dilation is connected to the minimal one by an isometry, we have

$$V(I_{2N+1})V(R^{N+1}) = U(V(R^N) \otimes V(I_{2N})) \tag{2.14}$$

for some U. This in turn gives

$$V(R^{N+1}) = (V(I_{2N+1})^\dagger U V(I_{2N}))(V(R^N) \otimes I_{2N}) \tag{2.15}$$

since $V(I_{2N+1})^\dagger V(I_{2N+1}) = I_{2N+1}$. The required isometry is thus

$$U' := V(I_{2N+1})^\dagger U V(I_{2N}) \tag{2.16}$$

which acts only on the ancillary spaces. ■

2.2 Axiomatic approach to quantum maps

We now give the axiomatic presentation of admissible quantum maps. In the axiomatic approach we define the requirements that quantum maps must fulfill if we want to preserve the probabilistic interpretation described in section 2.1. In the meanwhile, we will also discuss the type-theoretic aspects of higher order quantum maps. It turns out the the correct way is to assign types to the deterministic objects.

Consider a transformation $\mathscr{C}$. The probabilistic interpretation requires convexity $\mathscr{C}(\sum_i p_i\rho_i) = \sum_i p_i\mathscr{C}(\rho_i)$, along with the property $\mathscr{C}(p\rho) = p\mathscr{C}(\rho)$. This two conditions together implies linearity of $\mathscr{C}$.

On the other hand, if we want to apply the transformation $\mathscr{C}$ locally on one side of a bipartite state, and still be assured to obtain a legitimate output, we must require also complete positivity of $\mathscr{C}$.

Let us now consider a map $\tilde{\mathscr{S}}$ from linear maps $\mathscr{T} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ to linear maps $\mathscr{T}' : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_3)$. In order to have compatibility with the probabilistic structure of quantum mechanics, we have to require two conditions on the map $\tilde{\mathscr{S}}$:

i) it is linear

ii) it "completely" preserves complete positivity (i.e. it preserves complete positivity with respect to any extension).

In other words, condition ii) requires that $\tilde{\mathscr{S}}$ must preserve complete positivity, also when applied locally on some bipartite map. More explicitly, we require that if

$$\mathscr{R} : \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_2) \otimes \mathcal{L}(\mathcal{H}_B) \tag{2.17}$$

is CP, then

$$\mathscr{R}' := (\tilde{\mathscr{S}} \otimes \tilde{\mathscr{I}})(\mathscr{R}) : \mathcal{L}(\mathcal{H}_0) \otimes \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_3) \otimes \mathcal{L}(\mathcal{H}_B) \tag{2.18}$$

is also CP.

An equivalent characterization of these conditions can be obtained considering the conjugate map $\mathscr{S}$ of $\tilde{\mathscr{S}}$, defined as follows:

$$\mathscr{S} := \mathfrak{C} \circ \tilde{\mathscr{S}} \circ \mathfrak{C}^{-1}, \tag{2.19}$$

which transforms the Choi-Jamiołkowski operator T of $\mathscr{T}$ into the Choi-Jamiołkowski operator T' of $\mathscr{T}'$. Linearity of $\tilde{\mathscr{S}}$ is equivalent to linearity of $\mathscr{S}$. On the other hand $\tilde{\mathscr{S}}$ satisfies condition ii) if and only if $\mathscr{S}$ is CP.

We are ready to give the recursive definition of quantum combs:

Definition 3 *A* quantum 1-comb S *on* $(\mathcal{H}_0, \mathcal{H}_1)$ *is the Choi-Jamiołkowski operator of an admissible map* $\mathscr{S}^{(1)} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_1)$*.*

$$S = \mathfrak{C}(\mathscr{S}^{(1)}). \tag{2.20}$$

For $N \geqslant 2$*, a* quantum N-comb *on* $(\mathcal{H}_0, \dots, \mathcal{H}_{2N-1})$ *is the Choi-Jamiołkowski operator of an admissible* N*-map, i.e. an admissible map transforming* $(N-1)$*-combs on* $(\mathcal{H}_1, \dots, \mathcal{H}_{2N-2})$ *into* 1*-combs on* $(\mathcal{H}_0, \mathcal{H}_{2N-1})$*.*

A special class of combs are the deterministic combs:

Definition 4 *A* deterministic 1*-comb is the Choi-Jamiołkowski operator of a channel. A deterministic* N*-comb* $S^{(N)}$ *is the Choi-Jamiołkowski operator of a deterministic* N*-map, i.e. a map* $\mathscr{S}^{(N)}$ *that transforms deterministic* $(N-1)$*-combs into deterministic* 1*-combs.*

Definition 5 *An* N*-comb* $R^{(N)}$ *is* probabilistic *if there is a deterministic* N*-comb* $S^{(N)}$ *such that* $R^{(N)} \leqslant S^{(N)}$*.*

From a type-theoretic point of view we are defining an infinite hierarchy of types with the following recursive definition:

$$\begin{cases} 1 \text{ is the type of a quantum channel,} \\ (M+1) := M \to 1 \end{cases} \tag{2.21}$$

We have the following algebraic characterization of N-combs:

Theorem 5 *A positive operator $S^{(N)}$ on $\bigotimes_{i=0}^{2N-1}\mathcal{H}_i$ is a deterministic N-comb if and only if the following identity holds:*

$$\begin{aligned} &\mathrm{Tr}_{2j-1}[S^{(j)}] = I_{2j-2}\otimes S^{(j-1)}, \quad 2\leqslant j\leqslant N \\ &\mathrm{Tr}_1[S^{(1)}] = I_0, \end{aligned} \tag{2.22}$$

where $S^{(j)}, 1\leqslant j\leqslant N-1$ are deterministic j-combs. Equivalently:

$$\begin{aligned} &S^{(j)} * I_{2j-1} = S^{(j-1)} * I_{2j-2}, \quad 2\leqslant j\leqslant N \\ &S^{(1)} * I_1 = I_0. \end{aligned} \tag{2.23}$$

We introduce two lemmas which simplify the proof of the theorem.

Lemma 4 *The set of positive operators $R^{(N)}$ such that $R^{(N)}\leqslant S^{(N)}$ for some $S^{(N)}$ satisfying Eq. (2.22) generates the positive cone in $\mathcal{L}(\bigotimes_{i=0}^{2N-1}\mathcal{H}_i)$*

Proof. The operator $J^{(N)} := I/(d_1\dots d_{2N})$ satisfies Eq. (2.22). Now, given any operator $T^{(N)}$ on $\bigotimes_{i=0}^{2N-1}\mathcal{H}_i$, there is a positive number λ such that $R^{(N)} := T^{(N)}/\lambda\leqslant J^{(N)}$, with obviously $T^{(N)}=\lambda R^{(N)}$. ■

Lemma 5 *Consider two positive operators $R_i^{(N)}, i=1,2,$ such that $R_i^{(N)}\leqslant S_i^{(N)}$ for some $S_i^{(N)}$ satisfying Eq. (2.22). Moreover, suppose that*

$$\mathrm{Tr}_{2N-1}[R_1^{(N)}] = \mathrm{Tr}_{2N-1}[R_2^{(N)}]. \tag{2.24}$$

Then, there exist a nonnegative operator $T^{(N)}$ such that $O_i^{(N)} := R_i^{(N)} + T^{(N)}$ satisfy Eq. (2.22) for $i=1,2$.

Proof. Since $R_1^{(N)}\leqslant S_1^{(N)}$ we can define $T^{(N)} := S_1^{(N)} - R_1^{(N)}\geqslant 0$, giving $O_1^{(N)} = S_1^{(N)}$. Due to Eq. (2.24) we have that $O_2^{(N)} := R_2^{(N)} + T^{(N)}$ satisfies

$$\mathrm{Tr}_{2N-1}[O_2^{(N)}] = \mathrm{Tr}_{2N-1}[O_1^{(N)}], \tag{2.25}$$

hence, satisfying also Eq. (2.22). ■

Proof of Theorem 5. We proceed by induction. For $N=1$ the theorem is trivial. An operator $S^{(1)}\in\mathcal{L}(\mathcal{H}_0\otimes\mathcal{H}_1)$ is the Choi-Jamiołkowski operator for a quantum channel from $\mathcal{L}(\mathcal{H}_0)$ to $\mathcal{L}(\mathcal{H}_1)$ if and only if it satisfies

$$\mathrm{Tr}_1[S^{(1)}] = I_0. \tag{2.26}$$

Now, we suppose that the thesis holds for $1\leqslant M\leqslant N$ and prove that it holds for $N+1$.

Sufficient condition. Suppose that Eq. (2.22) holds. We want to prove that the map $\mathfrak{C}(S^{(N+1)})$ applied to some deterministic N-comb $R^{(N)}$ yields the Choi-Jamiołkowski operator of a channel, i.e

$$[\mathfrak{C}^{-1}(S^{(N+1)})](R^{(N)}) * I_{2N+1} = I_0. \tag{2.27}$$

Condition (2.22), rewritten as

$$S^{(N+1)} * I_{2N+1} = S^{(N)} * I_{2N-1}, \tag{2.28}$$

yields

$$\begin{aligned}[\mathfrak{C}^{-1}(S^{(N+1)})](R^{(N)}) * I_{2N+1} &= S^{(N+1)}) * R^{(N)} * I_{2N+1} = \\ &= S^{(N)} * I_{2N-1} * R^{(N)} = \\ &= S^{(N)} * R^{(N-1)} * I_{2N-2} = I_0,\end{aligned} \tag{2.29}$$

where third equality holds by induction hypothesis on $R^{(N)}$, and the last equality holds because $S^{(N)} * R^{(N-1)}$ is a channel by hypothesis on $S^{(N)}$.
Necessary condition. Let $S^{(N+1)}$ be a deterministic $(N+1)$-comb, i.e. the corresponding map $\mathfrak{C}^{-1}(S^{(N+1)})$ transforms any deterministic N-comb $O^{(N)} \in \mathrm{comb}(\mathcal{H}_1, \dots, \mathcal{H}_{2N})$ into a channel $\mathfrak{C}^{-1}(S^{(N+1)})(O^{(N)}) \in \mathrm{comb}(\mathcal{H}_0, \dots, \mathcal{H}_{2N-1})$. Consider now a couple of probabilistic N-combs $R_1^{(N)}$ and $R_2^{(N)}$ on $\bigotimes_{i=1}^{2N} \mathcal{H}_i$ such that

$$R_1^{(N)} * I_{2N} = R_2^{(N)} * I_{2N}. \tag{2.30}$$

Since $R_i^{(N)}$ is probabilistic, there exists a deterministic N-comb $Q_i^{(N)}$ such that $R_i^{(N)} \leqslant Q_i^{(N)}$. By lemma 5 can find a $T^{(N)}$ such that $O_i := R_i^{(N)} + T^{(N)}$ is deterministic for $i = 1, 2$. Then the following holds

$$[\mathfrak{C}^{-1}(S^{(N+1)})](O_1^{(N)}) * I_{2N+1} = I_0 = [\mathfrak{C}^{-1}(S^{(N+1)})](O_2^{(N)}) * I_{2N+1}, \tag{2.31}$$

and consequently

$$[\mathfrak{C}^{-1}(S^{(N+1)})](R_1^{(N)}) * I_{2N+1} = [\mathfrak{C}^{-1}(S^{(N+1)})](R_2^{(N)}) * I_{2N+1}. \tag{2.32}$$

Now, take

$$R_2 := (R_1 * I_{2N}) * \sigma_{2N} \tag{2.33}$$

for some state σ on system $\mathcal{H}_{2N}$. Then we have

$$\begin{aligned}[\mathfrak{C}^{-1}(S^{(N+1)})](R_2^{(N)}) * I_{2N+1}1 &= S^{(N+1)} * R_2^{(N)} * I_{2N+1} \\ &= S^{(N+1)} * (R_1 * I_{2N}) * \sigma_{2N} * I_{2N+1} = \\ &= (R_1 * I_{2N}) * (S^{(N+1)} * \sigma_{2N} * I_{2N+1}) = \\ &= R_1 * (S^{(N)} \otimes I_{2N})\end{aligned} \tag{2.34}$$

where, in the last equality, we introduced

$$S^{(N)} := S^{(N+1)} * \sigma_{2N} * I_{2N+1} \tag{2.35}$$

and used the fact that $S^{(N)}$ is not acting on $\mathcal{H}_{2N}$. Combining with Eq. (2.32) we have

$$S^{(N+1)} * I_{2N+1} * R_1^{(N)} = (S^{(N)} \otimes I_{2N}) * R_1^{(N)}. \tag{2.36}$$

Since this equation holds for all probabilistic N-combs, which generate the whole cone of positive operators, we obtain

$$S^{(N+1)} * I_{2N+1} = S^{(N)} \otimes I_{2N}. \tag{2.37}$$

To conclude the proof of the theorem, we need to prove that $S^{(N)}$ is a deterministic N-comb, that is, the map $\mathfrak{C}^{-1}(S^{(N)})$ transforms a deterministic $(N-1)$-comb $R^{(N-1)}$ into a channel. Indeed, we have

$$\begin{aligned}[\mathfrak{C}^{-1}(S^{(N)})](R^{(N-1)}) * I_{2N-1} &= S^{(N)} * R^{(N-1)} * I_{2N-1} = \\ &= S^{(N+1)} * \sigma_{2N} * I_{2N+1} * R^{(N-1)} * I_{2N-1} = \\ &= S^{(N+1)} * (\sigma_{2N} * R^{(N-1)} * I_{2N-1}) * I_{2N+1}.\end{aligned} \tag{2.38}$$

Since the operator

$$R^{(N)} := \sigma_{2N} * R^{(N-1)} * I_{2N-1} = \sigma_{2N} \otimes R^{(N-1)} \otimes I_{2N-1} \tag{2.39}$$

satisfies

$$R^{(N)} * I_{2N} = (\sigma_{2N} \otimes R^{(N-1)} \otimes I_{2N-1}) * I_{2N} = I_{2N-1} \otimes R^{(N-1)} \tag{2.40}$$

for any state σ, by inductive hypothesis on $R^{(N-1)}$ it is a deterministic N-comb. Concluding, we have that $S^{(N)}$ is deterministic because

$$S^{(N)} * R^{(N-1)} * I_{2N-1} = S^{(N+1)} * R^{(N)} * I_{2N+1} = I_0 \tag{2.41}$$

where the last equality follows by hypothesis on $S^{(N+1)}$. ■

Corollary 1 *A deterministic N-comb is also the Choi-Jamiołkowski operator of an N-partite memory channel.*

Proof. Immediate from Theorems 4 and 5. ■

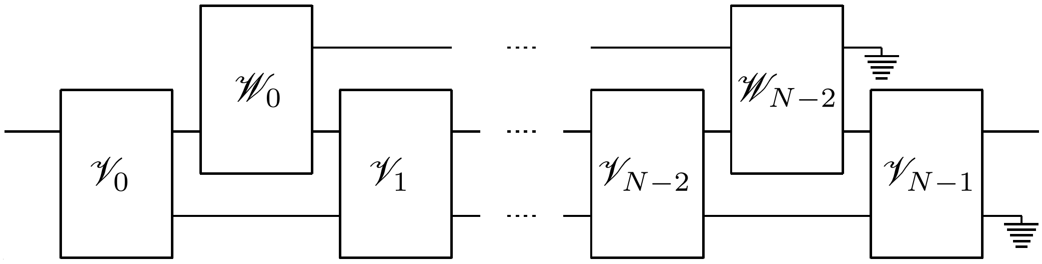


Figure 2.2

Theorem 6 (Realization of admissible N-maps) *For all N, any deterministic N-map $\tilde{\mathscr{S}}^{(N)}$ can be achieved by a physical scheme corresponding to the memory channel whose deterministic N-comb is $S^{(N)}$. Let $T^{(N-1)}$ be any $(N-1)$-comb in $\mathrm{comb}(\mathcal{H}_1, \ldots, \mathcal{H}_{2N-2})$. The transformation*

$$\tilde{\mathscr{S}}^{(N)} : \tilde{\mathscr{T}}^{(N-1)} \mapsto \tilde{\mathscr{T}}'^{(1)} = \tilde{\mathscr{S}}^{(N)} \left(\tilde{\mathscr{T}}^{(N-1)} \right) \tag{2.42}$$

can be achieved by connecting the two memory channels represented by $S^{(N)}$ and $T^{(N-1)}$ as in Fig. 2.2.

Proof. The statement is trivial for a deterministic 1-comb, which is a quantum channel. Now, by induction, suppose that the transformation $\tilde{\mathscr{T}}^{(N-1)}$ corresponding to a deterministic $N-1$ comb $T^{(N-1)}$ is realized by the $N-1$-partite memory channel having Choi-Jamiołkowski operator $T^{(N-1)}$. Let $W_0, i = 1, \ldots, N-2$ be the Choi-Jamiołkowski operators of the n interactions occurring in the memory channel, then $T^{(N-1)}$ can be expressed as

$$T^{(N-1)} = \bar{W}_{N-2} * W_{N-1} * \cdots * W_0, \tag{2.43}$$

where the Choi-Jamiołkowski operator $\bar{X}$ denotes the interaction described by X with the final ancilla traced out. By Corollary 1 also $S^{(N)}$ is the Choi-Jamiołkowski operator of a memory channel, then $S^{(N)}$ can be expressed as

$$S^{(N)} = \bar{V}_{N-1} * V_{N-2} * \cdots * V_0, \tag{2.44}$$

for suitable isometries V_i, where the link connects all the spaces representing ancillae. The application of $\mathscr{S}^{(N)} = \mathfrak{C} \circ \tilde{\mathscr{S}}^{(N)} \circ \mathfrak{C}^{-1}$ to $T^{(N-1)} = \mathfrak{C}(\tilde{\mathscr{T}}^{(N-1)})$ provides

$$\begin{aligned} \mathscr{S}^{(N)} \left(T^{(N-1)} \right) &= S^{(N)} * T^{(N-1)} \\ &= \bar{V}_{N-1} * \bar{W}_{N-2} * V_{N-2} * \cdots * W_0 * V_0, \end{aligned} \tag{2.45}$$

This proves that also the N-map $\tilde{\mathscr{S}}^{(N)}$ can be physically realized by a scheme corresponding to a memory channel. Clearly, Eq. (2.45) prescribes that the action of $\tilde{\mathscr{S}}^{(N)}$ on $\tilde{\mathscr{T}}^{(N-1)}$ corresponds to connecting the two memory channels associated to $S^{(N)}$ and $T^{(N-1)}$ as in Fig. 2.2. ■

Merging of teeth

Merging of adjacent teeth is just a regrouping of Hilbert spaces. Any N-comb on $(\mathcal{H}_0, \dots, \mathcal{H}_{2i-2}, \mathcal{H}_{2i-1}, \mathcal{H}_{2i}, \mathcal{H}_{2i+1}, \dots, \mathcal{H}_{2N-1})$ is also a $(N-1)$-comb on $(\mathcal{H}_0, \dots, \mathcal{H}_{2i-2} \otimes \mathcal{H}_{2i}, \mathcal{H}_{2i-1} \otimes \mathcal{H}_{2i+1}, \dots, \mathcal{H}_{2N-1})$, after the merging of i-th and $(i+1)$-th tooth.

It may be the case that an operator $S^{(N)} \in \mathcal{L}(\mathcal{H}_0 \otimes \cdots \otimes \mathcal{H}_{2N-1})$ is a N-comb compatible with more than one causal structure, i.e. we have that

$$S^{(N)} \in \mathrm{comb}(\mathcal{H}_0, \dots, \mathcal{H}_{2N-1}) \cap \mathrm{comb}(\mathcal{H}_{\pi(0)}, \mathcal{H}_{\pi(1)}, \dots, \mathcal{H}_{\pi(2N-1)}), \quad (2.46)$$

where π is a permutation of the indexes $0, \dots 2N-1$. A very simple example is any 2-comb obtained as convex combination of tensor product channels

$$S := \sum_j p_j R^{(j)} \otimes T^{(j)}. \quad (2.47)$$

These aspects will be discussed in section 2.2.2.

2.2.1 Maps from N-combs to M-combs

One can also consider admissible $(N \to M)$-maps, i.e. maps transforming N-combs into M-combs.

Definition 6 *An $(N \to M)$-*map *$\mathscr{S}^{(N \to M)}$ is a linear CP map transforming N-combs into M-combs. We say that $\mathscr{S}^{(N \to M)}$ is* deterministic *if it sends deterministic combs into deterministic combs.*

In this way we are introducing new types $N \to M$, which generally are not reducible to types $P \to 1$.

If M is of the form $M = P \to 1$ we can prove an analogue of the currying theorem for lambda calculus. Before presenting the theorem we need a preliminary discussion. Consider a pair of deterministic N- and M-combs, $R^{(N)}$ and $T^{(M)}$. We can define the product type

$$N \times M \quad (2.48)$$

to be the type of the tensor product $R^{(N)} \otimes T^{(M)}$. The natural definition of a map transforming $(N \times M)$-combs into channels requires linearity and the property of being locally CP. The latter is given by the following

Definition 7 *A map* $\mathscr{S}^{(N\times M\to 1)} : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_3, \mathcal{H}_4)$ *is* locally CP *if* $\mathscr{S} \otimes \mathscr{I}$ *is positive on positive tensor product operators* $R_1 \otimes R_2$ *with* $R_1 \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{K}_1)$ *and* $R_1 \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{K}_2)$.

We have the following

Lemma 6 *Consider a map* $\mathscr{S}^{(N\times M\to 1)} : \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_3, \mathcal{H}_4)$ *as in definition 7. Then,* $\mathscr{S}^{(N\times M\to 1)}$ *is also a CP map.*

Proof. The Choi-Jamiołkowski operator of $\mathscr{S}^{(N\times M\to 1)}$ is

$$S := \mathscr{S}^{(N\times M\to 1)} \otimes \mathscr{I}(|I\rangle\rangle\langle\langle I|_{\mathcal{H}_1\otimes\mathcal{H}_2,\mathcal{K}_1\otimes\mathcal{K}_2}). \tag{2.49}$$

Since $|I\rangle\rangle\langle\langle I|_{\mathcal{H}_1\otimes\mathcal{H}_2,\mathcal{K}_1\otimes\mathcal{K}_2} = |I\rangle\rangle\langle\langle I|_{\mathcal{H}_1,\mathcal{K}_1} \otimes |I\rangle\rangle\langle\langle I|_{\mathcal{H}_2,\mathcal{K}_2}$, the locally CP property of $\mathscr{S}^{(N\times M\to 1)}$ implies that the Choi-Jamiołkowski operator S is positive, hence the map $\mathscr{S}^{(N\times M\to 1)}$ is also CP. ■

Given this preliminary definitions, we can state the currying theorem for quantum maps.

Theorem 7 (Currying) *Let* $\mathscr{S}^{(N\to M+1)}$ *be a* $(N \to M+1)$*-map.* $\mathscr{S}^{(N\to M+1)}$ *is in one-to-one correspondence with a CP map* $\mathscr{S}^{(N\times M\to 1)}$ *that transforms tensor product operators* $R^{(N)} \otimes O^{(M)}$ *of* N*- and* M*-combs into* 1*-combs. Moreover,* $\mathscr{S}^{(N\to M+1)}$ *is deterministic if and only if* $\mathscr{S}^{(N\times M\to 1)}$ *transforms tensor product of deterministic combs into channels.*

Proof. Suppose that $\mathscr{S}^{(N\to M+1)}$ maps N-comb $R^{(N)}$ into $(M+1)$-comb $R'^{(M+1)} = \mathscr{S}^{(N\to M+1)}(R^{(N)})$. In term of Choi-Jamiołkowski operators we have

$$R'^{(M+1)} = \mathfrak{C}(\mathscr{S}^{(N\to M+1)}) * R^{(N)} \tag{2.50}$$

The map $\mathscr{R}'^{(M+1)}$ associated to $R'^{(M+1)}$ acts on M-comb $O^{(M)}$ as

$$\begin{aligned} \mathscr{R}'^{(M+1)}(O^{(M)}) &= \mathfrak{C}^{-1}(R'^{(M+1)})(O^{(M)}) = R'^{(M+1)} * O^{(M)} = \\ &= \mathfrak{C}(\mathscr{S}^{(N\to M+1)}) * R^{(N)} * O^{(M)} = \\ &= \mathfrak{C}(\mathscr{S}^{(N\to M+1)}) * (R^{(N)} \otimes O^{(M)}) \end{aligned} \tag{2.51}$$

The map $\mathscr{S}^{(N\to M+1)}$ thus induces a map $\mathscr{S}^{(N\times M\to 1)}$ on tensor product operators into 1-combs defined as

$$\mathscr{S}^{(N\times M\to 1)}(R^{(N)} \otimes O^{(M)}) := \mathfrak{C}(\mathscr{S}^{(N\to M+1)}) * (R^{(N)} \otimes O^{(M)}). \tag{2.52}$$

This map sends tensor product operators into a 1-comb, which is deterministic if $R^{(N)}$ and $O^{(N)}$ are deterministic. On the other hand, given a map $\mathscr{S}^{(N\times M\to 1)}$, we can define

$$\mathscr{S}^{(N\to M+1)}(R^{(N)}) := \mathfrak{C}(\mathscr{S}^{(N\times M\to 1)}) * R^{(N)}. \tag{2.53}$$

Clearly, if $\mathscr{S}^{(N\times M\to 1)}$ sends tensor product of deterministic combs into deterministic 1-combs, then $\mathscr{S}^{(N\to M+1)}$ is deterministic. ■

From a type-theoretic point of view, this theorem states that the following type isomorphism holds:

$$N \to (M \to 1) \cong N \times M \to 1. \tag{2.54}$$

The theorem can be easily generalized to $(N \to (M \to P))$-maps. Indeed, the essential part of the proof is the associativity of the link product of three Choi-Jamiołkowski operators, written for the special case in which the two operators $R^{(N)}$ and $O^{(M)}$ have no spaces in common

$$(\mathfrak{C}(\mathscr{S}^{(N\to(M\to P))}) * R^{(N)}) * O^{(M)} = \mathfrak{C}(\mathscr{S}^{(N\to(M\to P))}) * (R^{(N)} \otimes O^{(M)}), \tag{2.55}$$

which immediately leads to

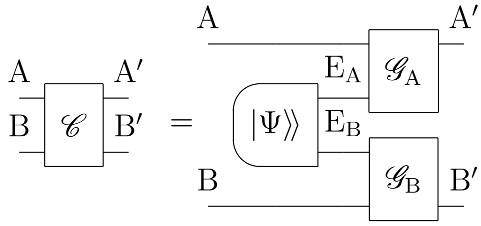
$$N \to (M \to P) \cong N \times M \to P. \tag{2.56}$$

Maps admissible on pairs, which, as seen, are in correspondence with higher order types, are not maps of type $N \to 1$, for any N. However, a pair $N \times M$ has globally $N + M$ teeth, so that every admissible map on $N \times M$ accepts as input an object with $N+M$ teeth. The hypothesis of *compatibility with remote connections* [43] requires that every map transforming an object with P teeth into channels is exactly a map of type $P \to 1$ for some arbitrary chosen ordering of the teeth. Under this hypothesis, the maps of type $N \to M$, begin also maps of type $N \times (M-1) \to 1$ by the currying theorem, must be maps of type $(N + M - 1) \to 1$ for some ordering. As we have seen, these are exactly the quantum combs. In other words, this hypothesis guarantees that the hierarchy of quantum maps collapses on the comb level. The price to be paid is the exclusion of genuinely higher order maps.

2.2.2 No-signaling channels

We now return to problem of characterizing combs satisfying more than one causal structure.

Definition 8 *The channel* $\mathscr{C} : \mathcal{L}(\mathrm{A}) \otimes \mathcal{L}(\mathrm{B}) \to \mathcal{L}(\mathrm{A}') \otimes \mathcal{L}(\mathrm{B}')$ *is "localizable" if it can be realized by local operations on* A *and* B *with a shared entangled ancilla* $|\Psi\rangle\rangle$ *on a couple of d-dimensional systems* $\mathrm{E_A}, \mathrm{E_B}$ *but without communication:*



. (2.57)

Definition 9 *A bipartite quantum channel* $\mathscr{C} : \mathcal{L}(\mathrm{A}) \otimes \mathcal{L}(\mathrm{B}) \to \mathcal{L}(\mathrm{A}') \otimes \mathcal{L}(\mathrm{B}')$ *is "*$\mathrm{A} \nrightarrow \mathrm{B}'$ *no-signaling" if* $\mathrm{Tr}_{\mathrm{A}'}[R_{\mathscr{C}}] = I_{\mathrm{A}} \otimes S_{\mathrm{BB}'}$ *where* $S_{\mathrm{BB}'}$ *is the Choi operator of some channel* $\mathscr{S} : \mathcal{L}(\mathrm{B}) \to \mathcal{L}(\mathrm{B}')$. *We say that* $\mathscr{C}$ *is " no-signaling" if it is both* $\mathrm{A} \nrightarrow \mathrm{B}'$ *no-signaling and* $\mathrm{B} \nrightarrow \mathrm{A}'$ *no-signaling.*

The following theorem holds

Theorem 8 *The following are equivalent:*

1. *The channel* $\mathscr{C} : \mathcal{L}(\mathrm{A}) \otimes \mathcal{L}(\mathrm{B}) \to \mathcal{L}(\mathrm{A}') \otimes \mathcal{L}(\mathrm{B}')$ *is no-signaling*

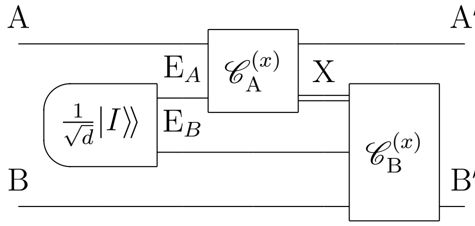
2. *There are equivalent d-dimensional quantum systems* $\mathrm{E}_{\mathrm{A}}, \mathrm{E}_{\mathrm{B}}$, *instruments* $\{\mathscr{C}_{\mathrm{A}}^{(x)}\}_{x \in \mathrm{X}}$ *and* $\{\mathscr{D}_{\mathrm{B}}^{(x)}\}_{x \in \mathrm{X}}$ *with outcome space* X, *and channels* $\mathscr{C}_{\mathrm{B}}^{(x)}, \mathscr{D}_{\mathrm{A}}^{(x)}$ *for each* $x \in \mathrm{X}$ *with*

$$\begin{aligned} \mathscr{C}_{\mathrm{A}}^{(x)} &: \mathcal{L}(\mathrm{A}) \otimes \mathcal{L}(\mathrm{E}_{\mathrm{A}}) \to \mathcal{L}(\mathrm{A}') \\ \mathscr{C}_{\mathrm{B}}^{(x)} &: \mathcal{L}(\mathrm{B}) \otimes \mathcal{L}(\mathrm{E}_{\mathrm{B}}) \to \mathcal{L}(\mathrm{B}') \\ \mathscr{D}_{\mathrm{B}}^{(x)} &: \mathcal{L}(\mathrm{B}) \otimes \mathcal{L}(\mathrm{E}_{\mathrm{B}}) \to \mathcal{L}(\mathrm{B}') \\ \mathscr{D}_{\mathrm{A}}^{(x)} &: \mathcal{L}(\mathrm{A}) \otimes \mathcal{L}(\mathrm{E}_{\mathrm{A}}) \to \mathcal{L}(\mathrm{A}') \end{aligned} \tag{2.58}$$

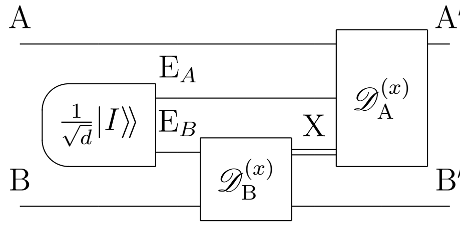
such that

$$\begin{aligned} \mathscr{C} &= \sum_{x \in X} \mathscr{C}_{\mathrm{B}}^{(x)} \circ \mathscr{C}_{\mathrm{A}}^{(x)} (d^{-1} |I\rangle\!\rangle\langle\!\langle I|_{\mathrm{E}_A \mathrm{E}_B}) \\ &= \sum_{x \in \mathrm{X}} \mathscr{D}_{\mathrm{A}}^{(x)} \circ \mathscr{D}_{\mathrm{B}}^{(x)} (d^{-1} |I\rangle\!\rangle\langle\!\langle I|_{\mathrm{E}_A \mathrm{E}_B}), \end{aligned} \tag{2.59}$$

namely, $\mathscr{C}$ *has the two equivalent circuit realizations*



, (2.60)

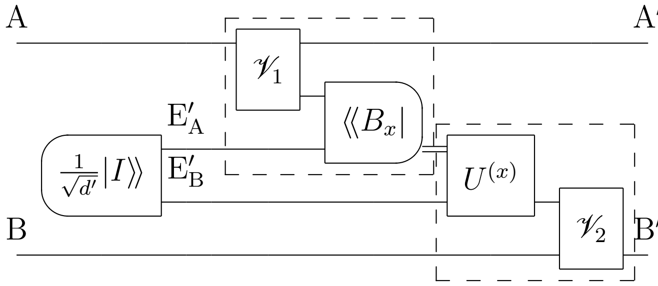


. (2.61)

Proof.

Proof of (1) $\Rightarrow$ (2).

$\mathscr{C}$ is B $\nrightarrow$ A$'$ no-signaling, therefore it can be realized as in Eq. (2.67), where E$'$ is a d'-dimensional system. This system can be teleported using the entangled state $\frac{1}{\sqrt{d'}}|I\rangle\rangle$ of systems $\mathrm{E}'_\mathrm{A}\mathrm{E}'_\mathrm{B}$, the Bell measurement $|B_x\rangle\rangle$ on systems E$'$ and E'_A, and classical communication of the outcome x followed by a controlled unitary U_x on system E'_B, corresponding to the circuit



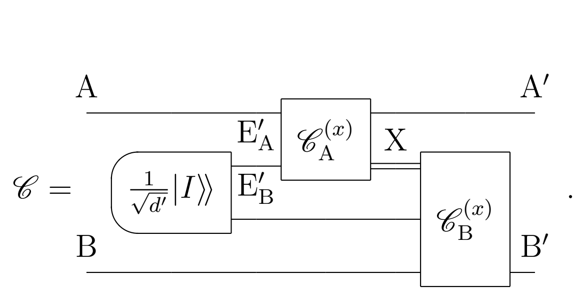
(2.62)

(the double wire represents the classical communication of the outcome x of the measurement).

The quantum operation $\mathscr{C}_\mathrm{A}^{(x)}$ and the channel $\mathscr{C}_\mathrm{B}^{(x)}$ are the grouped circuital elements in Eq. (2.62), and are given by

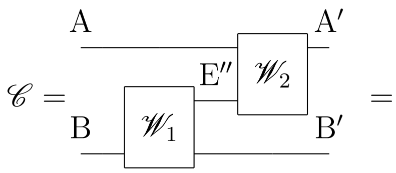
$$\begin{aligned}
\mathscr{C}_\mathrm{A}^{(x)}(\rho) &:= \langle\langle B_x|(\mathscr{V}_1 \otimes \mathscr{I}_{\mathrm{E}'_\mathrm{A}})(\rho)|B_x\rangle\rangle \\
\mathscr{C}_\mathrm{B}^{(x)}(\rho) &:= \mathscr{V}_2((U^{(x)} \otimes I_\mathrm{B})\rho(U^{(x)} \otimes I_\mathrm{B})^\dagger).
\end{aligned} \tag{2.63}$$

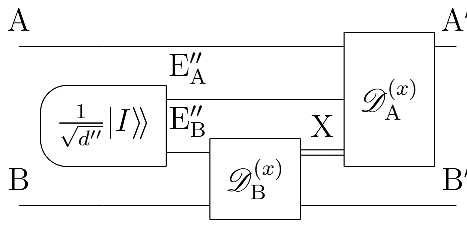
The final circuit is thus



(2.64)

Since the channel $\mathscr{C}$ is also $\mathrm{A} \nrightarrow \mathrm{B}'$ no-signaling, the same argument gives:





(2.65)

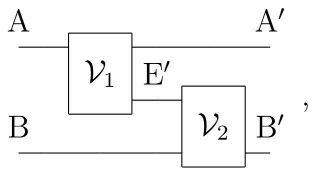
with $\mathscr{D}_{\mathrm{A}}^{(x)}$ and $\mathscr{D}_{\mathrm{B}}^{(x)}$ given by

$$\begin{aligned} \mathscr{D}_{\mathrm{B}}^{(x)}(\rho) &:= \langle\!\langle B_x | (\mathscr{W}_1 \otimes \mathscr{I}_{\mathrm{E}''_{\mathrm{B}}})(\rho) | B_x \rangle\!\rangle \\ \mathscr{D}_{\mathrm{A}}^{(x)}(\rho) &:= \mathscr{W}_2((U^{(x)} \otimes I_{\mathrm{A}})\rho(U^{(x)} \otimes I_{\mathrm{A}})^\dagger). \end{aligned} \tag{2.66}$$

We obtain the statement by defining $\mathrm{E_A}$ and $\mathrm{E_B}$ as d-dimensional systems, where $d := \max\{d', d''\}$, and embedding E'_{J} and $\mathrm{E}''_{\mathrm{J}}$ in $\mathrm{E_J}$, for $\mathrm{J} = \mathrm{A}, \mathrm{B}$.

Proof of (2) $\Rightarrow$ (1).

Suppose that $\mathscr{C}$ admits the realization circuit given in Eq. (2.60). We can group $\mathrm{E_B}$ and X in the composite system E'. Then $\mathscr{C}$ is also of the form of



, (2.67)

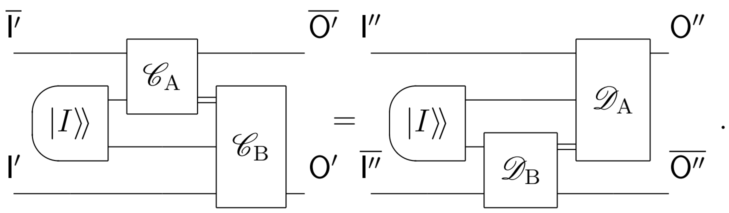
thus being $\mathrm{B} \nrightarrow \mathrm{A}'$ no-signaling, as proved in Ref. [43, 44]. In the same way, exploiting the second realization circuit in Eq. (2.61), one can prove that $\mathscr{C}$ is also $\mathrm{A} \nrightarrow \mathrm{B}'$ no-signaling. ■

Theorem 8 shows that the most general no-signaling channel differs from a localizable channel because it also admits a single round of classical communication, with the constraint that it must be possible to implement the channel exploiting communication in either directions.

For a multipartite channel satisfying two different no-signaling conditions, an analog of Theorem 8 holds. In fact, let us consider a a channel $\mathscr{C}$ with input systems labeled by a set of indices I and output systems labeled by a set O. Suppose that $\mathscr{C}$ satisfies the following no-signaling conditions

$$\begin{aligned} \mathrm{Tr}_{\mathsf{O}'}[R_{\mathscr{C}}] &= I_{\mathsf{I}'} \otimes S_{\overline{\mathsf{O}'} \cup \overline{\mathsf{I}'}} \\ \mathrm{Tr}_{\mathsf{O}''}[R_{\mathscr{C}}] &= I_{\mathsf{I}''} \otimes T_{\overline{\mathsf{O}''} \cup \overline{\mathsf{I}''}} \end{aligned} \quad , \tag{2.68}$$

for certain subsets $\mathsf{I}', \mathsf{I}'' \subseteq \mathsf{I}$ and $\mathsf{O}', \mathsf{O}'' \subseteq \mathsf{O}$, where $\overline{\mathsf{S}}$ represents the set complement of S, and for suitable Choi-Jamiołkowski operators S and T. Following the proof of Theorem 8 we can show that two circuits realizing $\mathscr{C}$ are



. (2.69)

In general the subsets $\mathsf{I}', \mathsf{I}''$ are not a partition of I. In this case we have that the circuits cannot be realized partitioning the systems between the two local parties A and B. In particular the input systems in $\overline{\mathsf{I}'} \cap \overline{\mathsf{I}''}$ are always assigned to the party which sends the classical message, and input systems in $\mathsf{I}' \cap \mathsf{I}''$ are assigned to the party which receives the classical message (and similarly for output systems). One can also consider more complex scenarios, i. e. channels with more than two no-signaling conditions of the kind in Eq. (2.68), or channels with nested conditions, for example when the Choi-Jamiołkowski operators S and T in Eq. (2.68) satisfy no-signaling conditions on their own. However the analysis of the classical communication required in these cases is complicated, and is an open problem.

No-signaling and localizable channels enjoy the remarkable property expressed in the following

Theorem 9 (Semigroupoid property) *Consider bipartite channels*

$$\begin{aligned} \mathscr{S} &\in \mathcal{L}(\mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_1), \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3)), \\ \mathscr{T} &\in \mathcal{L}(\mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3), \mathcal{L}(\mathcal{H}_4 \otimes \mathcal{H}_5)). \end{aligned} \tag{2.70}$$

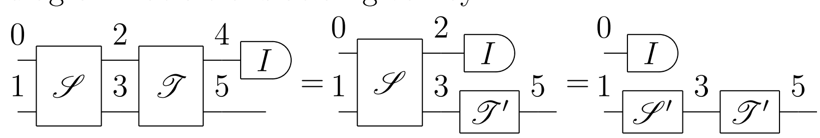
If they are both no-signaling (localizable), then their composition $\mathscr{T} \circ \mathscr{S}$ is no-signaling (localizable).

Proof.

(No-signaling case). The composition is no-signaling because it is no-signaling in both directions, in fact

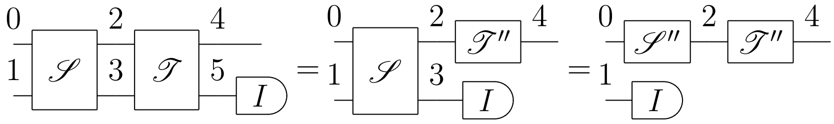
$$\begin{aligned} S * T * I_4 = S * T' * I_2 = S' * T' * I_0 = (S' * T') \otimes I_0, \\ S * T * I_5 = S * T'' * I_3 = S'' * T'' * I_1 = (S'' * T'') \otimes I_1. \end{aligned} \tag{2.71}$$

with diagrammatic translation given by



(2.72)

and

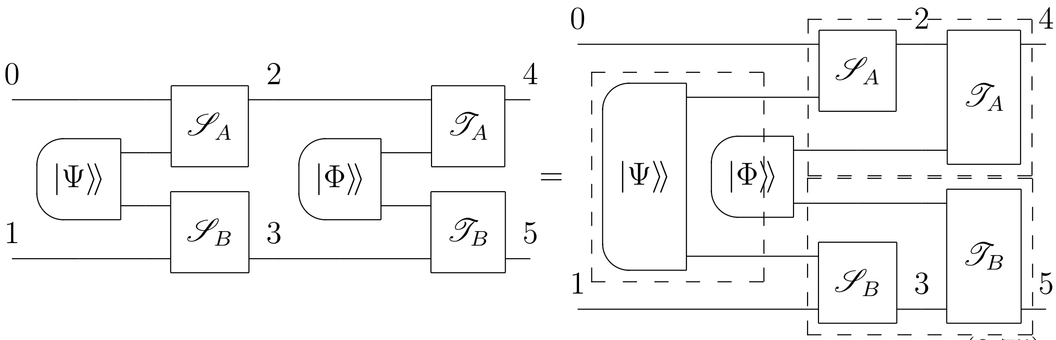


(2.73)

(Localizable case). The composition is

$$\begin{aligned}\mathfrak{C}(\mathscr{T}\circ\mathscr{S}) &= (R_\Psi * (S_A\otimes S_B)) * (R_\Phi * (T_A\otimes T_B)) = \\ &= R_\Psi * R_\Phi * S_A * S_B * T_A * T_B = \\ &= (R_\Psi * R_\Phi) * (S_A * T_A) * (S_B * T_B) = \\ &= (R_\Psi \otimes R_\Phi) * ((S_A * T_A)\otimes(S_B * T_B)),\end{aligned} \tag{2.74}$$

or, diagrammatically



(2.75)

which has the structure of a localizable channel. ■

We used the term "semigroupoid" instead of semigroup because a in a semigroup we require the every pair of elements are composable. Maps, on the other hand, can be composed only if input and output spaces match.

Pairs of channels (of type $N\times M$), are a special case of localizable channels, without the entangled resource. Hence, they are obviously no-signaling. A map admissible on no-signaling is automatically admissible on pairs. The converse is proved exploiting a result in [49], which we present here in a slightly different form.

Theorem 10 *Consider an Hermitian preserving map*

$$\mathscr{X}\in\mathcal{L}(\mathcal{L}(\mathrm{A}\otimes\mathrm{B}\otimes\mathrm{C}),\mathcal{L}(\mathrm{A}'\otimes\mathrm{B}'\otimes\mathrm{C}')) \tag{2.76}$$

along with its Choi-Jamiołkowski operator $X=\mathfrak{C}(\mathscr{X})$. Then, the following are equivalent

1.

$$X\in\mathrm{comb}(\mathrm{A},\mathrm{A}')\otimes\mathrm{comb}(\mathrm{B},\mathrm{B}',\mathrm{C},\mathrm{C}')\cap\mathrm{comb}(\mathrm{ABC},\mathrm{A}'\mathrm{B}'\mathrm{C}') \tag{2.77}$$

2.

$$X \in \mathrm{comb}(\mathrm{A},\mathrm{A}',\mathrm{B},\mathrm{B}',\mathrm{C},\mathrm{C}') \cap \mathrm{comb}(\mathrm{B},\mathrm{B}',\mathrm{A},\mathrm{A}',\mathrm{C},\mathrm{C}') \cap \\ \cap\, \mathrm{comb}(\mathrm{B},\mathrm{B}',\mathrm{C},\mathrm{C}',\mathrm{A},\mathrm{A}'). \tag{2.78}$$

Proof. (1) $\Rightarrow$ (2) Obvious.

(2) $\Rightarrow$ (1)

Let $\{E^i\}_{i\in\mathsf{J}}$ be an operator basis for $\mathcal{L}(\mathrm{A}\otimes\mathrm{A}')$. Then one can find operators X^i such that

$$X = \sum_{i\in\mathsf{J}} E^i_{\mathrm{AA}'} \otimes X^i_{\mathrm{BB}'\mathrm{CC}'} \tag{2.79}$$

Let us consider a dual set of operators $\tilde{E}^j$, such that

$$\tilde{E}^j * E^i = \delta_{ij} k_j \tag{2.80}$$

for some real numbers k_j. Exploiting the normalization of X we have that

$$(X * I_{\mathrm{C}'}) * \tilde{E}^j_{\mathrm{AA}'} = (I_C \otimes Y_{\mathrm{AA}'\mathrm{BB}'}) * \tilde{E}^j_{\mathrm{AA}'} = I_{\mathrm{C}} \otimes (Y_{\mathrm{AA}'\mathrm{BB}'} * \tilde{E}^j_{\mathrm{AA}'}), \tag{2.81}$$

for some operator Y. On the other hand, by the properties of link product we can also write

$$(X * \tilde{E}^j_{\mathrm{AA}'}) * I_{\mathrm{C}'} = \left[\left(\sum_{i\in\mathsf{J}} E^i_{\mathrm{AA}'} \otimes X^i_{\mathrm{BB}'\mathrm{CC}'}\right) * \tilde{E}^j_{\mathrm{AA}'}\right] * I_{\mathrm{C}'} = \\ = \sum_{i\in\mathsf{J}} E^i_{\mathrm{AA}'} * \tilde{E}^j_{\mathrm{AA}'} \otimes X^i_{\mathrm{BB}'\mathrm{CC}'} * I_{\mathrm{C}'} = k_j X^j_{\mathrm{BB}'\mathrm{CC}'} * I_{\mathrm{C}'}. \tag{2.82}$$

Posing $R^j_{\mathrm{BB}'} := k_j^{-1}(Y_{\mathrm{AA}'\mathrm{BB}'} * \tilde{E}^j_{\mathrm{AA}'})$ we can conclude that

$$X^j_{\mathrm{BB}'\mathrm{CC}'} * I_{\mathrm{C}'} = I_{\mathrm{C}} \otimes R^j_{\mathrm{BB}'}. \tag{2.83}$$

with $R^j_{\mathrm{BB}'} * I_{\mathrm{B}'}$ proportional to I_{B}. Thus we have proved that each X^j is proportional to an element in $\mathrm{comb}(\mathrm{B},\mathrm{B}',\mathrm{C},\mathrm{C}')$. Now, choosing $\mathsf{J}' \subset \mathsf{J}$ such that $\{X^j\}_{j\in\mathsf{J}'}$ is a maximally linearly independent subset, we can write

$$X = \sum_{j\in\mathsf{J}'} Z^j_{\mathrm{AA}} \otimes X^j_{\mathrm{BB}'\mathrm{CC}'} \tag{2.84}$$

for suitable operators Z^j. The same argument proves that each Z^j is proportional to an element in $\mathrm{comb}(\mathrm{A},\mathrm{A}')$.

■

We can attribute to the maps transforming pairs into channels the type $N \times M \to 1$.

The notation used so far for types is useful to mimic the recursive definitions of quantum combs, but it does not allow to refer to the *single* teeth of a comb and to their causal ordering. This is a serious drawback if we want to discuss causal properties, as in the definition of no-signaling channels. A notation which overcome this drawback can be introduced as follows: we attribute a symbol (such as A, B, $\ldots$) to every pair of input-output spaces (i.e. to every teeth), and represent a causal structure as an ordered string of such symbols.

For example, the set of combs with two teeth and with fixed causal ordering is represented by the string

$$AB. \tag{2.85}$$

To express the fact that a comb satisfies different causal orderings we introduce an intersection symbol, in such a way that, for example, the set of no-signaling channels (on the same Hilbert spaces), is represented by

$$AB \cap BA. \tag{2.86}$$

For pairs of (causally ordered) types we use the product symbol introduced before, for example a pair of channels is

$$A \times B. \tag{2.87}$$

The set of pairs of channels is a proper subset of the no-signaling, $A \times B \subset AB \cap BA$.

The set of no-signaling channels is maximal in the following sense: let $(AB \cap BA)^{\perp}$ be the set of maps admissible on no-signaling channels. The admissibility domain of these set of maps is the set of combs on which the considered maps are admissible. We can indicate pictorially this domain with $(AB \cap BA)^{\perp\perp}$. In principle, it can be larger than the set no-signaling channels, but it turns out that $AB \cap BA = (AB \cap BA)^{\perp\perp}$.

Consider now the set $A \times BC$, pairs consisting of a channel and a 2-comb. Clearly one has that

$$A \times BC \subset ABC \cap BAC \cap BCA \tag{2.88}$$

With notation in mind, we see that Theorem 10 proves that

$$A \otimes BC = (A \times BC)^{\perp\perp} = (ABC \cap BAC \cap BCA) \tag{2.89}$$

From this result follows as a particular case that the admissibility domain of the set of maps admissible on pairs is actually the set of no-signaling channels,

$$(A \times B)^{\perp\perp} = AB \cap BA. \tag{2.90}$$

It is an open problem to characterize the maps

$$(A \otimes B)^{\perp} = (A \times B)^{\perp} = (AB \cap BA)^{\perp} \tag{2.91}$$

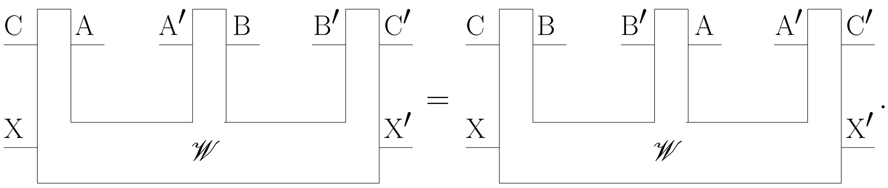
in terms of linear conditions, or in terms of a universal set.

2.3 No-switch theorem

A typical example of map which is admissible on no-signaling combs is the switch map $\mathscr{W}$. It takes as input a no-signaling comb and outputs a bipartite channel

$$\mathscr{W} : \mathrm{comb}(\mathrm{A}, \mathrm{A}', \mathrm{B}, \mathrm{B}') \cap \mathrm{comb}(\mathrm{B}, \mathrm{B}', \mathrm{A}, \mathrm{A}') \longrightarrow \mathrm{comb}(\mathrm{XC}, \mathrm{X}'\mathrm{C}') \tag{2.92}$$

where X, X' are qubit systems. Pictorially we have



(2.93)

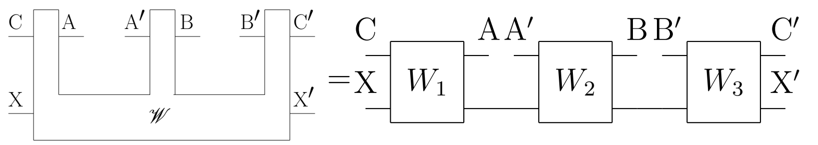
In the simplest version of the switch, $\mathrm{A}, \mathrm{A}', \mathrm{B}, \mathrm{B}', \mathrm{C}, \mathrm{C}'$ are quantum systems with the same dimension. The map $\mathscr{W}$ is defined as follows: on a pair of combs (F, G) representing channels $(\mathscr{F}, \mathscr{G})$ (i.e. on an object of *pair* type), it gives the composition $\mathfrak{C}(\mathscr{F} \circ \mathscr{G})$ or $\mathfrak{C}(\mathscr{G} \circ \mathscr{F})$, depending whether the control qubit is $|0\rangle$ or $|1\rangle$ respectively. The map is then extended by linearity to every state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ on the control qubit.

$$\mathscr{W}(F \otimes G) = |0\rangle\,\langle 0| \otimes \mathfrak{C}(\mathscr{F} \circ \mathscr{G}) + |1\rangle\,\langle 1| \otimes \mathfrak{C}(\mathscr{G} \circ \mathscr{F}) \tag{2.94}$$

This map is clearly admissible on pairs. By linearity we can now extend it to the set of no-signaling channels. A natural question is to ask whether the switch can be implemented as quantum comb. The answer is negative as proved by the following [54]

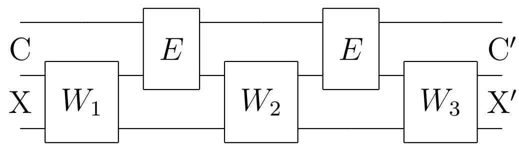
Theorem 11 (No-switch of boxes) *The map $\mathscr{W}$ is not realizable as deterministic quantum comb.*

Proof. Suppose by absurd that $\mathscr{W}$ is realizable as deterministic quantum comb, i.e. it admits a circuital realization as follows



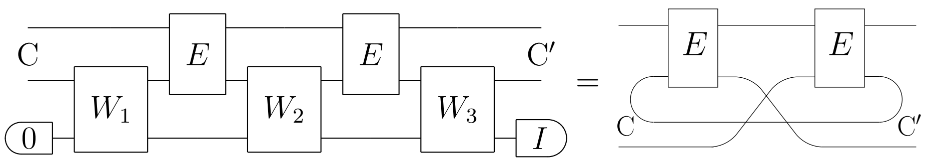
(2.95)

Then, we can apply it to a couple of linked swap gates as follows:



(2.96)

and obtain a properly normalized quantum channel. But, for $|\psi\rangle = |0\rangle$, the definition of the switch map leads to



(2.97)

(where the control qubit X′ has been traced away). The right hand side of this equation does not satisfy the normalization conditions for an admissible quantum comb. This contradiction implies that no such realization as Eq. (2.95) exists. ■

This theorem shows that the set of admissible maps on no-signaling channels is strictly larger then the set of quantum combs. It is not known how to characterize this set of maps in terms of a universal set. It is proposed that the following holds

Conjecture 1 *The set of admissible transformations on bipartite no-signaling channels is generated by the switch map $\mathscr{W}$ and the quantum combs.*

More generally, we can ask whether there is a universal set which generates every possible higher-order map, including all maps admissible on various sets of no-signaling combs. As we will see, this is relevant in order to extend the quantum lambda calculus to the whole hierarchy.

Chapter 3

Applications of higher order quantum maps

The quantum comb formalism and the conceptual apparatus of higher order quantum maps, while very useful for theoretical considerations, is also suitable for the analysis and resolution of concrete problems. In this chapter we describe three application of the theory introduced in the first part of the thesis.

3.1 Quantum learning of unitary transformations

A quantum memory would be an invaluable resource for Quantum Information Technology, and extensive experimental effort is in progress for its realization [7, 8, 9]. On a quantum memory one can store any unknown quantum state for later use. Can we exploit a quantum memory also to store an unknown quantum transformation, without keeping the device producing it?

Consider the scenario in which Alice puts at Bob's disposal N uses of a black box implementing an unknown unitary transformation U. Today Bob is allowed to exploit such uses at his convenience, running an arbitrary quantum circuit that makes N calls to Alice's black box. Tomorrow, however, Alice will withdraw the black box and ask Bob to reproduce U on a new input state $|\psi\rangle$ unknown to him. Alice will then test the output produced by Bob, and assign a score that is as higher as the output is closer to $U|\psi\rangle$. We refer to this two-party scenario as to an instance of *quantum learning* of the unitary U from a finite set of examples. More generally, Alice can ask Bob to reproduce U more than once, i.e. to produce $M \geq 1$ copies of U. In this case it is important to assess how the performance decays with the number of copies required, as in the case of quantum cloning [10].

Let us consider first the case of single input and output copies. Clearly, the only thing that Bob can do today is to use the black box on a known (generally entangled) input state $|\varphi\rangle$. After that, what remains available is the output state $|\varphi_U\rangle = (U \otimes I)|\varphi\rangle$, which Bob can store in a quantum memory. When Alice will provide the input state $|\psi\rangle$, Bob will send $|\psi\rangle$ and $|\varphi_U\rangle$ to an optimal *retrieving machine*, which extracts U and applies it to $|\psi\rangle$. When $N > 1$ input copies are available, Bob has also to find the best storing strategy: he can e.g. opt for a *parallel strategy* where U is applied on N different systems, yielding $(U^{\otimes N} \otimes I)|\varphi\rangle$, or for a *sequential strategy* where U is applied N times on the same system, generally alternated with other known unitaries, yielding $(UV_{N-1} \dots V_2 U V_1 U \otimes I)|\varphi\rangle$. The most general storing strategy is described by a *quantum circuit board*, i.e. a quantum network with open slots in which the input copies can be inserted [11]. In summary, finding the optimal quantum learning means finding the optimal storing board and the optimal retrieving machine.

An alternative to coherent retrieval is to estimate U, to store the outcome in a classical memory, and to perform the estimated unitary on the new input state. This incoherent strategy has the double advantage of avoiding the expensive use of a quantum memory, and of allowing one to reproduce U an unlimited number of times with constant quality. However, incoherent strategies are typically suboptimal for the similar task of quantum cloning [10], and this would suggest that a coherent retrieval achieves better performances. Surprisingly enough, we find that the incoherent strategies already achieve to ultimate performance of quantum learning. We analyzed the case in which U is a completely unknown unitary in a group G, and we found that the performances of the optimal retrieving machine are equal to those of optimal estimation. For an unknown qubit unitary with N input copies the maximum fidelity approaches unit asymptotically as $1/N^2$ and is achieved using N memory qubits. Our result can be also extended to solve the problem of *optimal inversion* of the unknown U, in which instead of performing U, Bob is asked to perform its inverse $U^\dagger$. In this case, our result provides the optimal approximate re-alignment of reference frames for the quantum communication scenario recently considered in Ref. [12].

We tackle the optimization of learning starting from the case $M = 1$. Referring to Fig. 3.1, we label the Hilbert spaces of quantum systems according to the following sequence: $(\mathcal{H}_{2n+1})_{n=0}^{N-1}$ are the inputs for the N examples of U, and $(\mathcal{H}_{2n+2})_{n=0}^{N-1}$ are the corresponding outputs. We denote by $\mathcal{H}_i = \bigotimes_{n=0}^{N-1} \mathcal{H}_{2n+1}$ ($\mathcal{H}_o = \bigotimes_{n=0}^{N-1} \mathcal{H}_{2n+2}$) the Hilbert spaces of all inputs

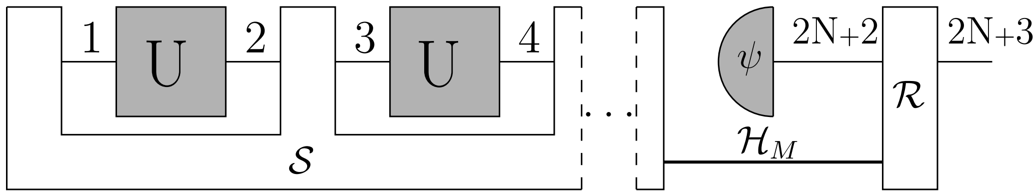


Figure 3.1: The learning process is described by a quantum comb R (in white) representing the quantum circuit board, in which the N uses of an oracle U are plugged, along with the state $|\psi\rangle$ (in gray). The wires represent the input-output Hilbert spaces. The output of the first comb is fed in a quantum memory, which later use in the retrieval stage is connected to the input of the second comb.

(outputs) of the N examples. Alice's input state $|\psi\rangle$ belongs to $\mathcal{H}_{2N+2}$, and the output state finally produced by Bob belongs to $\mathcal{H}_{2N+3}$. All spaces $\mathcal{H}_n$ considered here are $d-$dimensional, except the spaces $\mathcal{H}_0$ and $\mathcal{H}_{2N+1}$ which are one-dimensional, and are introduced just for notational convenience. The comb of the whole learning process is a positive operator L on the tensor of all Hilbert spaces, and satisfies the normalization condition [11]:

$$\mathrm{Tr}_{2k+1}[L^{(k)}] = I_{2k} \otimes L^{(k-1)} \qquad k = 0, 1, \ldots, N+1 \tag{3.1}$$

where $L^{(N+1)} = L$, $L^{(-1)} = 1$, and $L^{(k)}$ is a positive operator on the spaces $(\mathcal{H}_n)_{n=0}^{2k+1}$. When the N examples are connected with the learning board, Bob obtains a channel $\mathscr{C}_U$ with Choi operator given by

$$C_U = L * |U\rangle\rangle\langle\langle U|^{\otimes N} = \mathrm{Tr}_{i,o}\left[L\left(I_{2N+3} \otimes I_{2N+2} \otimes (|U\rangle\rangle\langle\langle U|^{\otimes N})^T\right)\right], \tag{3.2}$$

according to the definition of link product in Eq. (2).

As the figure of merit we maximize the fidelity of the output state $\mathscr{C}_U(|\psi\rangle\langle\psi|)$ with the target state $U|\psi\rangle\langle\psi|U^\dagger$, uniformly averaged over all pure states $|\psi\rangle$ and all unknown unitaries U in the group G. Apart from irrelevant constants, such optimization coincides with the maximization of the average channel fidelity between $\mathscr{C}_U$ and the target unitary, which is nothing but the fidelity between the Choi states C_U/d and $|U\rangle\rangle\langle\langle U|/d$:

$$F = \frac{1}{d^2} \int_G \langle\langle U|\langle\langle U^*|^{\otimes N} L |U^*\rangle\rangle^{\otimes N}|U\rangle\rangle \, \mathrm{d}\,U , \tag{3.3}$$

U^* being the complex conjugate of U in the computational basis, and $\mathrm{d}\,U$ denoting the normalized Haar measure. From the expression of F it is easy to prove that there is no loss of generality in requiring the commutation

$$[L, U_{2N+3} \otimes V^*_{2N+2} \otimes U_o^{*\otimes N} \otimes V_i^{\otimes N}] = 0 . \tag{3.4}$$

where U and V are arbitrary elements of G. Combining Eqs. (3.1) and (3.4) we then obtain

$$[L^{(N)}, U_o^{*\otimes N} \otimes V_i^{\otimes N}] = 0 \ . \tag{3.5}$$

Lemma 7 (Optimality of parallel storage) *The optimal storage of U can be achieved by applying $U^{\otimes N} \otimes I^{\otimes N}$ on a suitable input state $|\varphi\rangle \in \mathcal{H}_o \otimes \mathcal{H}_i$.*

Proof. The learning board $\mathscr{L}$ is obtained by connection of the storing board $\mathscr{S}$ with the retrieving channel $\mathscr{R}$, whence $L = R * S$. Denoting by $\mathcal{H}_M$ the Hilbert space of the quantum memory, $\mathscr{R}$ is a channel from $(\mathcal{H}_{2N+2} \otimes \mathcal{H}_M)$ to $\mathcal{H}_{2N+3}$, and satisfies the normalization condition

$$I_{2N+3} * R = I_{2N+2} \otimes I_M. \tag{3.6}$$

Using this fact, one gets

$$\begin{aligned} \mathrm{Tr}_{2N+3}[L] \equiv I_{2N+3} * L = (I_{2N+3} * R) * S = \\ = (I_{2N+2} \otimes I_M) * S = I_{2N+2} \otimes \mathrm{Tr}_M[S], \end{aligned} \tag{3.7}$$

which compared with Eq. (3.1) for $k = N+1$ implies $\mathrm{Tr}_M[S] = L^{(N)}$. Now, without loss of generality we take the storing board $\mathscr{S}$ to be a sequence of isometries, which implies that S is rank-one: $S = |\Phi\rangle\!\rangle\langle\!\langle\Phi|$. With this choice, the state S/d^N is a purification of $L^{(N)}/d^N$. Again, one can choose w.l.o.g. S/d^N to be a state on $(\mathcal{H}_o \otimes \mathcal{H}_i) \otimes (\mathcal{H}'_o \otimes \mathcal{H}'_i)$, with $\mathcal{H}'_o \simeq \mathcal{H}_o$ and $\mathcal{H}'_i = \mathcal{H}_i$, and assume $|\Phi\rangle\!\rangle = |L^{(N)\frac{1}{2}}\rangle\!\rangle$. Taking $V = I$ in Eq. (3.5) and using Eq. (1.5) we get

$$\left(U_o^{\otimes N} \otimes I_{i,o',i'}\right) |\Phi\rangle\!\rangle = \left(I_{o,i} \otimes U_{o'}^{T\otimes N} \otimes I_{i'}\right) |\Phi\rangle\!\rangle. \tag{3.8}$$

When the examples of U are connected to the storing board, the output is the state $\rho_U = S * |U\rangle\!\rangle\langle\!\langle U|_{o,i}^{\otimes N}$. Using the above relation we find that ρ_U is the projector on the state $|\varphi_U\rangle = (U_{o'}^{\otimes N} \otimes I_{i'})|\varphi\rangle$, where $|\varphi\rangle = \langle\!\langle I^{\otimes N}|_{o,i}|\Phi\rangle\!\rangle \in \mathcal{H}_{o'} \otimes \mathcal{H}_{i'}$. This proves that every storing board gives the same output of a parallel scheme. ■

Optimizing learning is then reduced to finding the optimal input state $|\varphi\rangle$ and the optimal retrieving channel $\mathscr{R}$. The fidelity can be computed substituting $L = R * S$ in Eq. (3.3), and using the relation

$$\begin{aligned} \langle\!\langle U|\langle\!\langle U^*|^{\otimes N}(R * S)|U\rangle\!\rangle|U^*\rangle\!\rangle^{\otimes N} = \\ = \langle\!\langle U|R|U\rangle\!\rangle * \langle\!\langle U^*|^{\otimes N} S|U^*\rangle\!\rangle^{\otimes N} = \langle\!\langle U|R|U\rangle\!\rangle * \rho_U, \end{aligned} \tag{3.9}$$

which gives

$$F = \frac{1}{d^2}\int_G \langle\!\langle U|\langle\varphi_U^*|R|U\rangle\!\rangle|\varphi_U^*\rangle \ \mathrm{d}\,U. \tag{3.10}$$

Lemma 8 (Optimal states for storage) *The optimal input state for storage can be taken of the form*

$$|\varphi\rangle = \bigoplus_j \sqrt{\frac{p_j}{d_j}} |I_j\rangle\rangle \in \widetilde{\mathcal{H}} \;, \tag{3.11}$$

where p_j are probabilities, $\widetilde{\mathcal{H}} = \bigoplus_j (\mathcal{H}_j \otimes \mathcal{H}_j)$ is a subspace of $\mathcal{H}_o \otimes \mathcal{H}_i$ carrying the representation $\widetilde{U} = \bigoplus_j (U_j \otimes I_j)$, I_j being the identity in $\mathcal{H}_j$, and the index j labeling the irreducible representations U_j contained in the decomposition of $U^{\otimes N}$.

Proof. Using Eqs. (1.5) and (3.5) it is possible to show that the local state $\rho = \mathrm{Tr}_i[|\varphi\rangle\langle\varphi|]$ is invariant under $U^{\otimes N}$. Decomposing $U^{\otimes N}$ into irreducible representations (irreps) we have $U^{\otimes N} = \bigoplus_j (U_j \otimes I_{m_j})$, where I_{m_j} is the identity on an m_j-dimensional multiplicity space $\mathbb{C}^{m_j}$. Therefore, ρ must have the form $\rho = \bigoplus_j p_j (I_j/d_j \otimes \rho_j)$, where ρ_j is an arbitrary state on the multiplicity space $\mathbb{C}^{m_j}$. Since $|\varphi\rangle$ is a purification of ρ, with a suitable choice of basis we have $|\varphi\rangle = |\rho^{\frac{1}{2}}\rangle\rangle = \bigoplus_j \sqrt{p_j/d_j}\; |I_j\rangle\rangle|\rho_j^{\frac{1}{2}}\rangle\rangle$, which after storage becomes $|\varphi_U\rangle = \bigoplus_j \sqrt{p_j/d_j}|U_j\rangle\rangle|\rho_j^{\frac{1}{2}}\rangle\rangle$. Hence, for every U the state $|\varphi_U\rangle$ belongs to the subspace $\widetilde{\mathcal{H}} = \bigoplus_j (\mathcal{H}_j^{\otimes 2} \otimes |\rho_j^{\frac{1}{2}}\rangle\rangle) \simeq \bigoplus_j \mathcal{H}_j^{\otimes 2}$. ■

We can then restrict our attention to the subspace $\widetilde{\mathcal{H}}$, and consider retrieving channels $\mathscr{R}$ from $(\mathcal{H}_{2N+2} \otimes \widetilde{\mathcal{H}})$ to $\mathcal{H}_{2N+3}$. The normalization of the Choi operator is then

$$\mathrm{Tr}_{2N+3}[R] = I_{2N+2} \otimes I_{\widetilde{\mathcal{H}}} \;. \tag{3.12}$$

Combining the expression of the fidelity (3.3) with that of the input state (3.11), it is easy to see that one can always use a covariant retrieving channel, satisfying

$$\left[R, U_{2N+3} \otimes V^*_{2N+2} \otimes \widetilde{U}^* \widetilde{V}'\right] = 0, \tag{3.13}$$

where $\widetilde{V}' = \bigoplus_j (I_j \otimes V_j)$ acts on $\widetilde{\mathcal{H}}$. We now exploit the decompositions $U \otimes U_j^* = \bigoplus_K \left(U_K \otimes I_{m_K^{(j)}}\right)$ and $V^* \otimes V_j = \bigoplus_L \left(V_L^* \otimes I_{m_L^{(j)}}\right)$, which yield

$$U_{2N+3} \otimes V^*_{2N+2} \otimes \widetilde{U}^* \widetilde{V} = \bigoplus_{K,L} (U_K \otimes V_L^* \otimes I_{m_{KL}}) \;. \tag{3.14}$$

Here $I_{m_{KL}}$ is given by $I_{m_{KL}} = \bigoplus_{j \in \mathsf{P}_{KL}} \left(I_{m_K^{(j)}} \otimes I_{m_L^{(j)}}\right)$, where P_{KL} is the set of values of j such that the irrep $U_K \otimes V_L^*$ is contained in the decomposition of $U \otimes V^* \otimes U_j^* \otimes V_j$. Relations (3.13) and (3.14) then imply

$$R = \bigoplus_{K,L} (I_K \otimes I_L \otimes R_{KL}) \ , \tag{3.15}$$

where R_{KL} is a positive operator on the multiplicity space

$$\mathbb{C}^{m_{JK}} = \bigoplus_{j \in \mathsf{P}_{KL}} \left(\mathbb{C}^{m_K^{(j)}} \otimes \mathbb{C}^{m_L^{(j)}} \right) . \tag{3.16}$$

Moreover, using the equality $I \otimes I_j = \bigoplus_K (I_K \otimes I_{m_K^{(j)}})$ we obtain

$$|I\rangle\!\rangle|\varphi^*\rangle = \bigoplus_{j=0}^{N/2} \sqrt{\frac{p_j}{d_j}} |I\rangle\!\rangle|I_j\rangle\!\rangle = \bigoplus_K |I_K\rangle\!\rangle|\alpha_K\rangle \ , \tag{3.17}$$

where $|I_K\rangle\!\rangle \in \mathcal{H}_K^{\otimes 2}$ and $|\alpha_J\rangle \in \mathbb{C}^{m_{KK}}$ is given by

$$|\alpha_K\rangle = \bigoplus_{j \in \mathsf{P}_{KK}} \sqrt{\frac{p_j}{d_j}} \ |I_{m_K^{(j)}}\rangle\!\rangle \ . \tag{3.18}$$

Exploiting Eqs. (3.15) and (3.17), the fidelity (3.10) can be rewritten as

$$F = \sum_K \frac{d_K}{d^2} \ \langle \alpha_K | R_{KK} | \alpha_K \rangle \ . \tag{3.19}$$

Theorem 12 (Optimal retrieving strategy) *The optimal retrieving of U from the memory state $|\varphi_U\rangle$ is achieved by measuring the ancilla with the optimal POVM $P_{\hat{U}} = |\eta_{\hat{U}}\rangle\langle\eta_{\hat{U}}|$ given by*

$$|\eta_{\hat{U}}\rangle = \bigoplus_j \sqrt{d_j} |\hat{U}_j\rangle\!\rangle, \tag{3.20}$$

and, conditionally to outcome $\hat{U}$, by performing the unitary $\hat{U}$ on the input system.

Proof. Let us denote by $P_{KL}^{(j)}$ the projector on the tensor product $\mathbb{C}^{m_K^{(j)}} \otimes \mathbb{C}^{m_L^{(j)}}$, and by $R_{KL}^{(j)} = P_{KL}^{(j)} R_{KL} P_{KL}^{(j)}$ the corresponding diagonal block of R_{KL}. Using Schur lemmas and Eq. (3.15) we obtain

$$\mathrm{Tr}_{2N+3}[R] = \sum_{K,L} \sum_{j \in \mathsf{P}_{KL}} (d_K I_j \otimes I_L \otimes \mathrm{Tr}_{m_K^{(j)}}[R_{KL}^{(j)}])/d_j. \tag{3.21}$$

The normalization of Eq. (3.12) then becomes

$$\sum_{K\in \mathsf{S}_{jL}} \frac{d_K}{d_j}\ \mathrm{Tr}_{m_K^{(j)}}[R^{(j)}_{KL}] = I_{m_L^{(j)}}\ , \tag{3.22}$$

having defined S_{jL} as the set of all K such that j belongs to P_{KL}. For the fidelity (3.19) we then have the bound

$$\begin{aligned} F &= \sum_K \frac{d_K}{d^2} \sum_{j,j'\in \mathsf{P}_{KK}} \sqrt{\frac{p_j p_{j'}}{d_j d_{j'}}} \langle\langle I_{m_K^{(j)}}|R_{KK}|I_{m_K^{(j')}}\rangle\rangle & (3.23)\\ &\leq \sum_K \frac{d_K}{d^2} \left(\sum_{j\in \mathsf{P}_{KK}} \sqrt{\frac{p_j \langle\langle I_{m_K^{(j)}}|R^{(j)}_{KK}|I_{m_K^{(j)}}\rangle\rangle}{d_j}} \right)^2 & (3.24)\\ &\leq \sum_K \frac{\left(\sum_{j\in \mathsf{P}_{KK}} m_K^{(j)}\sqrt{p_j}\right)^2}{d^2} = F_{est}\ , & (3.25) \end{aligned}$$

having used the positivity of R_{KK} for the first bound and the normalization of $R^{(j)}_{KK}$ in Eq. (3.22) for the second. Regarding the last equality, it can be proved as follows. First, the Choi operator of the measure-and-prepare strategy is

$$R_{est} = \int_G |\hat{U}\rangle\rangle\langle\langle\hat{U}| \otimes |\eta^*_{\hat{U}}\rangle\langle\eta^*_{\hat{U}}|\, \mathrm{d}\,\hat{U}. \tag{3.26}$$

Using Eq. (3.17) with $|\varphi^*\rangle$ replaced by $|\eta^*_I\rangle$ and performing the integral we obtain

$$R_{est} = \bigoplus_K (I_K^{\otimes 2} \otimes \widetilde{R}_{KK})/d_K, \tag{3.27}$$

where $\widetilde{R}_{KK} = |\beta_K\rangle\langle\beta_K|$ and

$$|\beta_K\rangle = \bigoplus_{j\in \mathsf{P}_{KK}} \sqrt{d_j}|I_{m_K^{(j)}}\rangle\rangle. \tag{3.28}$$

Eq. (3.19) then gives $F_{est} = \sum_K\ |\langle\alpha_K|\beta_K\rangle|^2/d^2$. ■

Using the above result it becomes easy to optimize the input state for storing. In fact, such a state is just the optimal state for the estimation of the unknown unitary U [13], whose expression is known in most relevant cases. For example, when U is an unknown qubit unitary in $SU(2)$, learning becomes equivalent to optimal estimation of an unknown rotation in the

Bloch sphere [14]. For large number of copies, the optimal input state is given by

$$|\varphi\rangle \approx \sqrt{4/N} \sum_{j=j_{\min}}^{N/2} \frac{\sin(2\pi j/N)}{\sqrt{2j+1}} \, |I_j\rangle\!\rangle, \tag{3.29}$$

with $j_{\min} = 0$ for N even and $j_{\min} = 1/2$ for N odd, and the fidelity is

$$F \approx 1 - \frac{2\pi^2}{N^2}. \tag{3.30}$$

Remarkably, this asymptotic scaling can be achieved without using entanglement between the set of N qubits that are rotated and an auxiliary set of N rotationally invariant qubits: the optimal storing is achieved just by applying $U^{\otimes N}$ on a the optimal N-qubit state [14]. Another example is that of an unknown phase-shift $U = \exp[i\theta\sigma_z]$. In this case, for large number of copies the optimal input state is

$$|\varphi\rangle = \sqrt{2/(N+1)} \sum_{m=-N/2}^{N/2} \sin[\pi(m+1/2)/(N+1)]|m\rangle \tag{3.31}$$

and the fidelity is [15]

$$F \approx 1 - \frac{2\pi^2}{(N+1)^2}. \tag{3.32}$$

Again, the optimal state can be prepared using only N qubits.

Our result can be immediately extended to the case where Bob has to reproduce $M > 1$ copies of the unknown unitary U. Indeed, let $\mathscr{C}_U$ be the M-partite channel obtained by Bob, and $\mathscr{C}_U^{(1)}$ be the local channel

$$\mathscr{C}_U^{(1)}(\rho) = \mathrm{Tr}_{\bar{1}}[\mathscr{C}_U(\rho \otimes (I/d)^{\otimes M-1})], \tag{3.33}$$

where $\mathrm{Tr}_{\bar{1}}$ denotes the trace over all spaces except the first. The local channel $\mathscr{C}_U^{(1)}$ describes the evolution of the first input of $\mathscr{C}_U$ when a randomly chosen state is sent to the remaining $(M-1)$ inputs. Of course, the fidelity between $\mathscr{C}_U^{(1)}$ and the unitary U cannot be larger than the optimal fidelity F_{est} of Eq. (3.25), and the same holds for any local channel $\mathscr{C}_U^{(i)}$, in which all but the i-th input system are discarded. Therefore, the measure-and-prepare strategy is optimal also for the maximization of the single-copy fidelity of all local channels, and such fidelity does not decrease with increasing M. Moreover, our result can be extended to the maximization of the global fidelity between $\mathscr{C}_U$ and $U^{\otimes M}$, just by replacing U with $U^{\otimes M}$ in all derivations. Again, the optimal retrieving is obtained by measuring the optimal POVM $P_{\hat{U}}$ and by performing $\hat{U}^{\otimes M}$ conditionally to outcome $\hat{U}$. Finally, we note that the same result holds as well when the input (output) uses are not identical copies $U^{\otimes N}$ ($U^{\otimes M}$), but generally N (M) different unitaries, each of them belonging to a different representation of the group G.

We conclude by extending our result to the *optimal inversion* of an unknown unitary U. For this task the fidelity of the learning board is

$$F' = \frac{1}{d^2} \int_G \langle\langle U^\dagger | \langle\langle U^* |^{\otimes N} \; L' \; | U^\dagger \rangle\rangle | U^* \rangle\rangle^{\otimes N} \, \mathrm{d}\, U, \tag{3.34}$$

as obtained by substituting U with $U^\dagger$ in the target of Eq. (3.3). From this expression it is easy to see that one can always assume

$$[L', V_{2N+3} \otimes U^*_{2N+2} \otimes U_o^{*\otimes N} \otimes V_i^{\otimes N}] = 0. \tag{3.35}$$

Therefore, the optimal inversion is obtained from our derivations by simply substituting $U_{2N+3} \to V^*_{2N+3}$ and $V^*_{2N+2} \to U_{2N+3}$. Accordingly, the optimal inversion is achieved by measuring the optimal POVM $P_{\hat{U}}$ on the optimal state $|\varphi_U\rangle$ and by performing $\hat{U}^\dagger$ conditionally to outcome $\hat{U}$. This provides the optimal approximate re-alignment of reference frames in the quantum communication scenario recently considered in Ref. [12]. In this scenario, the state $|\varphi\rangle \in \widetilde{\mathcal{H}}$ serves as a token of Alice's reference frame, and is sent to Bob along with a quantum message $|\psi\rangle \in \mathcal{H}$. Due to the mismatch of reference frames, Bob receives the decohered state

$$\sigma_\psi = \int_G |\varphi_U\rangle\langle\varphi_U| \otimes U|\psi\rangle\langle\psi|U^\dagger \, \mathrm{d}\, U, \tag{3.36}$$

from which he tries to retrieve the message $|\psi\rangle$ with maximum fidelity

$$f = \int \mathrm{d}\, \psi \; \langle\psi|\mathscr{R}'(\sigma_\psi)|\psi\rangle \, \mathrm{d}\, \psi, \tag{3.37}$$

where $\mathscr{R}'$ is the retrieving channel and $\mathrm{d}\, \psi$ denotes the uniform probability measure over pure states. The maximization of f is equivalent to the maximization of the channel fidelity

$$F' = \int_G \langle\langle U^\dagger | \langle \varphi^*_U | R' | U^\dagger \rangle\rangle | \varphi^*_U \rangle \, \mathrm{d}\, U, \tag{3.38}$$

which is the figure of merit for optimal inversion. It is worth stressing that the state $|\varphi_{fid}\rangle$ that maximizes the fidelity is not the state

$$|\varphi_{lik}\rangle = \bigoplus_j \sqrt{\frac{d_j}{\sum_j d_j^2}} |I_j\rangle\rangle, \tag{3.39}$$

that maximizes the likelihood [16, 17]. For $G = SU(2), U(1)$ the state $|\varphi_{fid}\rangle$ gives an average fidelity that approaches 1 as $1/N^2$, while for $|\varphi_{lik}\rangle$ the scaling is $1/N$. On the other hand, Ref. [12] shows that $|\varphi_{lik}\rangle$ allows a prefect correction of the misalignment errors with probability of success $p = 1 - 1/N^{350000}$, a thing which is not possible for $|\varphi_{fid}\rangle$. The determination of the best input state to maximize the probability of success, and the study of the trade-off between probability and fidelity remain open interesting problems for future work.

In conclusions, in this we found the optimal storing and retrieving of an unknown unitary transformation with N input and M output copies, proving the optimality of incoherent "measure-and-rotate" strategies under general hypotheses. The result has been extended to the optimal inversion of U, with application to the optimal approximate re-alignment of reference frames for quantum communication.

3.2 Optimal quantum tomography

A crucial issue in quantum information theory is the precise determination of states and processes. The procedure by which this task can be accomplished is known as *quantum tomography* [26, 19, 20].

Tomographing an unknown state ρ of a quantum system means performing a suitable POVM $\{P_i\}$ such that every expectation value can be evaluated from the probability distribution $p_i = \mathrm{Tr}[\rho P_i]$. In particular the expectation value of an operator A can be obtained when it is possible to expand A over the POVM as follows

$$A = \sum_i f_i[A] P_i, \tag{3.40}$$

$f_i[A]$ denoting suitable expansion coefficients. The expectation of A is then obtained as

$$\langle A \rangle = \sum_i f_i[A] \langle P_i \rangle. \tag{3.41}$$

When expansion (3.40) holds for all operators $\mathcal{B}(\mathcal{H})$—i. e. $\mathcal{B}(\mathcal{H}) = \mathsf{span}\{P_i\}$—the POVM is called *informationally complete* [22, 23].

Information-completeness of the POVM along with convergence of the series (3.40) rewrite as follows

$$a\|A\|_2^2 \leqslant \sum_{i=1}^{N} |\langle\langle P_i|A\rangle\rangle|^2 \leqslant b\|A\|_2^2, \quad A \in B(\mathcal{H}), \tag{3.42}$$

with $0 < a \leqslant b < \infty$. Sets of vectors $|P_i\rangle\rangle$ satisfying condition (3.42) are known as *frames* [24, 25]. This condition is equivalent to invertibility of the *frame operator* $F = \sum_i |P_i\rangle\!\rangle\langle\!\langle P_i|$. The expansion in Eq. (3.40) can be written as follows

$$|A\rangle\!\rangle = \sum_i \langle\!\langle D_i|A\rangle\!\rangle |P_i\rangle\!\rangle, \tag{3.43}$$

in terms of a *dual frame* $\{D_i\}$, namely a set of operators satisfying the identity $\sum_i |P_i\rangle\!\rangle\langle\!\langle D_i| = I$. For linearly dependent frame $\{P_i\}$ the dual $\{D_i\}$ is not unique.

The request for the POVM $\{P_i\}$ to be informationally complete can be relaxed if we have some prior information about the state ρ. If we know that the state belongs to a given subspace $\mathcal{V} \subseteq \mathcal{B}(\mathcal{H})$ the expectation value is

$$\langle A\rangle = \langle\!\langle \rho|A\rangle\!\rangle = \langle\!\langle \rho|Q_{\mathcal{V}}|A\rangle\!\rangle \tag{3.44}$$

$Q_{\mathcal{V}}$ orthogonal projector on $\mathcal{V}$, whence the set $\{P_i\}$ is required to span only $\mathcal{V}$.

For the estimation of the expectation $\langle A\rangle$ of an observable A, optimality means minimization of the cost function given by the variance $\delta(A)$ of the random variable $\langle\!\langle D_i|A\rangle\!\rangle$ with probability distribution $\mathrm{Tr}[\rho P_i]$, namely

$$\delta(A) := \sum_i |\langle\!\langle D_i|A\rangle\!\rangle|^2 \,\mathrm{Tr}[\rho P_i] - |\,\mathrm{Tr}[\rho A]|^2. \tag{3.45}$$

In a Bayesian scheme the state ρ is randomly drawn from an ensemble $\mathcal{S} = \{\rho_k, p_k\}$ of states ρ_k with prior probability p_k, with the variance averaged over $\mathcal{S}$, leading to

$$\delta_{\mathcal{S}}(A) := \sum_i |\langle\!\langle D_i|A\rangle\!\rangle|^2 \,\mathrm{Tr}[\rho_{\mathcal{S}} P_i] - \sum_k p_k |\,\mathrm{Tr}[\rho_k A]|^2 \tag{3.46}$$

where $\rho_{\mathcal{S}} = \sum_k p_k \rho_k$. Moreover, a priori we can be interested in some observables more than other ones, and this can be specified in terms of a weighted set of observables $\mathcal{G} = \{A_n, q_n\}$, with weight $q_n > 0$ for the observable A_n. Averaging over $\mathcal{G}$ we have

$$\delta_{\mathcal{S},\mathcal{G}} := \sum_i \langle\!\langle D_i|G|D_i\rangle\!\rangle \,\mathrm{Tr}[\rho_{\mathcal{S}} P_i] - \sum_{k,n} p_k q_n |\,\mathrm{Tr}[\rho_k A_n]|^2 \tag{3.47}$$

where $G = \sum_n q_n |A_n\rangle\!\rangle\langle\!\langle A_n|$. The weighted set $\mathcal{G}$ yields a representation of the state, given in terms of the expectation values. The representation is faithful when $\{A_n\}$ is an operator frame, e. g. when it is made of the dyads $|i\rangle\langle j|$ corresponding to the matrix elements $\langle j|\rho|i\rangle$.

Notice that only the first term of $\delta_{\mathcal{S},\mathcal{G}}$ depends on $\{P_i\}$ and $\{D_i\}$. If $\rho_i \in \mathcal{V}$ for all states $\rho_i \in \mathcal{S}$, reminding Eq. (3.44) the first term of Eq. (3.47) becomes

$$\eta = \sum_i \langle\langle D_i | Q_\mathcal{V} G Q_\mathcal{V} | D_i \rangle\rangle \operatorname{Tr}[\rho_\mathcal{S} P_i]. \tag{3.48}$$

We now generalize this approach to tomography of quantum operations, keeping generally different input and output Hilbert spaces $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$, respectively. This has the advantage that the usual tomography of states comes as the special case of one-dimensional $\mathcal{H}_{in}$, whereas tomography of POVMs corresponds to one-dimensional $\mathcal{H}_{out}$.

A quantum operation is a trace non increasing CP-map $\mathcal{T} : \mathcal{B}(\mathcal{H}_{in}) \longrightarrow \mathcal{B}(\mathcal{H}_{out})$. In order to gather information about a quantum operation $\mathcal{T}$, the most general procedure consists in: *i*) preparing a state $\rho \in \mathcal{B}(\mathcal{H}_{in} \otimes \mathcal{H}_A)$ where $\mathcal{H}_A$ is an ancillary system with the same dimension of $\mathcal{H}_{in}$; *ii*) measuring the state $(\mathcal{T} \otimes \mathcal{I}_A)(\rho)$ with a POVM $\{P_i\}$. The probability of obtaining a generic outcome i is given by

$$p_i = \operatorname{Tr}[(\mathcal{T} \otimes \mathcal{I}_A)(\rho) P_i], \tag{3.49}$$

which, using the Choi-Jamiołkowski isomorphism [48],

$$\mathcal{T}(\rho) = \operatorname{Tr}_{in}[(I_{out} \otimes \rho^T) R_\mathcal{T}], \;\; R_\mathcal{T} = \mathcal{T} \otimes I_{in}(|I\rangle\rangle\langle\langle I|) \tag{3.50}$$

becomes

$$\operatorname{Tr}[\operatorname{Tr}_{in}[(I_A \otimes R_\mathcal{T})(\rho^{\theta_{in}} \otimes I_{out})] P_i] = \operatorname{Tr}[R_\mathcal{T} \Pi_i^{(\rho)}], \tag{3.51}$$

where θ is the transposition w.r.t. the orthonormal basis in Eq. (1.5), and

$$\Pi_i^{(\rho)} = \{\operatorname{Tr}_A[(\rho \otimes I_{out})(I_{in} \otimes P_i^{\theta_{out}})]\}^T. \tag{3.52}$$

It is convenient to use here the notion of *tester* along with the theoretical framework introduced in [11, 28]. A tester is the natural generalization of the concept of POVM from states to transformations, and is represented by a set of positive operators $\{\Pi_i\}$ with

$$\sum_i \Pi_i = I \otimes \sigma, \qquad \operatorname{Tr}[\sigma] = 1 \tag{3.53}$$

The probability distribution in Eq. (3.51) is precisely represented by a Born-rule with the tester $\{\Pi_i\}$ in place of $\{P_i\}$, and the operator $R_\mathcal{T}$ in place of ρ. Such generalized Born rule can be rewritten in terms of the usual one as follows [11, 28, 29]

$$p_i = \operatorname{Tr}[R_\mathcal{T} \Pi_i] = \operatorname{Tr}[\mathcal{T} \otimes \mathcal{I}(\nu) P_i], \tag{3.54}$$

with

$$\nu = |\sqrt{\sigma}\rangle\!\rangle\langle\!\langle\sqrt{\sigma}|, \qquad P_i = (I \otimes \sigma^{-1/2})\Pi_i(I \otimes \sigma^{-1/2}). \tag{3.55}$$

This method allows a straightforward generalization of the tomographic method from states to transformation. Now tomographing a quantum operation means using a suitable tester Π_i such that the expectation value of any other possible measurement can be inferred by the probability distribution $p_i = \mathrm{Tr}[R_{\mathcal{T}}\Pi_i]$. In order to achieve this task we have to require that $\{\Pi_i\}$ is an operator frame for $\mathcal{B}(\mathcal{H}_{out} \otimes \mathcal{H}_{in})$. This means that we can expand any operator on $\mathcal{H}_{out} \otimes \mathcal{H}_{in}$ as follows

$$A = \sum_i \langle\!\langle\Delta_i|A\rangle\!\rangle \Pi_i \qquad A \in \mathcal{B}(\mathcal{H}_{out} \otimes \mathcal{H}_{in}). \tag{3.56}$$

where $\{\Delta_i\}$ is a possible dual of the frame $\{\Pi_i\}$, that is the condition

$$\sum_i |\Pi_i\rangle\!\rangle\langle\!\langle\Delta_i| = I_{out} \otimes I_{in} \tag{3.57}$$

holds.

Optimizing the tomography of quantum operations means minimizing the statistical error in the determination of the expectation of a generic operator A as in Eq. (3.56). This is provided by the variance

$$\delta(A) = \sum_i |\langle\!\langle\Delta_i|A\rangle\!\rangle|^2 \,\mathrm{Tr}[R_{\mathcal{T}}\Pi_i] - |\,\mathrm{Tr}[R_{\mathcal{T}}A]|^2 \tag{3.58}$$

We assume an ensemble $\mathcal{E} = \{R_k, p_k\}$ of possible transformations and a weighted set $\mathcal{G} = \{A_n, q_n\}$ of possible observables. Averaging the statistical error over these ensembles we obtain

$$\delta_{\mathcal{E},\mathcal{A}} := \sum_i \langle\!\langle\Delta_i|G|\Delta_i\rangle\!\rangle \,\mathrm{Tr}[R_{\mathcal{E}}\Pi_i] - \sum_{k,n} p_k q_n |\,\mathrm{Tr}[R_k A_n]|^2. \tag{3.59}$$

Optimizing this figure of merit means: *i*) optimizing the choice of the dual frame $\{\Delta_i\}$; *ii*) optimizing the choice of the frame $\{\Pi_i\}$. The optimization of the set $\{\Pi_i\}$ reflects in both choosing the best input state for the quantum operation and the best final measurement.

In the following, for the sake of clarity we will consider

$$\dim(\mathcal{H}_{in}) = \dim(\mathcal{H}_{out}) =: d, \tag{3.60}$$

and focus on the "symmetric" case $G = I$; this happens for example when the set $\{A_n\}$ is an orthonormal basis, whose elements are equally weighted.

Moreover, we assume that the averaged channel of the ensemble $\mathcal{E}$ is the maximally depolarizing channel, whose Choi operator is $R_{\mathcal{E}} = d^{-1} I \otimes I$.

With these assumptions the relevant term of figure of merit becomes

$$\eta = \sum_i \langle\langle \Delta_i | \Delta_i \rangle\rangle d^{-1} \operatorname{Tr}[\Pi_i]. \tag{3.61}$$

Since $R_{\mathcal{E}}$ is invariant under the action of $SU(d) \times SU(d)$ we now show that it is possible to impose the same covariance also on the tester without increasing the value of η. Let us define

$$\Pi_{i,g,h} := (U_g \otimes V_h) \Pi_i (U_g^\dagger \otimes V_h^\dagger), \tag{3.62}$$

$$\Delta_{i,g,h} := (U_g \otimes V_h) \Delta_i (U_g^\dagger \otimes V_h^\dagger). \tag{3.63}$$

It is easy to check that $\Delta_{i,g,h}$ is a dual of $\Pi_{i,g,h}$ by evaluating the group average after the sum on i. Then we observe that the normalization of $\Pi_{i,g,h}$ gives

$$\sum_i \int dg dh \, \Pi_{i,g,h} = d^{-1} I \otimes I \tag{3.64}$$

corresponding to $\sigma = d^{-1} I$ in Eq. (3.55), namely one can choose $\nu = d^{-1} |I\rangle\rangle\langle\langle I|$. In the last identity dg and dh are invariant measures normalized to unit.

It is easy to verify that the figure of merit for the covariant tester is the same as for the non covariant one, whence, w.l.o.g. we optimize the covariant tester. The condition that the covariant tester is informationally complete w.r.t. the subspace of transformations to be tomographed will be verified after the optimization.

We note that a generic covariant tester is obtained by Eq. (3.62), with operators Π_i becoming "seeds" of the covariant POVM, and now being required to satisfy only the normalization condition

$$\sum_i \operatorname{Tr}[\Pi_i] = d \tag{3.65}$$

(analogous of covariant POVM normalization in [30, 21]). The problem of optimization of the dual frame has been solved in [31]. With the optimal dual, the figure of merit simplifies as

$$\eta = \operatorname{Tr}[\tilde{X}^{-1}], \tag{3.66}$$

where

$$\tilde{X} = \sum_i \int dg dh \, \frac{d |\Pi_{i,g,h}\rangle\rangle\langle\langle \Pi_{i,g,h}|}{\operatorname{Tr}[\Pi_{i,g,h}]} = \int dg dh \, W_{g,h} X W_{g,h}^\dagger \tag{3.67}$$

with $W_{g,h} = U_g \otimes U_g^* \otimes V_h \otimes V_h^*$ and

$$X = \sum_i \frac{d|\Pi_i\rangle\!\rangle\langle\!\langle\Pi_i|}{\mathrm{Tr}[\Pi_i]}. \tag{3.68}$$

We label the four spaces in such a way that $W_{g,h} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2 \otimes \mathcal{H}_4)$, with $\mathcal{H}_1 = \mathcal{H}_{out}$ and $\mathcal{H}_2 = \mathcal{H}_{in}$. Using Schur's lemma we have

$$\begin{aligned}
\tilde{X} &= P_1 + AP_2 + BP_3 + CP_4,\\
P_1 &= \Omega_{13} \otimes \Omega_{24},\\
P_2 &= (I_{13} - \Omega_{13}) \otimes \Omega_{24},\\
P_3 &= \Omega_{13} \otimes (I_{24} - \Omega_{24}),\\
P_4 &= (I_{13} - \Omega_{13}) \otimes (I_{24} - \Omega_{24}),
\end{aligned} \tag{3.69}$$

having posed $\Omega = |I\rangle\!\rangle\langle\!\langle I|/d$ and

$$\begin{aligned}
A &= \frac{1}{d^2-1}\left\{\sum_i \frac{\mathrm{Tr}[(\mathrm{Tr}_2[\Pi_i])^2]}{\mathrm{Tr}[\Pi_i]} - 1\right\}\\
B &= \frac{1}{d^2-1}\left\{\sum_i \frac{\mathrm{Tr}[(\mathrm{Tr}_1[\Pi_i])^2]}{\mathrm{Tr}[\Pi_i]} - 1\right\}\\
C &= \frac{1}{(d^2-1)^2}\left\{\sum_i \frac{d\,\mathrm{Tr}[\Pi_i^2]}{\mathrm{Tr}[\Pi_i]} - (d^2-1)(A+B) - 1\right\}.
\end{aligned} \tag{3.70}$$

One has

$$\mathrm{Tr}[\tilde{X}^{-1}] = 1 + (d^2-1)\left(\frac{1}{A} + \frac{1}{B} + \frac{(d^2-1)}{C}\right). \tag{3.71}$$

We note that if the ensemble of transformations is contained in a subspace $\mathcal{V} \subseteq \mathcal{B}(\mathcal{H}_{out} \otimes \mathcal{H}_{in})$ the figure of merit becomes $\eta = \mathrm{Tr}[\tilde{X}^{\ddagger} Q_{\mathcal{V}}]$, where $\tilde{X}^{\ddagger}$ is the Moore-Penrose pseudoinverse. We now carry on the minimization for three relevant subspaces:

$$\begin{aligned}
\mathcal{Q} &= \mathcal{B}(\mathcal{H}_{out} \otimes \mathcal{H}_{in}),\\
\mathcal{C} &= \{R \in \mathcal{Q},\ \mathrm{Tr}_{out}[R] = I_{in}\}\\
\mathcal{U} &= \{R \in \mathcal{Q},\ \mathrm{Tr}_{out}[R] = I_{in}, \mathrm{Tr}_{in}[R] = I_{out}\}
\end{aligned} \tag{3.72}$$

corresponding respectively to quantum operations, general channels and unital channels. The subspaces $\mathcal{C}$ and $\mathcal{U}$ are invariant under the action of the group $\{W_{g,h}\}$ and thus the respective projectors decompose as

$$\begin{aligned}
Q_{\mathcal{C}} &= P_1 + P_2 + P_4,\\
Q_{\mathcal{U}} &= P_1 + P_4
\end{aligned} \tag{3.73}$$

Without loss of generality we can assume the operators $\{\Pi_i\}$ to be rank one. In fact, suppose that Π_i has rank higher than 1. Then it is possible to decompose it as $\Pi = \sum_j \Pi_{i,j}$ with $\Pi_{i,j}$ rank 1. The statistics of Π_i can be completely achieved by $\Pi_{i,j}$ through a suitable post-processing. For the purpose of optimization it is then not restrictive to consider rank one Π_i, namely $\Pi_i = \alpha_i |\Psi_i\rangle\rangle\langle\langle\Psi_i|$, with $\sum_i \alpha_i = d$. Notice that all multiple seeds of this form lead to testers satisfying Eq. (3.65).

In the three cases under examination, the figure of merit is then

$$\begin{aligned} \eta_{\mathcal{Q}} &= \mathrm{Tr}[\tilde{X}^{-1}] = 1 + (d^2-1)\left(\frac{2}{A} + \frac{(d^2-1)^2}{1-2A}\right) \\ \eta_{\mathcal{C}} &= \mathrm{Tr}[\tilde{X}^{\ddagger} Q_{\mathcal{C}}] = 1 + (d^2-1)\left(\frac{1}{A} + \frac{(d^2-1)^2}{1-2A}\right) \\ \eta_{\mathcal{U}} &= \mathrm{Tr}[\tilde{X}^{\ddagger} Q_{\mathcal{U}}] = 1 + (d^2-1)\left(\frac{(d^2-1)^2}{1-2A}\right) \end{aligned} \tag{3.74}$$

where

$$0 \leqslant A = (d^2-1)^{-1}\big(\sum_i \alpha_i \,\mathrm{Tr}[(\Psi_i\Psi_i^{\dagger})^2] - 1\big) \leqslant \frac{1}{d+1} < \frac{1}{2}. \tag{3.75}$$

The minimum can simply be determined by derivation with respect to A, obtaining $A = 1/(d^2+1)$ for quantum operations, $A = 1/(\sqrt{2}(d^2-1)+2)$ for general channels and $A = 0$ for unital channels. The corresponding minimum for the figure of merit is

$$\begin{aligned} \eta_{\mathcal{Q}} &\geqslant d^6 + d^4 - d^2 \\ \eta_{\mathcal{C}} &\geqslant d^6 + (2\sqrt{2}-3)d^4 + (5-4\sqrt{2})d^2 + 2(\sqrt{2}-1) \\ \eta_{\mathcal{U}} &\geqslant (d^2-1)^3 + 1. \end{aligned} \tag{3.76}$$

The same result for quantum operations and for unital channels has been obtained in [32] in a different framework.

These bounds are simply achieved by a single seed $\Pi_0 = d|\Psi\rangle\rangle\langle\langle\Psi|$, with

$$\mathrm{Tr}[(\Psi\Psi^{\dagger})^2] = \frac{2d}{d^2+1}, \quad \frac{\sqrt{2}(d^2-1)+3}{d(\sqrt{2}(d^2-1)+2)}, \quad 1 \tag{3.77}$$

respectively for quantum operations, general channels and unital channels, namely with

$$\Psi = [d^{-1}(1-\beta)I + \beta\,|\psi\rangle\,\langle\psi|]^{\frac{1}{2}} \tag{3.78}$$

where $\beta = \sqrt{(d+1)/(d^2+1)}$ for quantum operations, $\beta = [(d-1)(2+\sqrt{2}(d^2-1))]^{-1/2}$ for general channels and $\beta = 0$ for unital channels, and

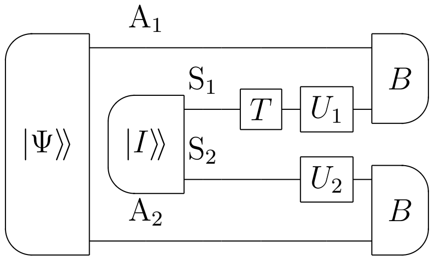
$|\psi\rangle$ is any pure state. The informationally completeness is thus verified *a posteriori* (see [30]).

The same procedure can be carried on when the operator G has the more general form $G = g_1 P_1 + g_2 P_2 + g_3 P_3 + g_4 P_4$, where P_i are the projectors defined in (3.69). In this case Eq. (3.71) becomes

$$\mathrm{Tr}[\tilde{X}^{-1} G] = g_1 + (d^2 - 1)\left(\frac{g_2}{A} + \frac{g_3}{B} + \frac{(d^2-1)g_4}{C}\right), \tag{3.79}$$

which can be minimized along the same lines previously followed. G has this form when optimizing measuring procedures of this kind: i) preparing an input state randomly drawn from the set $\{U_g \rho U_g^\dagger\}$; ii) measuring an observable chosen from the set $\{U_h A U_h^\dagger\}$.

We now show how the optimal measurement can be experimentally implemented with the circuit



(3.80)

The bipartite system carrying the Choi operator of the transformation is indicated with the labels S_1 and S_2. We prepare a pair of ancillary systems A_1 and A_2 in the joint state $|\Psi\rangle\!\rangle\langle\!\langle\Psi|$, then we apply two random unitary transformations U_1 and U_2 to S_1 and S_2, finally we perform a Bell measurement on the pair $\mathrm{A}_1\mathrm{S}_1$ and another Bell measurement on the pair $\mathrm{A}_2\mathrm{S}_2$. This experimental scheme realizes the continuous measurement by randomizing among a continuous set of discrete POVM; this is a particular application of a general result proved in [33]. The scheme proposed is feasible using e. g. the Bell measurements experimentally realized in [34]. We note that choosing $|\Psi\rangle\!\rangle$ maximally entangled (as proposed for example in [35]) is generally not optimal, except for the unital case.

With the same derivation starting from Eq. (3.61), but keeping $\dim(\mathcal{H}_{in}) \neq \dim(\mathcal{H}_{out})$, one obtains the optimal tomography for general quantum operations. The special case of $\dim(\mathcal{H}_{in}) = 1$ (one has $P_3 = P_4 = 0$ in Eq. (3.69)) corresponds to optimal tomography of states, whereas case $\dim(\mathcal{H}_{out}) = 1$ ($P_2 = P_4 = 0$) gives the optimal tomography of POVMs. The corresponding experimental schemes are obtained by removing the upper/lower branch for POVMs/states, respectively. In the remaining branch the bipartite detector becomes a mono-partite, performing a von Neumann measurement for the qudit, preceded by a random unitary in $SU(d)$. Moreover, for the case of POVM, the state $|\Psi\rangle\rangle$ is missing, whereas, for state-tomography, both bipartite states are missing. The optimal η in Eq. (3.48) is given by $\eta = d^3+d^2-d$, in both cases (for state-tomography compare with Ref. [36]).

The general method for optimizing quantum tomography presented here is very versatile, allowing to consider arbitrary prior ensemble and representation. We provided the optimal experimental schemes for tomography of states and various kinds of process tomography, giving the corresponding performance, all schemes being feasible with the current technology.

3.3 Localizability and entanglement-breaking

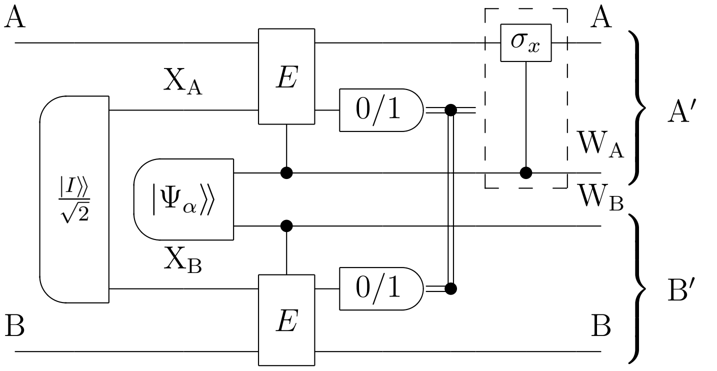
In Ref. [42] is was shown that not every no-signaling channel is *localizable* (see also Definition 8). The problem is how to generate "superquantum" correlations—i.e stronger than those arising from entanglement—without signaling, as for PR-boxes [56]. In Ref. [37], is was proposed the following

Conjecture 2 *All no-signaling channels are mixtures of entanglement-breaking and localizable channels.*

The conjecture was based on the only known quantum realization of a PR-box, which was made with an entanglement-breaking channel. Such conjecture, however, implicitly forbids truly coherent super-quantum correlations. This corresponds to perfect monogamy of correlations, in the sense that when the channel violates the Cirel'son bound [45] the entanglement of the input systems with other ones is broken. We will show that Conjecture 2 is false, allowing for more flexibility, with a trade-off between generated correlations and preserved entanglement, and with a violation of the Cirel'son bound achieved coherently, in the full range between the quantum bound and the maximum possible correlation.

We provide a counterexample to Conjecture 2, in terms of a no-signaling channel that is atomic, (i. e. it cannot be written as a convex combination of different channels whence also of no-signaling channels) and that is neither entanglement-breaking nor localizable [55].

Let $\mathrm{A}, \mathrm{B}, \mathrm{X}_\mathrm{A}, \mathrm{X}_\mathrm{B}, \mathrm{W}_\mathrm{A}, \mathrm{W}_\mathrm{B}$ be qubits. We define the channel $\mathscr{R}_\alpha$ depending on $\alpha, 0 \leqslant \alpha \leqslant 1$:

$\mathscr{R}_\alpha =$  (3.81)

where E is the swap operator,

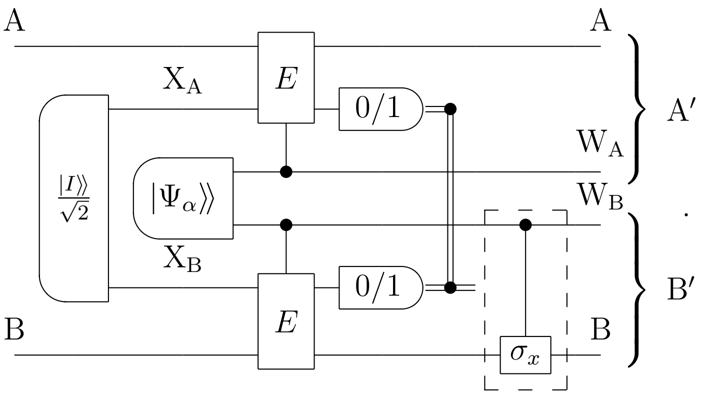
$$|\Psi_\alpha\rangle\rangle := \sqrt{\alpha}\,|0\rangle\,|0\rangle + \sqrt{1-\alpha}\,|1\rangle\,|1\rangle\,, \tag{3.82}$$

the two-qubit gate in the dashed box is a controlled-σ_x given by

$$\Sigma_{\mathrm{AW_A}} := |1\rangle\,\langle 1|_{\mathrm{W_A}} \otimes (\sigma_x)_{\mathrm{A}} + |0\rangle\,\langle 0|_{\mathrm{W_A}} \otimes I_{\mathrm{A}} \tag{3.83}$$

classically controlled by the outcomes of the measures on the computational basis (represented by the circuital element $-\boxed{0/1}$). Notice that the classical control works as a logical AND, implying that the box $\Sigma_{\mathrm{AW_A}}$ is performed if and only if both outcomes of the measurements $-\boxed{0/1}$ are equal to 1.

We notice that circuit $\mathscr{R}_\alpha$ in Eq. (3.81) is implemented using local operations, entanglement, and one round of classical communication from Bob to Alice, thus being of the form of Eq. (2.61). One can verify that $\mathscr{R}_\alpha$ can be equivalently realized applying the controlled-σ_x on systems B and $\mathrm{W_B}$ as follows

$\mathscr{R}_\alpha =$  . (3.84)

Consequently $\mathscr{R}_\alpha$ also admits a realization of the form given in Eq. (2.60). By Theorem 8, we can conclude that this is a no-signaling channel. The

Choi-Jamiołkowski operator of $\mathscr{R}_\alpha$ is:

$$R_\alpha = \sum_{m,n=0}^{1} |K_{mn}^\alpha\rangle\!\rangle\langle\!\langle K_{mn}^\alpha| \tag{3.85}$$

with

$$|K_{mn}^\alpha\rangle\!\rangle = \left[(\Sigma_{\mathrm{AW_A}})^{mn} \otimes \langle m|_{\mathrm{X_A}} \langle n|_{\mathrm{X_B}}\right] |\Phi_\alpha\rangle\!\rangle \tag{3.86}$$

and

$$|\Phi_\alpha\rangle\!\rangle = (\mathbb{E} \otimes I_{\mathrm{AB}}) \left(|I\rangle\!\rangle_{\mathrm{AB,AB}} \otimes \frac{1}{\sqrt{2}} |I\rangle\!\rangle_{\mathrm{X_A X_B}} \otimes |\Psi_\alpha\rangle\!\rangle \right), \tag{3.87}$$

where $\mathbb{E}$ denotes the tensor product of the two controlled-swap.

Using Mathematica, we prove that $\mathscr{R}_{\tilde{\alpha}}$ with $\tilde{\alpha} := 1/6$ is a counterexample by showing that it satisfies the following properties:

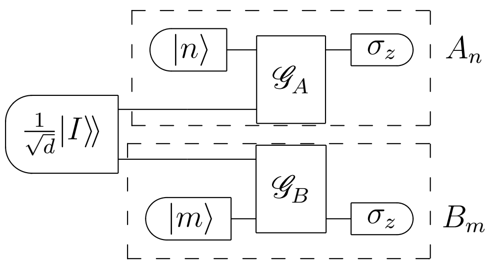
1. It is not entanglement-breaking,
2. It is not localizable
3. It is atomic.

Proof of (1)

$\mathscr{R}_{\tilde{\alpha}}$ is not entanglement breaking. A channel is entanglement breaking if and only if the corresponding Choi-Jamiołkowski operator is separable. Thus, we can prove that $\mathscr{R}_{\tilde{\alpha}}$ is not entanglement breaking by showing that $R_{\tilde{\alpha}}$ violates the Peres-Horodecki criterion for separability [46, 47]. According to the criterion, if a state is separable it has a positive definite partial transpose. Numerically one can check that $R_{\tilde{\alpha}}$ has a partial transpose with negative eigenvalues, whence we conclude that it is entangled and $\mathscr{R}_{\tilde{\alpha}}$ is not entanglement-breaking.

Proof of (2)

$\mathscr{R}_{\tilde{\alpha}}$ is not localizable. If $\mathscr{R}_\alpha$ were localizable (see Eq. (2.57)), the following observables A_n, B_m



(3.88)

($-\boxed{\sigma_z}$ represents the measurement of σ_z) would verify the Cirel'son bound [42]:

$$c_\alpha := |\langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle| \leqslant 2\sqrt{2}. \tag{3.89}$$

We have that

$$\langle A_n B_m\rangle = \mathrm{Tr}[(\sigma^z_A \otimes |n\rangle \langle n|_A \otimes \sigma^z_B \otimes |m\rangle \langle m|_B \otimes I_{W_A W_B}) R_\alpha] \tag{3.90}$$

whence (using expression in Eq. (3.85) for R_α) one finds $c_\alpha = |4-6\alpha|$. Since $c_{\tilde{\alpha}} = 3 > 2\sqrt{2}$, the Cirel'son bound is violated and $\mathscr{R}_{\tilde{\alpha}}$ cannot be localizable.

Proof of (3)

$\mathscr{R}_{\tilde{\alpha}}$ is extremal. One can check that the matrices $\{K^{\tilde{\alpha}\dagger}_{mn} K^{\tilde{\alpha}}_{m'n'}\}$ are linearly independent. By theorem 3 the channel $\mathscr{R}_{\tilde{\alpha}}$ is extremal.

In conclusion, we have provided the general realization scheme of no-signaling channels, giving a counterexample to the conjecture of Ref. [37], stating that such channels are mixtures of entanglement-breaking and localizable channels. The general scheme allows for more flexibility of entanglement monogamy, opening the new problem of determining the trade-off between generated correlations and preserved entanglement. The Cirel'son bound is violated coherently, in the full range up to the maximum possible correlation. The general realization scheme looks counter-intuitive, due to the presence of classical communication. However, the nontrivial constraint is the fact that an equivalent scheme must exist, with communication in the reverse direction, and remarkably this suffices to make the channel no-signaling.

The result has an intrinsic foundational relevance, involving the pivotal role of causality in theoretical physics and computer science.

Chapter 4

Quantum lambda calculus

Lambda calculus is a model of computation proposed by Church, based on the concept of "function". Traditionally, there are two different ways to define functions. In the first we consider functions as operations which transform an input into an output according to some rule. For example, let us consider the identity function on some set S. Intensionally, it is the operation which takes an element $x \in \mathsf{S}$ and gives the same element as output, i.e. we can think of it as the operation

$$x \mapsto x. \tag{4.1}$$

This is called the "intensional" interpretation. On the other hand, the "extensional" interpretation defines functions as relations between its input and output set. The extensional definition of the identity function is a subset F of $\mathsf{S} \times \mathsf{S}$, in particular

$$\mathsf{F} := \{(x, x)\ ,\ x \in \mathsf{S}\}. \tag{4.2}$$

The operative "meaning" of the function is lost.

A formalization of the intensional viewpoint led Alonzo Church and his collaborators to the Untyped lambda calculus. This model of computable functions turned out to be equivalent to a different model of computation based on the formalization of the operations carried out by a human when he actually compute a function. This second model is called "Turing Machine", after the name of the mathematician who proposed it.

The essential idea behind lambda calculus is that every computable function is recursively built from a set of variables and elementary operation. Each well-formed expression of this kind is called a "term".

In the next sections we sketch the main features of the lambda calculus, both in the untyped and in the simple typed version, essentially basing our presentation on [50] (see also [51, 52]).

4.1 Untyped Lambda calculus

The simplest version of lambda calculus is the untyped lambda calculus. Because of this simplicity, which is actually a lack of constraints on the formalism, the untyped lambda calculus is the most powerful version, being fully equivalent to the Turing Machine. As we will see, this power also yields to paradoxical behaviors. For this reason, it is ill-suited to be used as a formal description of quantum computation, or of any physical process in general: physical processes always behave consistently! It will be necessary to introduce types which mimic the fact the physical transformations acts between particular types of *systems*. For example, a two dimensional unitary transformation can only act on 1/2 spins.

It is nevertheless useful to begin with the untyped calculus, in order to settle the notation and introduce the fundamental concepts.

4.1.1 Syntax

In the "pure" untyped lambda calculus, the set of lambda terms is recursively defined by the following

Definition 10 *Assume given an infinite set $\mathcal{V}$ of variables. Let A be an alphabet consisting of the elements of $\mathcal{V}$, the special symbols "(", ")", "λ", and ".". The set of lambda terms is the smallest subset $\Lambda \subseteq A^*$ such that:*

- *If $x \in \mathcal{V}$ then $x \in \Lambda$*
- *If $M, N \in \Lambda$ then $(MN) \in \Lambda$*
- *If $x \in \mathcal{V}$ and $M \in \Lambda$ then $\lambda x.M \in \Lambda$*

This definition can be abbreviated with the following Backus-Naur Form (BNF):

$$Term \quad M, N ::= x \mid (MN) \mid (\lambda x.M) \tag{4.3}$$

The intended meaning of these expressions is the following: (MN) is the application of the function M to the argument N (in the usual notation for functions this would be denoted as $M(N)$); $\lambda x.M$ is the function which take x as input and produces as output the term M. For example, the identity function is represented by the term

$$\lambda x.x. \tag{4.4}$$

4.1.2 α-equivalence

The terms $\lambda x.x$ and $\lambda y.y$ are essentially the same, as is common in mathematical notation: $\int f(x)dx$ is the same as $\int f(y)dy$. Lambda calculus is a formal calculus, so this equivalence (called α-equivalence) needs to be clearly stated.

The set of free variables of a term M is denoted $FV(M)$ and it is defined as follows:

$$\begin{aligned} FV(x) &= \{x\} \\ FV(MN) &= FV(M) \cup FV(N) \\ FV(\lambda x.M) &= FM(M) \setminus \{x\} \end{aligned} \tag{4.5}$$

Variables can be *renamed*. Formally, renaming is defined as

$$\begin{aligned} x\{y/x\} &\equiv y, \\ z\{y/x\} &\equiv z, \quad \text{if } x \neq z, \\ (MN)\{y/x\} &\equiv (M\{y/x\})(N\{y/x\}), \\ (\lambda x.M)\{y/x\} &\equiv \lambda y.(M\{y/x\}), \\ (\lambda z.M)\{y/x\} &\equiv \lambda z.(M\{y/x\}), \quad \text{if } x \neq z. \end{aligned} \tag{4.6}$$

We define α-equivalence to be the smallest congruence relation on lambda terms such that for all terms M and all variables y not appearing in M,

$$\lambda x.M =_\alpha \lambda y.(M\{y/x\}). \tag{4.7}$$

Definition 11 *The* substitution *of N for all free occurrences of a variable x in a lambda term is recursively defined by the rules*

$$\begin{aligned} &x[N/x] = N, \\ &y[N/x] = y, \quad x \neq y, \\ &(MP)[N/x] = (M[N/x])(P(N/x]), \\ &(\lambda y.M)[N/x] = \lambda z.(M[z/y][N/x]), \quad z \neq FV(M) \cup FV(N) \end{aligned} \tag{4.8}$$

4.1.3 Operational semantics

The operational semantic of the lambda calculus is a set or rules which allow a progressive evaluation of a term. The evaluation is the formal equivalent of the *computation* of the term: a term which can be reducted (or evaluated) is like a program whose execution is not yet terminated. Formally we give the following

Definition 12 *A* redex *is any term of the form* $(\lambda x.M)N$. *Any term containing a redex as a subterm is called* reducible. *A lambda term without any redexes is said to be in* normal form.

The aim of reduction is exactly to reach a *normal* term.

The single-step reduction is given by the following rules

$$
\begin{array}{ll}
(\beta) & \dfrac{}{(\lambda x.M)N \to_\beta M[N/x]} \\[2ex]
(op) & \dfrac{M \to_\beta M'}{MN \to_\beta M'N} \\[2ex]
(arg) & \dfrac{N \to_\beta N'}{MN \to_\beta MN'} \\[2ex]
(\xi) & \dfrac{M \to_\beta M'}{\lambda x.M \to_\beta \lambda x.M'}
\end{array}
\tag{4.9}
$$

These rules do not allow an unambiguous choice of the reduction step during the evaluation of a term. If we need an explicitly determined reduction path for a term, we must choose an *evaluation strategy.* The most common choices are the *Call-by-value* and the *Call-by-name* strategies.

In a Call-by-value (CBV) reduction the argument of a function must be evaluated before evaluating the body of the function. A *value* v is any variable x or any abstraction $\lambda x.M$.

$$
Value \quad v ::= x \mid \lambda x.M \tag{4.10}
$$

The CBV single-step reduction rules are the following

$$
\begin{array}{ll}
(\beta_v) & \dfrac{}{(\lambda x.M)v \to_v M[v/x]} \\[2ex]
(op) & \dfrac{M \to_v M'}{MN \to_v M'N} \\[2ex]
(arg_v) & \dfrac{N \to_v N'}{vN \to_v vN'}
\end{array}
\tag{4.11}
$$

The Call-by-name (CBN) evaluation strategy is defined by the following single-step reduction rules

$$
\begin{array}{ll}
(\beta) & \dfrac{}{(\lambda x.M)N \to_n M[N/x]} \\[2ex]
(op) & \dfrac{M \to_n M'}{MN \to_n M'N}
\end{array}
\tag{4.12}
$$

In the CBN strategy the β-reduction can be applied unconditionally, while in the CBV strategy it can be applied only if the argument of the function is a *value.*

4.2 Simple typed lambda calculus

The untyped lambda calculus allows to define every computable function, as proved by Church and Kleene. On the other hands, it give rise to paradoxical behaviors, because it is possible to apply a function to itself. This leads to terms whose process of reduction is infinite. To avoid this effects (which is the formal equivalent to the paradoxes of set theory), one must forbid the unconditional formation of terms, and allow only well-formed terms according to some criterion. In set theory, paradoxes are avoided thanks to a type theory which forbids that a set can be element of itself. In the lambda calculus, one is led to introduce a suitable notion of *type* for each term, and imposing that terms can be combined only when the respective type are matching.

4.2.1 Syntax

The set of terms is given by

$$\begin{aligned} Term \quad M, N, P ::= & x \mid (MN) \mid (\lambda x^A.M) \mid \langle M, N \rangle \mid \\ & \pi_1 M \mid \pi_2 M \mid * \mid inj_1(M) \mid inj_2(M) \mid \\ & case\ M\ of\ x^A \mapsto N, y^B \mapsto P \mid \epsilon_A M \end{aligned} \tag{4.13}$$

where A and B represent *types* (see below). These expressions have an intuitive which is the roughly the following. For the conjunctive part: $\langle M, N \rangle$ is a pair of terms; if M is a pair, $\pi_1 M$ and $\pi_2 M$ are the first and the second element of the pair, respectively; the symbol $*$ is the term standing for the truth value "true". For the disjunctive part: $inj_1(M)$ and $inj_2(M)$ "inject" the term M into a disjunction where M is the first or the second possibility, respectively; the *case* term is essentially and if-then-else construct, which evaluates a disjunction and returns N or P. Finally, $\epsilon_A M$ is the lambda calculus equivalent of the "ex falso quodlibet" logical rule. This intuitive explanation will take a precise form in the typing rules and in the reduction rules.

Not every term built with this recursive procedure is admissible. Types are introduced precisely to select only admissible terms.

In the *simple typed* lambda calculus types are recursively defined starting from a set of *base types* $A, B, \ldots$, by the following BNF

$$Type \quad A, B ::= A \to B \mid A \times B \mid A + B \mid 1 \mid 0 \tag{4.14}$$

The meaning of simple types is the following: $A \to B$ is the type of function from A to B, $A \times B$ is the type of conjunctions, i.e. pairs $\langle M, N \rangle$ where M

has type A and N has type B, $A+B$ is the type of a disjunction, i.e. either a term of type A or a term of type B. The unit types for the conjunction and for the disjunction are 1 and 0 respectively.

To avoid terms which are not well-formed, we introduce *typing rules*. We admit only terms which can by typed according to the typing rules. A *typing judgment* is an expression of the form

$$x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n \vdash M : A \tag{4.15}$$

which means: "assuming x_1 of type A_1, x_2 of type A_2, etc., then M is a term of type A". In the following, we will indicate with Γ a set of typing assumptions for variables. In this way, the previous typing judgment can be written as $\Gamma \vdash M : A$, meaning that M is of type A under the assumptions contained in Γ.

A typing rule is an expression of the form

$$\frac{J_1 \qquad J_2 \qquad \ldots}{J} \tag{4.16}$$

where $J, J_1, J_2, \ldots$ are typing judgments. The intended meaning of a typing rule is that, whenever the typing judgments above the horizontal line hold, then we can state also the typing judgment below the line.

The first group of typing rules is

$$\begin{array}{ll}
(var) & \dfrac{}{\Gamma, x : A \vdash x : A} \\[2ex]
(app) & \dfrac{\Gamma \vdash M : A \to B \qquad \Gamma \vdash N : A}{\Gamma \vdash MN : B} \\[2ex]
(abs) & \dfrac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x^A.M : A \to B}
\end{array} \tag{4.17}$$

The rule (var) is an axiom, namely it can be used to produce a (tautological) typing judgment since there are no hypothesis above the line.

The rules for the conjunction are

$$\begin{array}{ll}
(pair) & \dfrac{\Gamma \vdash M : A \qquad \Gamma \vdash N : B}{\Gamma \vdash \langle M, N \rangle : A \times B} \\[2ex]
(\pi_1) & \dfrac{\Gamma \vdash M : A \times B}{\Gamma \vdash \pi_1 M : A} \\[2ex]
(\pi_2) & \dfrac{\Gamma \vdash M : A \times B}{\Gamma \vdash \pi_2 M : B} \\[2ex]
(*) & \dfrac{}{\Gamma \vdash * : 1}
\end{array} \tag{4.18}$$

We note that also $(*)$ is an axiom as it allows to state an unconditional typing judgment.

The rules for the disjunction are

$$
\begin{array}{ll}
(inj_1) & \dfrac{\Gamma \vdash M : A}{\Gamma \vdash inj_1(M) : A + B} \\
(inj_2) & \dfrac{\Gamma \vdash N : B}{\Gamma \vdash inj_2(N) : A + B} \\
(case) & \dfrac{\Gamma \vdash M : A + B \qquad \Gamma, x : A \vdash N : C \qquad \Gamma, y : B \vdash P : C}{\Gamma \vdash case\ M\ of\ x^A \mapsto N, y^B \mapsto P : C} \\
(\epsilon) & \dfrac{\Gamma \vdash M : 0}{\Gamma \vdash \epsilon_A M : A}
\end{array}
\tag{4.19}
$$

which the *case* rule acts as an if-then-else construct. We note that this is the only rule requiring three hypothesis. We also see the precise sense of the $\epsilon_A M$ term: the (ϵ) rule allow to derive a term of whatever type A we want, provided we have already derived a term of type 0.

Assigning a type to a term means deriving a typing judgment for the term starting from axioms and applying the rules. *Well typed terms* are exactly those terms for which a valid typing judgment can be derived.

4.2.2 Operational semantics

Reduction in simple typed lambda calculus is very similarly to the untyped case. Only, we need some more rules to describe the reduction of new terms

$$
\begin{array}{lc}
(\beta_{\to}) & (\lambda x^A.M)N \to M[N/x] \\
(\beta_{\times 1}) & \pi_1\langle M, N\rangle \to M \\
(\beta_{\times 2}) & \pi_2\langle M, N\rangle \to N \\
(\beta_{+1}) & (case\ inj_1(M)\ of\ x^A \mapsto N, y^B \mapsto P) \to N[M/x] \\
(\beta_{+2}) & (case\ inj_2(M)\ of\ x^A \mapsto N, y^B \mapsto P) \to P[M/y]
\end{array}
\tag{4.20}
$$

Note that the product type $A \times B$ is a *Cartesian product*, which can be reconstructed from its projections, i.e. if $P = \langle M, N\rangle$ then

$$\langle \pi_1(P), \pi_2(P)\rangle \to P. \tag{4.21}$$

On the other hand, in the quantum lambda calculus the product will be more similar to a *tensor* product, and no reduction like the one in Eq. (4.21) exists.

4.3 Quantum lambda calculus

In Ref. [53] it is introduced a lambda calculus for quantum computation, which is the most complete and refined in the literature so far. However, if one want to represent the complete hierarchy of quantum theory this calculus turns out to be inadequate. In this section we review the main features of the quantum lambda calculus and show its drawbacks.

4.3.1 Syntax

In our presentation we restrict to the quantum lambda calculus without classical recursion, since it introduce technical complications and it is not relevant to our discussion. Terms are given by the following BNF

$$\begin{aligned} Term \quad M, N, P ::= {} & c \mid x \mid \lambda x.M \mid MN \mid \langle M, N\rangle \mid * \\ & let\ \langle x, y\rangle = M\ in\ N \mid inj_1(M) \mid inj_2(M) \mid \\ & case\ M\ of\ x^A \mapsto N, y^B \mapsto P \end{aligned} \tag{4.22}$$

The symbol c ranges over a given set of term constants. All the other term constructors have been explained in the section for the simple typed lambda calculus. The only difference is that the projectors $\pi_i M$ have been replaced by a different term $let\ \langle x, y\rangle = M\ in\ N$. The projectors allow to access the single components of a pair, which acts as a *Cartesian* pair. The *let* term, instead, only allow to substitute the whole pair in another term. As usual, the precise meaning of this difference appears in the reduction rules.

The set of term constants contains

- *new*, which takes as input a classical bit, 0 or 1 and creates a quantum bit $|0\rangle$ or $|1\rangle$, respectively.
- *meas*, which takes as input a quantum bit and outputs a classical bit given by the of a measure in the standard basis $|0\rangle$, $|1\rangle$.
- a set of constants U, ranging over some fixed set of universal gates. This includes the Hadamard gate H and the controlled-not two-qubit gate $CNOT$.

We note that in classical lambda calculus the pair can be eliminated by the projections. In the quantum lambda calculus the only elimination rule for the pair acts by substituting both elements at the same time.

The types are formally given by

$$Type \quad A, B ::= qbit \mid !A \mid (A \multimap B) \mid (A \otimes B) \mid (A \oplus B) \mid \top \tag{4.23}$$

The meaning of types in quantum lambda calculus is similar to the classical calculus. In particular, *qbit* is the type of a qubit system, $A \multimap B$ is the type of maps from A to B, $A \otimes B$ is the conjunction of quantum objects (i.e. a tensor product), with its unit type $\top$. $A \oplus B$ is the usual disjunction (we stress that it does not mean the superposition of quantum states). The type $!A$ means that the term is *duplicable*, i.e. it can be re-used. For example, a term like $\lambda x.\langle x, x\rangle$ only makes sense if x has type $!A$ for some A.

As in classical lambda calculus, there are typing rules which allow to derive typing judgment for quantum lambda terms. Because of the formal complicatedness of this typing system we only say that it is cleverly devised (through a notion of *subtyping*) to take into account the presence of duplicable terms.

Many useful quantum computational terms are provided, for example the Deutsch algorithm is represented by a term **Deutsch** defined by

$$\begin{aligned}\mathbf{Deutsch}\ U_f :=& let\ \langle x, y\rangle = (\mathbf{tens}\ H\ (\lambda x.x))(U_f\langle H(new\ 0), H(new\ 1)\rangle)\ in \\ & meas\ x,\end{aligned} \tag{4.24}$$

where

$$\mathbf{tens}\ f\ g := \lambda\langle x, y\rangle.\langle fx, gy\rangle \tag{4.25}$$

(here and in the following we use boldface for the name of a specific lambda term). EPR pair generation corresponds to the term

$$\mathbf{EPR} := \lambda x.CNOT\langle H(new\ 0), new\ 0\rangle \tag{4.26}$$

One can derive the typing judgments

$$\vdash \mathbf{Deutsch}\ :\ !((qbit \otimes qbit \multimap qbit \otimes qbit) \multimap qbit) \tag{4.27}$$

and

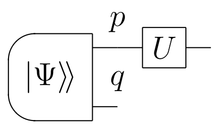
$$\vdash \mathbf{EPR}\ :\ !(\top \multimap (qbit \otimes\ qbit)). \tag{4.28}$$

A quantum lambda term is essentially a classically controlled algorithm which can operate on a quantum state of one ore more qubits. In the calculus, a term applied to a specific quantum state is represented by a *quantum closure*, which is a triple $[Q, L, M]$ where

- Q is a normalized state of n qubits, $Q \in \otimes_{i=1}^{n}\mathbb{C}^2$, for some $n \geqslant 0$
- L is a list of distinct term variables, written as $|x_1, \ldots, x_n\rangle$
- M is a lambda term whose free variables all appear in L

For example, we have the equivalence of a quantum closure with a quantum circuit:

$$[|\Psi\rangle\rangle, |pq\rangle\,, \langle Up, q\rangle] = \tag{4.29}$$



4.3.2 Operational semantics

The operational semantics for quantum lambda calculus has the peculiarity of being a *probabilistic* rewriting procedure on quantum closures. The classical control is treated as in the call-by-value beta reduction for classical lambda calculus

$$\begin{aligned}
&[Q, L, (\lambda x.N)M] \to_1 [Q, L, N[M/x]\\
&[Q, L, let\ \langle x, y\rangle = \langle V, W\rangle\ in\ N] \to_1 [Q, L, N[V/x, W/y]]\\
&[Q, L, case\ inj_1(V)\ of\ x^A \mapsto N, y^B \mapsto P] \to_1 [Q, L, N[V/x]]\\
&[Q, L, case\ inj_2(W)\ of\ x^A \mapsto N, y^B \mapsto P] \to_2 [Q, L, P[W/y]]
\end{aligned} \tag{4.30}$$

We note that the rules for projections are replaced by a rule for the *let* term. This is required, since the tensor product type contains also entangled pairs.

On the other hand, the reduction rules for operations on quantum data are

$$\begin{aligned}
&[Q, |x_1 \dots x_n\rangle\,, U\langle x_{j_1}, \dots, x_{j_n}\rangle] \to_1 [Q', |x_1 \dots x_n\rangle\,, \langle x_{j_1}, \dots, x_{j_n}\rangle]\\
&[\alpha\,|Q_0\rangle + \beta\,|Q_1\rangle\,, |x_1 \dots x_n\rangle\,, meas\ x_i] \to_{|\alpha|^2} [|Q_0\rangle\,, |x_1 \dots x_n\rangle\,, 0]\\
&[\alpha\,|Q_0\rangle + \beta\,|Q_1\rangle\,, |x_1 \dots x_n\rangle\,, meas\ x_i] \to_{|\beta|^2} [|Q_1\rangle\,, |x_1 \dots x_n\rangle\,, 0]\\
&[Q, |x_1 \dots x_n\rangle\,, new\ 0] \to_1 [Q \otimes |0\rangle\,, |x_1 \dots x_{n+1}\rangle\,, x_{n+1}]\\
&[Q, |x_1 \dots x_n\rangle\,, new\ 1] \to_1 [Q \otimes |1\rangle\,, |x_1 \dots x_{n+1}\rangle\,, x_{n+1}]
\end{aligned} \tag{4.31}$$

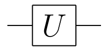
where $\to_p$ means that the reduction happens with probability p. In the first rule Q' is the state obtained applying the unitary represented by the constant U to the state Q. In the rule for measurement, $|Q_0\rangle$ and $|Q_1\rangle$ are normalized states of the form

$$\begin{aligned}
|Q_0\rangle &= \sum_j \alpha_j \left|\phi_j^0\right\rangle \otimes |0\rangle \otimes \left|\psi_j^0\right\rangle,\\
|Q_1\rangle &= \sum_j \beta_j \left|\phi_j^1\right\rangle \otimes |1\rangle \otimes \left|\psi_j^1\right\rangle
\end{aligned} \tag{4.32}$$

where $\left|\phi_j^0\right\rangle$ and $\left|\phi_j^1\right\rangle$ are states of $(i-1)$ qubits, so that the reduction actually simulates the measurement of the i-th qubit.

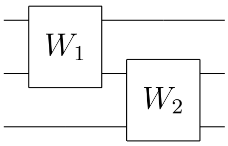
In Ref. [53] it is also proved the *safety property*, namely, a quantum closure with a well-typed lambda term (which is called *quantum program*) always terminates in a value state, and the total probability of the various reduction branches is one.

Now, suppose that we have a unitary channel represented by a term $\mathbf{u}$



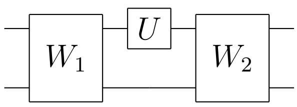
$$\mathbf{u} := U \tag{4.33}$$

and a memory channel represented by a term $\mathbf{t}$



$$\mathbf{t} := let\ \langle z, x\rangle = W_2\langle p, a\rangle\ in\ \langle z, W_1\langle x, q\rangle\rangle \tag{4.34}$$

There is no way to represent the usage of this memory channel as supermap, for example the application of $\mathbf{t}$ to $\mathbf{u}$ as in



(4.35)

This supermap can be represented by a different term $\mathbf{t}'$ defined as

$$\mathbf{t}'\ U := \lambda\langle x, y\rangle.W_2(\mathbf{tens}\ U\ (\lambda x.x))W_1\langle x, y\rangle \tag{4.36}$$

but it is not clear from the calculus that $\mathbf{t}$ and $\mathbf{t}'$ are essentially the same quantum object.

Moreover, whilst the quantum lambda calculus allows to represent the circuits built from the elementary set, it does not take into account the admissible quantum maps which are defined only axiomatically. If the whole set of admissible maps is generated by a suitable universal set of elementary gates, there is a possible workaround: in order to represent every map in the calculus, one only needs to add constant terms for the new elementary gates.

4.4 Categorical aspects

In this section we translate some facts about higher order maps in a categorical language. There is a well-known correspondence between category theory and lambda calculus in the classical setting. Stating the properties of higher order quantum maps in categorical terms can help in understanding their role as a denotational semantics for some formal system.

The category **FdHilb** of finite dimensional Hilbert spaces and linear maps forms a monoidal category with the vectorial tensor product. This product has exponentials, in fact

$$\mathbf{FdHilb}(\mathcal{H}_0 \otimes \mathcal{H}_1, \mathcal{H}_2) \cong \mathbf{FdHilb}(\mathcal{H}_0, \mathcal{H}_1 \otimes \mathcal{H}_2), \tag{4.37}$$

the exponential object $\mathcal{H}_2^{\mathcal{H}_1}$ being isomorphic to $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Positive operators and completely positive maps also form a monoidal category **CPM** with exponentials, thanks to Choi-Jamiołkowski isomorphism. The link product is the evaluation, satisfying

$$\begin{aligned} &\mathfrak{C}(\rho \mapsto S * \rho) = S \\ &\mathfrak{C}(\mathscr{S}) * \rho = \mathscr{S}(\rho). \end{aligned} \tag{4.38}$$

This category does not account for normalization of operators and maps, as is required by probabilistic structure of quantum theory. The following construction is an attempt to build a category with normalized objects. The guiding principle is "every normalization condition is represented by a different object". This is inspired by the classical correspondence between types in lambda calculus and objects in Cartesian closed categories.

We will indicate with $\mathbf{N}$ the set of natural numbers $0 \leqslant n \leqslant (N-1)$. Suppose that the set of Hilbert spaces $\{\mathcal{H}_i\}_{i\in I}$ is labeled by a set of indexes I. We define a category **QK** by the following data:

Indicate with $P_{fin}(I)$ the set of finite subsets of I. Then, an object is a function $\pi^N : \mathbf{2N} \to P_{fin}(I)$, such that

$$n \neq m \Rightarrow \pi^N(n) \cap \pi^N(m) = \emptyset. \tag{4.39}$$

We define the set $\mathrm{comb}(\pi^N)$ of positive operators $R^{(N)}$ on

$$\mathcal{H}^{(N)} := \bigotimes_{j\in\pi^N(n), n\in\mathbf{2N}} \mathcal{H}_j \tag{4.40}$$

satisfying the conditions

$$\begin{aligned} &R^{(j)} * I_{\pi^N(2j-1)} = R^{(j-1)} * I_{\pi^N(2j-2)}, \quad 2 \leqslant j \leqslant N \\ &R^{(1)} * I_{\pi^N(1)} = I_{\pi^N(0)}. \end{aligned} \tag{4.41}$$

The arrows from π^N to π^M are the completely positive maps $\mathscr{S}$,

$$\mathscr{S} : \mathcal{L}(\mathcal{H}^{(N)}) \to \mathcal{L}(\mathcal{H}^{(M)}) \tag{4.42}$$

such that if $R^{(N)} \in \mathrm{comb}(\pi^N)$ then $\mathscr{S}(R^{(N)}) \in \mathrm{comb}(\pi^M)$.

We are looking for categories which are closed in some suitable sense. Theorems in section 2.2.1 prove that some pair of objects in **QK** have an exponential object. Consider for example:

$$\mathbf{QK}(\pi^N, \pi^1) \cong \mathbf{QK}(\mathbb{C}, \pi^{(N+1)}). \tag{4.43}$$

The category **CPTP** is defined as the full subcategory of **QK** obtained retaining only state objects, i.e objects of the form π^1 with $\pi^1(0) = \mathbb{C}$ the tensor identity. **CPTP** is monoidal with tensor product given by the vectorial tensor product. The product of state objects π_1^1 and π_2^1 is the state object $\pi_1^1 \otimes \pi_2^1$ defined by

$$\begin{aligned} &(\pi_1^1 \otimes \pi_2^1)(0) := \mathbb{C} \\ &(\pi_1^1 \otimes \pi_2^1)(1) := \pi_1^1(1) \cup \pi_2^1(1). \end{aligned} \tag{4.44}$$

A state object will be indicated with σ_i where i is a distinguishing label.

The exponential objects of **CPTP** live in **QK**

$$\mathbf{CPTP}(\sigma_0 \otimes \sigma_1, \sigma_2) = \mathbf{QK}(\sigma_0 \otimes \sigma_1, \sigma_2) \cong \mathbf{QK}(\sigma_0, \pi^1_{1|2}). \tag{4.45}$$

An open problem is to define (if possible!) the correct monoidal product on **QK**. A possibility is the vectorial tensor product, but since it requires to assign a specific causal structure to the product object (as discussed in section 2.2.1), it rules out truly higher order objects. Another possibility is the "causal" product, defined as the disjoint union $\pi^N \cup \pi^M$, but in this way one needs to enlarge the category to accommodate such undefined causal structures.

Bibliography

[1] K. Kraus, *States, Effects, and Operations* (Springer, Berlin, 1983).

[2] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, New York, 1973).

[3] W. F. Stinespring, Proc. Amer. Math. Soc.**6**, 211 (1955).

[4] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[5] M. A. Naimark, Iza. Akad. Nauk USSR, Ser. Mat. **4**, 277 (1940).

[6] G. Gutoski and J. Watrous, in *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computation (STOC)*, pag. 565-574 (2007).

[7] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurasek, and E. S. Polzik, Nature 432, 482 (2004);

[8] C. F. Roos, M. Riebe, H. Hấffner,1 W. Hấnsel,1 J. Benhelm, G. P. T. Lancaster, C. Becher, F. Schmidt-Kaler, R. Blatt, Science **304** 1478 (2004)

[9] P. Rabl, D. DeMille, J. M. Doyle, M. D. Lukin, R. J. Schoelkopf, and P. Zoller, Phys. Rev. Lett. **97** 033003 (2006)

[10] V. Scarani *et al*, Rev. Mod. Phys. **77**, 1225 (2005).

[11] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 060401 (2008).

[12] S. D. Bartlett, T. Rudolph, R. W. Spekkens, and P. S. Turner, arXiv:0812.5040.

[13] G. Chiribella, G. M. D'Ariano and M. F. Sacchi, Phys. Rev. A **72** 042338 (2005)

[14] G. Chiribella, G. M. D'Ariano, P. Perinotti, M. F. Sacchi, Phys. Rev. Lett **93**, 180503 (2004)

[15] V. Bužek, R. Derka, and S. Massar, Phys. Rev. Lett. **82**, 2207 (1999).

[16] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. A **70**, 062105 (2004)

[17] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Int. J. Quantum Inf. **4**, 453 (2006)

[18] *Quantum State Estimation*, Lecture Notes in Physics, **649** M. Paris and J. Rehacek Eds., (Springer, Berlin 2004)

[19] G. M. D'Ariano, P. Lo Presti, Phys. Rev. Lett. **86**, 4195 (2001).

[20] G. M. D'Ariano, L. Maccone, and P. Lo Presti Phys. Rev. Lett. **93**, 250407 (2004).

[21] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North Holland, Amsterdam, 1982.

[22] E. Prugovečki, Int. J. Theor. Phys, **16** 321 (1977)

[23] P. Busch, Int. J. Theor. Phys., **30** 1217 (1991).

[24] R. J. Duffin and A. C. Schaeffer, Trans. Am. Math. Soc. **72**, 341 (1952)

[25] P. G. Casazza, Taiw. J. Math. **4**, 129 (2000).

[26] *Quantum State Estimation*, Lecture Notes in Physics, **649** M. Paris and J. Rehacek Eds., (Springer, Berlin 2004)

[27] M.-D. Choi, Lin. Alg. Phys. **10**, 285 (1975).

[28] M. Ziman, Phys. Rev. A **77**, 062112 (2008)

[29] G. Chiribella, G. M. D'Ariano, P. Perinotti, Phys. Rev. Lett. **101**, 180501 (2008).

[30] G. M. D'Ariano, P. Perinotti, M. F. Sacchi, J. Opt. B: Quantum and Semicl. Optics, **6**, S487 (2004).

[31] G. M. D'Ariano, P. Perinotti, Phys. Rev. Lett. **98**, 020403 (2007).

[32] A. J. Scott, J. Phys. A **41**, 055308 (2008).

[33] G. Chiribella, G. M. D'Ariano, D. M. Schlingemann, Phys. Rev. Lett. **98**, 190403 (2007)

[34] P. Walther, A. Zeilinger, Phys. Rev. A **72**, 010302(R) (2005)

[35] M. Mohseni, A. T. Rezakhani, D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).

[36] A. J. Scott, J. Phys. A **39**, 13507 (2006).

[37] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **74**, 012305 (2006).

[38] J. Pearl, *Causality*, (Cambridge University Press, Cambridge, 2009)

[39] N. Herbert, Found. Phys. 12 1171 (1982)

[40] A. Einstein, *The Born-Einstein Letters; Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955* (Walker, New York, 1971).

[41] G. Chiribella, G. M. D'Ariano, P. Perinotti, Phys. Rev. A **81**, 062348 (2010)

[42] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, Phys. Rev. A **64**, 052309 (2001).

[43] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **80**, 022339 (2009).

[44] T. Eggeling, D. Schlingemann, and R. F. Werner, Europhys. Lett. **57**, 782 (2002).

[45] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).

[46] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[47] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[48] M. D. Choi, Linear Algebra Appl. **10**, 285 (1975).

[49] G. Gutoski, Quantum Inf. Comp., **9** 739 (2009)

[50] P. Selinger, *Lecture notes on the Lambda calculus*, available online

[51] H. P. Barendregt, *The Lambda Calculus, its Syntax and Semantics*, North Holland, 1984

[52] J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*, Cambridge University Press, 1989

[53] P. Selinger, B. Valiron, book chapter in S. Gay and I. Mackie, editors, *Semantic Techniques in Quantum Computation*, Cambridge University Press, 2009

[54] G. Chiribella, G. M. D'Ariano, P. Perinotti, B. Valiron, `arXiv:0912.0195`

[55] S. Facchini, G. M. D'Ariano, P. Perinotti, `arXiv:1007.1177`

[56] S. Popescu and D. Rohrlich, Found. Phys. **24** 379 (1994).

[57] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Rev. A **81**, 032324 (2010)

[58] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Rev. Lett. **102** 010404 (2009)

[59] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, IEEE JSTQE vol. 15 no. 6, 1646 (2009)

List of publications

- A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, P. Perinotti *Optimal quantum learning of a unitary transformation*, Phys. Rev. A **81**, 032324 (2010)
- A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, P. Perinotti *Optimal quantum tomography for states, measurements, and transformations*, Phys. Rev. Lett. **102** 010404 (2009)
- A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, P. Perinotti *Optimal quantum tomography*, IEEE JSTQE, vol. 15 no. 6, 1646 (2009)