



**MSc in Computer Science 2017-18**

**Project Dissertation**

**Project Dissertation title:** A computational model for quantum information processing with indefinite causal order

**Term and year of submission:** Trinity term 2018

**Candidate name:** Iskren Vankov

**Title of Degree the dissertation is being submitted under:** MSc in Computer Science

**Word count:** 12923

# Abstract

Higher-order quantum computation is a novel paradigm in computation where inputs and outputs can be quantum gates. In a series of recent works, higher-order computation has been shown to yield advantages over first-order quantum computation. One prominent example of higher-order computation is the quantum SWITCH, which takes two quantum channels as input and combines them in an indefinite causal order. There is currently no established framework for reasoning about protocols with indefinite causal order, where our intuition may easily fail. In this thesis we develop a series of computational models and graphical languages for dealing with higher-order computations that include the quantum SWITCH. We present an analysis of the allowed types of transformations arising in our models. Using our framework, we construct a novel higher-order algorithm. We argue that this algorithm offers an exponential advantage over all algorithms with definite causal order.

# Acknowledgements

I would like to thank my supervisor Giulio Chiribella first for introducing me to the concepts discussed in this thesis, and second for his continuous guidance in my research. I also thank William Simmons for the overview of quantum lambda calculus he provided me with.

I am extremely grateful to Kosta Yanev for his sponsorship of my Oxford Master's degree, the prime result of which is the current thesis.

This work was been partly supported by a grant from the John Templeton Foundation. The opinions expressed in this work are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Contributions list . . . . .	2
1.2	Thesis structure . . . . .	3
<b>2</b>	<b>Density operators and quantum channels</b>	<b>4</b>
2.1	Postulates of quantum mechanics . . . . .	4
2.2	Density operators formalism . . . . .	5
2.3	Quantum channels . . . . .	5
2.4	Quantum information processing . . . . .	7
<b>3</b>	<b>Higher-order quantum computation</b>	<b>8</b>
3.1	Quantum supermaps . . . . .	8
3.2	Indefinite causal order through the quantum SWITCH . . . . .	9
<b>4</b>	<b>Non-signalling supercircuits with definite causal order</b>	<b>12</b>
4.1	Basic computational model . . . . .	12
4.2	Characterisation of the transformations . . . . .	14
4.3	Derived second-order operations . . . . .	15
<b>5</b>	<b>Non-signalling supercircuits with indefinite causal order</b>	<b>18</b>
5.1	Computational model . . . . .	18
5.2	Characterisation of the transformations . . . . .	20
5.3	Derived second-order operations . . . . .	21
<b>6</b>	<b>A novel second-order quantum algorithm</b>	<b>23</b>
6.1	The DQC1 algorithm . . . . .	23
6.2	$N$ -channel DQC1 with indefinite causal order . . . . .	24
6.3	$N$ -channel DQC1 using 2-SWITCHes . . . . .	27
6.4	Measurement modifications to the protocol . . . . .	28
6.5	Arguments for the advantage of our algorithm . . . . .	29
<b>7</b>	<b>Towards higher order computation with signalling channels</b>	<b>31</b>
7.1	Computational model on signalling channels . . . . .	31
7.2	Characterisation of the transformations . . . . .	33
7.3	No-go theorem for naïve SWITCHing of signalling channels . . . . .	34
7.4	The causal $N$ -SWITCH as a direct sum . . . . .	36
<b>8</b>	<b>Conclusions</b>	<b>38</b>
8.1	Summary of results . . . . .	38
8.2	Discussion on further work . . . . .	39

# Chapter 1

## Introduction

Most of the existing quantum algorithms - including the famous Shor's factorising [1] and Grover's search [2] algorithms - are formulated in the circuit model of computation [3]. An important characteristic of the circuit model is that quantum gates are used in a fixed causal order. Quantum mechanics, however, allows for scenarios where the gates' order is not fixed [4, 5, 6].

An example operation that produces indefinite causal order is the quantum SWITCH, proposed by Chiribella, *et. al.* [5]. Given two channels  $\mathcal{A}$  and  $\mathcal{B}$ , the quantum SWITCH can output a combination of the two channels in a superposition of both compositional orders  $\mathcal{A} \circ \mathcal{B}$  and  $\mathcal{B} \circ \mathcal{A}$ .

The quantum SWITCH is a higher-order operation that acts on quantum channels. It is an example of a more general class of maps, called *quantum supermaps* [7], which transform quantum gates into quantum gates.

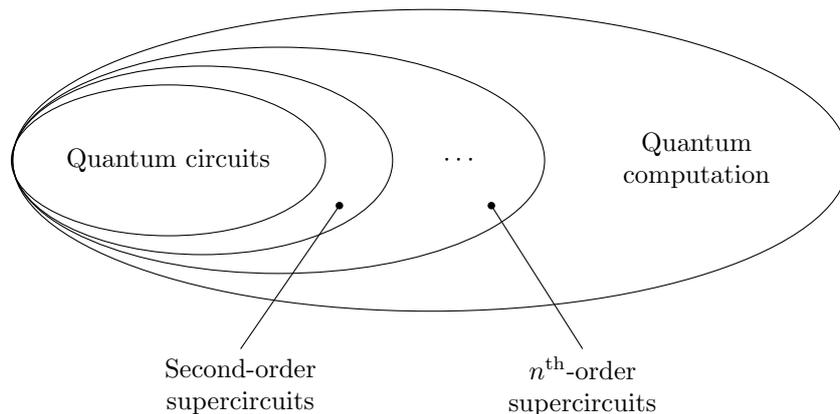
Going further, maps from supermaps to supermaps can be defined, and so on indefinitely. In this way one obtains an infinite hierarchy of higher-order transformations, where the first order describes quantum operations acting on quantum states, the second order quantum supermaps acting on quantum operations, the third order maps between quantum supermaps, and so on. This is the hierarchy of *higher-order quantum computation*. In principle, a full-blown model of quantum computation should include all levels of this hierarchy (Figure 1.1).

There already exist results that indicate advantages of higher-order computation with indefinite causal order over algorithms with fixed causal order. Chiribella in Ref. [8] showed an advantage of the quantum SWITCH in a toy problem where the task is to decide whether two operators commute or anti-commute. Araújo, *et. al.* have proposed in Ref. [9] a problem which is conjectured to have exponential complexity classically, quadratic for quantum with fixed causal order, and only linear for quantum with indefinite causal order. (The number of queries to  $n$  black boxes is  $O(2^n)$ ,  $O(n^2)$ , and  $O(n)$  respectively.)

Despite the existence of the above algorithms, no formal model of quantum computation with indefinite causal order has been proposed so far. In this thesis we fill this gap by rigorously defining a model for second-order computation using the quantum SWITCH. We proceed to evaluate our model in terms of the types of maps it allows, thus exploring the potential of our model of second-order computation.

We also develop a graphical language for describing second-order circuits, which we call *supercircuits*. Our motivation is that there does not yet exist a convenient way of expressing circuits with indefinite causal order.

The importance of having a graphical language is highlighted in cases where our intuition may fail. An example is the communication protocol proposed by Ebler, *et al.* in Ref. [10]. It describes how non-zero communication via two completely depolarising channels is made possible by applying them with indefinite causal order. The result is counter-intuitive because the two fixed composition orders of two identical channels are the same, and so one



**Figure 1.1:** A diagram of the infinite hierarchy of higher-order quantum computation.

would not expect to gain any advantage by putting them in superposition.

We utilise our framework in order to construct a novel second-order algorithm, inspired by the DQC1 algorithm [14]. The algorithm accepts  $N$  black box quantum channels as input and estimates a certain quantity that depends on them. We provide evidence that our algorithm has an exponential advantage over any algorithm that uses the black boxes in definite causal order. If this expectation is proven correct, our algorithm would be the first particular example of exponential advantage of computation with indefinite causal order over computation with fixed causal order. In comparison, the strongest currently known results for such advantage is only polynomial (examples are Refs. [9, 15]). While finding a formal proof of the expected advantage is still an open problem, in this thesis we provide evidence for it. As a matter of fact, proving an asymptotic advantage over arbitrary algorithms with definite causal order is a challenging problem that has not been solved for any of the algorithms presented so far in the literature.

Our work is timely because the quantum SWITCH has recently been implemented experimentally, hence verifying computation with indefinite causal order. Procopio *et. al.* [15] then designed and executed an experiment that solves this task using the quantum SWITCH, hence providing experimental evidence for its physical realisation.

Further experimental evidence was provided by Rubino, *et. al.* in Ref. [16] with another experiment involving indefinite causal order. There they define a *causal witness* to be a set of measurements such that some outcome is incompatible with *any* fixed causal order. Using that causal witness, their experiment guarantees that genuine indefinite causal order was used.

These positive experimental results indicate that higher-order computation with indefinite causal order can be utilised in a practical way, rather than it being a purely theoretical framework. This highlights the practical importance of defining well the computational model made possible by higher-order maps, and shows our motivation for doing so.

## 1.1 Contributions list

- Non-signalling supercircuits with definite causal order
  - Defined computational model
  - Developed graphical language
  - Characterised the allowed transformations
- Non-signalling supercircuits with indefinite causal order

- Defined computational model
- Developed graphical language
- Characterised the allowed transformations
- Signalling supercircuits with definite causal order
  - Defined computational model
  - Developed graphical language
  - Characterised the allowed transformations
- Signalling supercircuits with indefinite causal order
  - Proved a no-go result on the decomposability of the causal  $N$ -SWITCH
- DQC1 with indefinite causal order
  - Proposed a novel quantum algorithm
  - Provided arguments why there is an expected exponential advantage over algorithms with fixed causal order
- Miscellaneous
  - Developed LaTeX tools for typesetting supercircuits

## 1.2 Thesis structure

This thesis is structured as follows. We begin by introducing the mathematical framework of density matrices and quantum channels in Chapter 2. In Chapter 3 we overview recent developments on higher order quantum computation that serve as the basis for our work. Our original contributions begin in Chapter 4 where we define a model for non-signalling supercircuits with definite causal order, and we prove results on the transformations it allows. In the same chapter we present the core of our graphical language. In Chapter 5 we enrich our model with the quantum SWITCH, hence introducing indefinite causal order. We again prove results on the allowed transformations and extend the graphical language to include the quantum SWITCH. In this framework, we construct a novel second-order quantum algorithm, presented in Chapter 6. In the same chapter we provide evidence for an exponential advantage of our algorithm over algorithms with definite causal order. A further extension of our model is considered in Chapter 7 where we include arbitrary signalling channels. In this model we show that the quantum SWITCH of  $N$  channels cannot be decomposed into elementary SWITCHes on 2 channels. Finally, in Chapter 8 we summarise our results and envisage further developments of our research.

## Chapter 2

# Density operators and quantum channels

In this chapter we introduce the mathematical framework of density operators and quantum channels, which is used throughout this thesis.

### 2.1 Postulates of quantum mechanics

Summarised concisely, quantum mechanics can be described by 4 postulates. (For more details, see e.g. Shankar's book on quantum mechanics [17].) These 4 basic postulates form the following mathematical framework. (P1) The state of a quantum system is fully described by a vector in a Hilbert space. In finite dimensions a Hilbert space is an inner product space over the complex numbers. States are represented in Dirac notation as *kets*  $|\psi\rangle$ , and their Hermitian conjugates as *bras*  $\langle\psi|$ . Usually it is assumed that these states are normalized, which means that their respective vectors in the space are of unit length:

$$\langle\psi|\psi\rangle = 1 \tag{2.1}$$

(P2) The evolution of a closed system's state is described by a unitary operator  $U$ :

$$|\psi\rangle \xrightarrow{U} U|\psi\rangle \tag{2.2}$$

Measurement allows us to extract information about a system's state. However, then that system cannot be treated as closed, and hence it is not unitary and in general not reversible. (P3) Measurement is mathematically described by a set of operators  $\{M_i\}$  satisfying the completeness relation

$$\sum_i M_i^\dagger M_i = I, \tag{2.3}$$

needed to conserve total probability. Each of these operators corresponds to a particular outcome  $i$  of the measurement which is yielded with probability

$$\mathbb{P}(i) = \langle\psi| M_i^\dagger M_i |\psi\rangle \tag{2.4}$$

and for each of them respectively, the system is transformed (non-unitarily) into another state:

$$|\psi\rangle \longrightarrow \frac{M_i |\psi\rangle}{\left(\langle\psi| M_i^\dagger M_i |\psi\rangle\right)^{1/2}} \tag{2.5}$$

Finally, (P4) composite systems are treated by taking the *tensor product* of their individual Hilbert spaces.

## 2.2 Density operators formalism

Having briefly described the basic postulates of quantum mechanics, we proceed to show how they give rise to the formalism of *density operators*, first introduced by von Neumann in Ref. [18]. This alternative framework is equivalent mathematically to the one described above. However, it is more convenient to work with, and as such is the one used throughout this thesis.

A density operator is a positive operator with trace one that acts on the state space. A system's state is completely described by its density operator. Let us elaborate on the correspondence of this statement to the first postulate quoted above - that a system's state is completely described by a vector. Consider a system which is in one of some states  $\{|\psi_i\rangle\}$  with probability  $p_i$  for each, respectively. This state is described by the following density operator:

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \quad (2.6)$$

A state is called *pure* if it can be represented by a single ket  $|\psi\rangle$ , i.e. its density matrix is of the form  $\rho = |\psi\rangle \langle\psi|$ . Otherwise, the state is described by a convex combination of pure states and is referred to as *mixed*. This distinction illustrates the convenience of the density operators formalism within which both pure and mixed states are represented by a single object  $\rho$ . In the previous framework, however, mixed states need be represented rather clumsily as a set  $\{p_i, |\psi_i\rangle\}$ .

Having reformulated the way we describe a system's state, we proceed to accordingly reformulate the other postulates from the previous section. Unitary evolution (Equation 2.2) for density matrices becomes:

$$\rho \xrightarrow{U} U\rho U^\dagger \quad (2.7)$$

Regarding measurement, Equations 2.4 and 2.5 become respectively

$$\mathbb{P}(i) = \text{Tr} \left[ M_i^\dagger M_i \rho \right] \quad (2.8)$$

and

$$\rho \longrightarrow \frac{M_i \rho M_i^\dagger}{\text{Tr} \left[ M_i^\dagger M_i \rho \right]} \quad (2.9)$$

Composite systems are still described by taking the tensor product of their Hilbert spaces. More precisely, given  $n$  systems in states  $\rho_1 \in \mathcal{H}_1, \dots, \rho_n \in \mathcal{H}_n$ , their composite state is  $\rho \in \mathcal{H}$  where

$$\rho = \bigotimes_{i=1}^n \rho_i \quad \text{and} \quad \mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i \quad (2.10)$$

In order to obtain the reduced density operator  $\rho_A$  of a subsystem  $A$  of a composite system  $A \otimes B$ , we use the *partial trace*:

$$\rho_A = \text{Tr}_B [\rho] \quad (2.11)$$

## 2.3 Quantum channels

Quantum channels describe quantum operations beyond unitary evolutions and measurements. Kraus showed in Ref. [19] that it is sufficient and necessary for an arbitrary quantum operation to be completely positive (CP) and non-trace-increasing. Quantum channels are

a subset of quantum operations, described by completely positive trace-preserving (CPTP) maps. They describe deterministic quantum operations that do not include measurement (which is not a trace-preserving operation).

Kraus' theorem provides a particular form in which all quantum operations can be expressed. It states that for any map  $\mathcal{E}$  which transforms a density matrix of a Hilbert space  $\mathcal{H}_1$  into a density matrix of a Hilbert space  $\mathcal{H}_2$ , there exists a set of operators  $\{E_i\} : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ , referred to as Kraus operators, such that

$$\sum_i E_i^\dagger E_i \leq I, \quad (2.12)$$

where the action of the map on the density matrix is described by

$$\rho \xrightarrow{\mathcal{E}} \sum_i E_i \rho E_i^\dagger \quad (2.13)$$

For CPTP maps the completeness constraint (Equation 2.12) becomes strict:

$$\sum_i E_i^\dagger E_i = I \quad (2.14)$$

The Kraus decomposition is not unique, i.e. different sets of Kraus operators can implement the same quantum map.

A convenient way to represent quantum channels is using the Choi-Jamiołkowski isomorphism [20, 21, 22]. It establishes a correspondence between quantum channels (CPTPs) and quantum states (density operators). In order to give the mathematical definition of the Choi-Jamiołkowski isomorphism, we need to first introduce a bit of notation.

Let  $\mathcal{L}(\mathcal{H})$  denote the set of linear operators on a finite-dimensional Hilbert space  $\mathcal{H}$ . Let the set of linear operators from a finite-dimensional Hilbert space  $\mathcal{H}_1$  to a finite-dimensional Hilbert space  $\mathcal{H}_2$  be denoted by  $\mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2)$ . The set of linear operators from  $\mathcal{L}(\mathcal{H}_1)$  to  $\mathcal{L}(\mathcal{H}_2)$  is denoted by  $\mathcal{L}(\mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2))$ . Consider a quantum channel  $\mathcal{E} : \mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2)$  which transforms density matrices in a Hilbert space  $\mathcal{H}_1$  into density matrices in a Hilbert space  $\mathcal{H}_2$ . We can construct a density matrix  $\rho$  in the joint Hilbert space  $\mathcal{H}_1 \otimes \mathcal{H}_2$  in the following way:

$$\rho = (\mathcal{E} \otimes \mathcal{I})(|I\rangle\rangle \langle\langle I|), \quad (2.15)$$

where  $\mathcal{I}$  is the identity map on  $\mathcal{H}_1$ , and  $|I\rangle\rangle$  is the maximally entangled state

$$|I\rangle\rangle = \sum_i |i\rangle \otimes |i\rangle \in \mathcal{H}_1^{\otimes 2}. \quad (2.16)$$

(The double ket notation was first introduced by Royer in Ref. [23] and further developed by D'Ariano, *et. al.* in Ref. [24].) The constructed density operator  $\rho$  is referred to as the *Choi matrix* of  $\mathcal{E}$ . The Choi-Jamiołkowski isomorphism describes the correspondence between  $\mathcal{E}$  and  $\rho$  in the following way:

$$\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1) \rightarrow \mathcal{L}(\mathcal{H}_2)) \cong \rho \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2). \quad (2.17)$$

The inverse of Equation 2.15 determines the action of the channel  $\mathcal{E}$  acting on an arbitrary density operator  $\sigma$  to be

$$\mathcal{E}(\sigma) = \text{Tr}_{\mathcal{H}_1} \left[ (I \otimes \sigma^T) \rho \right]. \quad (2.18)$$

Here  $\sigma^T$  is the transpose of  $\sigma$  in the same basis  $\{|i\rangle\}$  used in Equation 2.16.

Non-signalling channels are a particular type of quantum channels which are of special interest in this thesis. By definition, they are multipartite channels which allow any causal order between their parts. This intrinsic property of theirs will be of use to us when we

discuss indefinite causal order. Their formal definition is given as follows.

**Definition 2.1 (Non-signalling channels)**

Consider a channel  $\mathcal{C}$  that transforms systems  $A, B, \dots$  into  $A', B', \dots$  respectively:

$$\begin{array}{ccc}
 \text{---} A & & A' \text{---} \\
 \text{---} B & & B' \text{---} \\
 \vdots & & \vdots \\
 \text{---} & & \text{---}
 \end{array}
 \begin{array}{c}
 \boxed{\mathcal{C}} \\
 \\
 \\
 \\
 \end{array}
 \begin{array}{ccc}
 \text{---} A' & & A' \\
 \text{---} B' & & B' \\
 \vdots & & \vdots \\
 \text{---} & & \text{---}
 \end{array}
 \quad (2.19)$$

We call the channel  $\mathcal{C}$  non-signalling or causal (Ref. [25]) if it has the following property:

$$\begin{array}{ccc}
 \text{---} A & & A' \text{---} \boxed{\text{Tr}} \\
 \text{---} B & & B' \text{---} \\
 \vdots & & \vdots \\
 \text{---} & & \text{---}
 \end{array}
 \begin{array}{c}
 \boxed{\mathcal{C}} \\
 \\
 \\
 \\
 \end{array}
 =
 \begin{array}{ccc}
 \text{---} A \boxed{\text{Tr}} & & A' \\
 \text{---} B & & B' \\
 \vdots & & \vdots \\
 \text{---} & & \text{---}
 \end{array}
 \begin{array}{c}
 \boxed{\mathcal{C}} \\
 \\
 \\
 \\
 \end{array}
 \quad (2.20)$$

and similarly for all the other transformations  $B \rightarrow B'$ , etc.

Non-signalling channels are characterised in Ref. [5] to be affine combinations of product channels. The characterisation is as follows. Let  $\text{CPTP}(A \rightarrow A')$  be the set of CPTP maps (quantum channels) from system  $A$  to system  $A'$ . Then a quantum channel  $\mathcal{C} \in \text{CPTP}(AB \rightarrow A'B')$  is non-signalling if and only if

$$\mathcal{C} = \sum_i \lambda_i \mathcal{F}_i \otimes \mathcal{G}_i \quad (2.21)$$

where  $\forall i. \lambda_i \in \mathbb{R}, \mathcal{F}_i \in \text{CPTP}(A \rightarrow A'), \mathcal{G}_i \in \text{CPTP}(B \rightarrow B')$ , and  $\sum_i \lambda_i = 1$ .

## 2.4 Quantum information processing

The standard circuit model describes the evolution of finite-dimensional quantum systems through a sequence of quantum gates. Here we provide one possible way to define a quantum circuit. It is obtained by slicing the circuit into sequential timesteps and representing each timestep by a quantum channel.

**Definition 2.2 (Quantum circuits)**

A quantum circuit describes how an ordered sequence of quantum channels transforms the state of a quantum system. It is fully described by the following three objects:

- A finite-dimensional quantum system  $A$  (e.g. a set of qubits)
- A set of channels that act on  $A$
- A total order on the channels which fixes their order in time.

*Qubits*, the most commonly used type of quantum system, are described by states in two-dimensional Hilbert spaces  $\mathcal{H}_2$ . Different physical implementations of such two-level quantum systems could be electron spins (up/down), photon polarisation (horizontal/vertical), and others. Throughout this thesis, however, we are not concerned with the physical realisation of quantum computation, but rather its theoretical side. Three-level systems in  $\mathcal{H}_3$  are called *qutrits*, and general higher-dimensional systems are referred to as *qudits*. We denote the trivial system type, which is the type of measurement outcomes, simply as 1, where  $\mathcal{H}_1 = \mathbb{C}$ . A more general discussion on the standard model of quantum computation can be found for instance in Nielsen and Chuang's book [26]. In this thesis we consider higher orders of computation, which generalise this basic scenario.

# Chapter 3

## Higher-order quantum computation

### 3.1 Quantum supermaps

Quantum supermaps are maps between quantum operations. In Ref. [7] two types of quantum supermaps are distinguished: *deterministic* supermaps that transform channels into channels (CPTP to CPTP maps), and *probabilistic* supermaps that transform channels into operations (CPTP to CP maps). Deterministic supermaps on quantum channels take a central place in the research presented in this thesis. Therefore we deem important to quote here their formal definition provided in Ref. [5], mostly following the original notation used in that paper. A few intermediate definitions need to be introduced first. Let  $\text{St}(A)$  denote the set states of a system  $A$ . As discussed in section 2.2, this is exactly the unit-trace positive operators subset of  $\mathcal{L}(\mathcal{H}_A)$ . Denote the Hermitian-preserving subset of  $\mathcal{L}(\mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_{A'}))$  by  $\text{Herm}(A \rightarrow A')$ . The *marginal* on  $A \rightarrow A'$  of a channel  $\mathcal{C} \in \text{CPTP}(AE \rightarrow AE')$  relative to a state  $\sigma \in \text{St}(E)$  is defined to be the channel  $\mathcal{C}_\sigma$ , where

$$\mathcal{C}_\sigma(\rho) = \text{Tr}_{E'}[\mathcal{C}(\rho \otimes \sigma)] \quad (3.1)$$

Finally, the *extension of a set of channels*  $S \subseteq \text{CPTP}(A \rightarrow A')$  in  $\text{CPTP}(AE \rightarrow AE')$ , for some systems  $E, E'$  is the set

$$\text{Ext}_{E \rightarrow E'}(S) := \{ \mathcal{C} \in \text{CPTP}(AE \rightarrow AE') \mid \forall \sigma \in \text{St}(E). \mathcal{C}_\sigma \in S \} \quad (3.2)$$

Using all these new terms, Ref. [5] defines a deterministic supermap on quantum channels in the following way.

**Definition 3.1 (Deterministic supermaps on quantum channels)**

A deterministic supermap of input type  $\text{CPTP}(A \rightarrow A')$  and output type  $\text{CPTP}(B \rightarrow B')$  is a map  $\mathcal{S}$  from  $\text{Herm}(A \rightarrow A')$  to  $\text{Herm}(B \rightarrow B')$  such that for any systems  $E, E'$

$$\forall \mathcal{C} \in \text{Ext}_{E \rightarrow E'}[\text{CPTP}(A \rightarrow A')]. (\mathcal{S} \otimes \mathcal{I}_{E \rightarrow E'}) (\mathcal{C}) \in \text{Ext}_{E \rightarrow E'}[\text{CPTP}(B \rightarrow B')] \quad (3.3)$$

Two types of generalisations to supermaps are proposed by Chiribella, *et al.* in Ref. [5]. The first generalisation is a recursively generated infinite hierarchy of higher-order maps. The first order describes quantum operations acting on quantum states. The second order are the already described quantum supermaps, the third order are maps between supermaps, etc. Every tier gives rise to the next one by considering maps between its elements. Throughout this thesis we used the terms “first-order computation” and “second-order computation” in the sense of the hierarchy just described. Quantum circuits (Definition 2.2) are an example

of first-order computation.

The other generalisation from Ref. [5] regards supermaps between restricted sets of channels, rather than only  $\text{CPTP} \rightarrow \text{CPTP}$ , as in Definition 3.1. Ref. [5] proceeds in this direction by rigorously defining deterministic supermaps on product and non-signalling channels. We shall not present the definitions here for the sake of brevity since they are similar in spirit to Definition 3.1, and are readily available for further reference in the paper. In principle, an even further generalisation could be considered by mixing the two already mentioned. Namely, one could consider maps between restricted maps of different levels of the hierarchy. For instance, a third-order map from (restricted first-order) non-signalling channels to (second-order) supermaps.

## 3.2 Indefinite causal order through the quantum SWITCH

Some higher-order maps can be realised by quantum circuits. Ref. [27] characterises such maps for all tiers of the hierarchy. In general, however, higher-order maps cannot be realised using quantum circuits. One example is the novel higher-order quantum resource, called the quantum SWITCH [5].

In order to understand the quantum SWITCH, it is useful to start from its classical version. The classical SWITCH is an operation that accepts two unknown quantum boxes  $\mathcal{F}$  and  $\mathcal{G}$ , and a classical control bit  $x$ , and outputs a compositional ordering of the two boxes depending on the control:

$$\text{SWITCH}(x, \mathcal{F}, \mathcal{G}) = \begin{cases} \mathcal{G} \circ \mathcal{F} & \text{if } x = 1 \\ \mathcal{F} \circ \mathcal{G} & \text{if } x = 0 \end{cases} \quad (3.4)$$

The classical SWITCH is an example of a higher-order operation that cannot be expressed as a circuit. In particular, Ref. [5] proves a no-go theorem that the SWITCH, as defined in Equation 3.4, cannot be implemented by a circuit, in which the oracles  $\mathcal{F}$  and  $\mathcal{G}$  are called a single time in a fixed causal order. The proof shows that a circuit implementation implies the existence of deterministic time travel, and is hence impossible *ad absurdum*.

In order to appreciate what qualities of the SWITCH make it non-computable by a regular circuit, let us discuss the four ways around the no-go theorem, proposed in the paper. Relaxing any of these four requirements makes the SWITCH simulatable by a quantum circuit. First,  $\mathcal{F}$  and  $\mathcal{G}$  are given as black boxes, i.e. we have zero knowledge about these channels. Second, they can be called only once by the circuit. In fact, if the channels are unitaries, a circuit realisation exists that only needs two copies of one of them. Third, it is assumed that closed timelike curves do not exist. This requirement is more a constraint on physics, rather than an intrinsic property of the SWITCH. It is important nonetheless, since the SWITCH could be implemented by utilising signalling from the future to the past. Finally, the no-go theorem applies to *deterministic* circuits only. Otherwise, using post-selected teleportation, a circuit could implement the SWITCH probabilistically. It is shown, however, that the probability for such a probabilistic circuit to successfully compute the SWITCH is upper-bounded by an exponentially small quantity in the number of qubits the channels act on.

Having discussed the classical SWITCH, let us consider possible generalisations. In principle, the SWITCH outputs superpositions of permutations of its input channels. The channels need not be composed in order to produce a single channel as output. This leads to the more general SWITCH, shown in Figure 3.1, which simply reorders its input channels. The permuted channels could then be composed, achieving the same result, but further transformations are also allowed, such as inserting another channel in between them.

A second important generalisation is allowing the SWITCH to have quantum control, instead of just classical. That establishes quantum correlation between the two orders ( $\mathcal{A} \otimes \mathcal{B}$

$$\text{SWITCH} \left( \boxed{\mathcal{A}}, \boxed{\mathcal{B}} \right) = \begin{array}{|c|} \hline \boxed{\mathcal{A}} \\ \hline \boxed{\mathcal{B}} \\ \hline \end{array} + \begin{array}{|c|} \hline \boxed{\mathcal{B}} \\ \hline \boxed{\mathcal{A}} \\ \hline \end{array}$$

**Figure 3.1:** An illustration of the quantum SWITCH acting as a second-order map on two quantum channels. The SWITCH maps a pair of quantum channels into a superposition of their permutations. The permutations can correspond either to sequential ( $\circ$ ) or parallel ( $\otimes$ ) composition. Some authors (e.g. [9, 10]) use a variation of the SWITCH that composes sequentially its input channels, i.e. the output is  $\mathcal{B} \circ \mathcal{A} + \mathcal{A} \circ \mathcal{B}$ . Instead, we consider a more general SWITCH that does not compose its input channels. This gives us the additional freedom to insert other channels between  $\mathcal{A}$  and  $\mathcal{B}$  after they are SWITCHed.

and  $\mathcal{B} \otimes \mathcal{A}$  for input channels  $\mathcal{A}$  and  $\mathcal{B}$ ) and the state of the external control system. The result is a way of creating and utilising superpositions of causal orders - *indefinite causal order*. A formal definition goes as follows.

**Definition 3.2 (Quantum SWITCH)**

Consider two quantum channels  $\mathcal{C}^{(1)}, \mathcal{C}^{(2)}$  with Kraus decompositions

$$\mathcal{C}^{(i)}(\rho) = \sum_j K_j^{(i)} \rho \left( K_j^{(i)} \right)^\dagger \quad \text{for } i = 1, 2 \quad (3.5)$$

The quantum SWITCH acting on these two channels is then defined (Ref. [5]) as

$$\mathfrak{S} \left( \mathcal{C}^{(1)}, \mathcal{C}^{(2)} \right) (\rho) = \sum_{ij} W_{ij} \rho W_{ij}^\dagger \quad (3.6)$$

where the Kraus operators  $W_{ij}$  are given by

$$W_{ij} = |0\rangle \langle 0| \otimes K_i^{(1)} \otimes K_j^{(2)} + |1\rangle \langle 1| \otimes K_j^{(2)} \otimes K_i^{(1)} \quad (3.7)$$

**Remark 3.1** Notice that we define the SWITCH without composing the channels at its output. This allows for further SWITCHing of them and for inserting other channels in between them. Some authors use an alternative definition of the SWITCH, where the channels are composed at the output. More precisely, Equation 3.7 becomes

$$W_{ij} = |0\rangle \langle 0| \otimes K_j^{(2)} \circ K_i^{(1)} + |1\rangle \langle 1| \otimes K_i^{(1)} \circ K_j^{(2)} \quad (3.8)$$

Throughout the thesis we use our definition (Equation 3.7) where the channels are not composed at the output.

A further extension is to define a supermap that operates similarly to the quantum SWITCH but acts on an arbitrary number of channels. The  $N$ -SWITCH is defined to be such a map that outputs permutations of  $N$  channels. There are  $N!$  permutations, therefore we choose the control of the  $N$ -SWITCH to be an  $N!$ -level quantum system. In order to give a formal definition, we first need a bit of notation. Let  $\text{Perm}(N)$  be the set of all permutations of  $N$  elements. We denote the  $j$ th element of a permutation  $\pi \in \text{Perm}(N)$  by  $\pi(j)$ .

**Definition 3.3 (Quantum  $N$ -SWITCH)**

Consider a set  $\{\mathcal{C}^{(i)}\}$  of  $N$  quantum channels with Kraus decompositions

$$\mathcal{C}^{(i)}(\rho) = \sum_{j=1}^{d_i} K_j^{(i)} \rho \left( K_j^{(i)} \right)^\dagger, \quad (3.9)$$

where  $d_i$  is the number of Kraus operators of the  $i^{\text{th}}$  channel. Consider an  $N!$ -dimensional quantum system  $C$  with orthonormal basis  $\{|\pi\rangle \mid \pi \in \text{Perm}(N)\}$ . The quantum  $N$ -SWITCH is a supermap that, given the set of channels  $\{\mathcal{C}^{(i)}\}$  as input and using  $C$  as control, outputs a quantum map with Kraus operators  $W_{i_1, \dots, i_N}$  given by

$$W_{i_1, \dots, i_N} = \sum_{\pi \in \text{Perm}(N)} |\pi\rangle \langle \pi| \otimes \bigotimes_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \quad (3.10)$$

An important question is whether the  $N$ -SWITCH can be decomposed into a number of instances of the 2-SWITCH. This has been shown to be true in Ref. [28] for non-signalling channels with  $O(N^2)$  uses of the 2-SWITCH. This result has been strengthened to  $O(N \log_2 N)$  in Ref. [29].

When decomposing the  $N$ -SWITCH into a sequence of  $K$  2-SWITCHes, the control system  $E$  is a set of  $K$  qubits. In general  $2^K$  need not be equal to  $N!$ . Therefore some of  $E$ 's orthogonal states will correspond to identical permutations in  $\text{Perm}(N)$ . This deviates from Definition 3.3 but does not change how the  $N$ -SWITCH works. The mapping between states of  $E$  and resulting permutations in  $\text{Perm}(N)$  can be defined arbitrarily, as long as each permutation is related to at least one state.

**Remark 3.2** Consider a physical realisation of the  $N$ -SWITCH as a sequence of 2-SWITCHes that only allows product states of their controls. This can be true when, for instance, an application of the SWITCH involves preparing a new qubit to be used as control, without allowing it to interact with other external systems. For such product state-controlled SWITCHes it is no longer true that any superposition of permutations can be obtained. For instance, the equally weighted superposition for three channels is impossible to generate. It is an interesting question we leave open to categorise which superpositions are obtainable and which are not, using only product states as control.

The  $N$ -SWITCH has been utilised in Ref. [9] which proposes a problem where its use is conjectured to reduce computational complexity. On the experimental side, an optical implementation of the  $N$ -SWITCH has been proposed in Ref. [30] as a basis for further experiments.

## Chapter 4

# Non-signalling supercircuits with definite causal order

Higher-order maps form an infinite hierarchy with very sophisticated structure. We now limit our attention to a fragment of this hierarchy, which is still more expressive than the usual quantum circuits.

The standard quantum circuit model consists of wires representing qubits, and boxes representing quantum channels. We now consider a model of second-order *supercircuits* where the wires represent time slots to which quantum channels are allocated, and the boxes represent higher-order transformations. We call these wires *timelines*.

### 4.1 Basic computational model

The first and simplest model for such supercircuits we can define goes as follows.

**Definition 4.1 (Non-signalling supercircuits with definite causal order)**

*A non-signalling supercircuit with definite causal order describes how a sequence of quantum supermaps acts on an  $N$ -partite non-signalling channel. In order to process such input, we allow the following three supermaps:*

- Sequential composition - *the sequential composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{B} \circ \mathcal{A}$ .*
- Parallel composition - *the parallel composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{A} \otimes \mathcal{B}$ .*
- Swap of timelines - *deterministic reordering in time of channels.*

In this model the channels are composed only with a definite causal order. Also, there are no allowed operations between supermaps. Hence the model is of second-order computation with definite causal order.

Now let us consider how to depict graphically supercircuits on such non-signalling channels. First, let any non-signalling channel  $\mathcal{C}$  that transforms systems of dimensionalities  $A, B, \dots$  into  $A', B', \dots$  respectively in the first-level circuit, be represented as:

$$\left( \begin{array}{c} \overline{A \rightarrow A'} \\ \overline{B \rightarrow B'} \\ \vdots \\ \overline{\phantom{A \rightarrow A'}} \end{array} \right) \mathcal{C} \quad (4.1)$$

in the supercircuit. The labels  $A, A', B, B', \dots$  can stand for the Hilbert spaces of quantum systems of different dimensionalities. For instance, a qubit -  $\mathcal{H}_2$ , a qutrit -  $\mathcal{H}_3$ , two qubits -  $\mathcal{H}_4$ , etc. Note how *an operation* in a circuit becomes *a state* in a supercircuit. As described above, the wires coming out of it are the *timelines* that represent *when* this operation will be applied. In order to differentiate between first and second order wires, we use a thick style for second-order ones, i.e. timelines. Without loss of generality, let timelines be ordered from top to bottom, i.e. a timeline *above* another means that its corresponding operation happens *before or at the same time as* the operation of the other one. An  $N$ -partite input channel is depicted with  $N$  timelines.

The labels only carry information about the *dimensionality* of the systems acted on. Hence operations with matching dimensionalities could act on the same system, or on different systems. In order to illustrate this point, consider two channels, both of type  $\mathcal{H}_2 \rightarrow \mathcal{H}_2$ , i.e. from a single qubit to a single qubit. We may want to indicate that they act on two separate qubits, or that they act on the same qubit, in some order. This is achieved by utilizing the two second-order composition operations - sequential and parallel. The former - sequential composition - allows the application of two operations on *the same* system under the following condition. It is only defined for operations where the output type of the top (earlier) one matches the input type of the bottom (later) one. Suppose we have operations  $\mathcal{A} : A \rightarrow B$  and  $\mathcal{B} : B \rightarrow C$ . Graphically we denote their sequential composition  $\mathcal{B} \circ \mathcal{A}$  on the same physical system as:

$$\mathcal{B} \circ \mathcal{A} \quad \longleftrightarrow \quad \begin{array}{c} \textcircled{\mathcal{A}} \xrightarrow{A \rightarrow B} \\ \textcircled{\mathcal{B}} \xrightarrow{B \rightarrow C} \end{array} \quad \begin{array}{c} \text{---} \\ \diagdown \\ \text{---} \\ \diagup \\ \text{---} \end{array} \quad \textcircled{\mathcal{A} \rightarrow \mathcal{B}} \xrightarrow{A \rightarrow C} \quad (4.2)$$

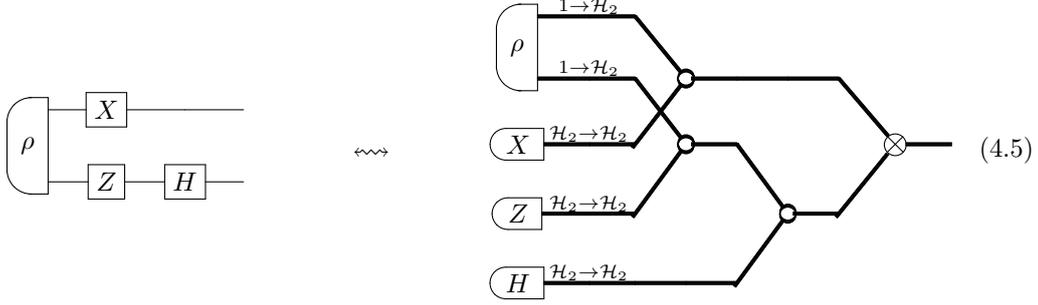
Parallel composition, in contrast, denotes operations applied on *different* systems. Naturally, there is no restriction on the types since the operations do not act on the same system. Suppose we have arbitrary operations  $\mathcal{A} : A \rightarrow A'$  and  $\mathcal{B} : B \rightarrow B'$  the types of which need not match in any way. We express graphically their application to different systems as:

$$\mathcal{A} \otimes \mathcal{B} \quad \longleftrightarrow \quad \begin{array}{c} \textcircled{\mathcal{A}} \xrightarrow{A \rightarrow A'} \\ \textcircled{\mathcal{B}} \xrightarrow{B \rightarrow B'} \end{array} \quad \begin{array}{c} \text{---} \\ \diagdown \\ \text{---} \\ \diagup \\ \text{---} \end{array} \quad \textcircled{\mathcal{A} \otimes \mathcal{B}} \xrightarrow{A \otimes B \rightarrow A' \otimes B'} \quad (4.3)$$

Finally, the swap of timelines is depicted as:

$$\begin{array}{c} \textcircled{\mathcal{A}} \xrightarrow{A \rightarrow A'} \\ \textcircled{\mathcal{B}} \xrightarrow{B \rightarrow B'} \end{array} \quad \begin{array}{c} \text{---} \\ \diagdown \\ \text{---} \\ \diagup \\ \text{---} \end{array} \quad \begin{array}{c} \text{---} \\ \diagdown \\ \text{---} \\ \diagup \\ \text{---} \end{array} \quad \begin{array}{c} \text{---} \\ \diagdown \\ \text{---} \\ \diagup \\ \text{---} \end{array} \quad (4.4)$$

Using the notation discussed above, we can express the preparation of a physical system to an initial state as a channel from the trivial type which we denote simply as “1”, where  $\mathcal{H}_1 = \mathbb{C}$ . We do not draw timelines with the trivial type  $1 \rightarrow 1$ . The graphical language expresses the rule that second-order operations have a fixed order by considering their application time from left to right. Example 4.1 shows how a simple circuit can be expressed as a supercircuit in our model.

**Example 4.1**


**Remark 4.1** *The circuit model of computation is widely used in literature and is the framework within which a lot of algorithms have been discovered. For that reason we define our graphical language in a way that resembles the circuit model. There does already exist another graphical framework for dealing with higher-order computation and indefinite causal order. It was first presented in Ref. [11] by Kissinger and Uijlen, and then developed further by them in Ref. [12] where it incorporates and visualises the process matrices formalism proposed in Ref. [6]. That framework, however, is based on string diagrams (see e.g. Coecke and Kissinger's book - Ref. [13]). Our motivation for the approach we have chosen lies in this difference between circuits and string diagrams.*

## 4.2 Characterisation of the transformations

We now consider what kind of maps can be described by the model defined above.

**Proposition 4.1** *Non-signalling supercircuits with definite causal order (Definition 4.1) are maps from non-signalling channels to non-signalling channels.*

**Proof.** In order to prove the result we prove that all allowed supermaps in our model map non-signalling channels to non-signalling channels. First, consider the parallel composition supermap that outputs the tensor product of two channels. By Equation 2.21, any product channel is also a non-signalling channel. Hence parallel composition preserves the non-signalling property. Next, consider the sequential composition supermap. We explicitly show it preserves the non-signalling property in the following equation:

$$\begin{array}{c}
 \begin{array}{|c|} \hline A \\ \hline \vdots \\ \hline I \\ \hline J \\ \hline K \\ \hline \vdots \\ \hline Z \\ \hline \end{array} \quad \begin{array}{|c|} \hline A' \\ \hline \vdots \\ \hline I' \\ \hline J' \\ \hline K' \\ \hline \vdots \\ \hline Z' \\ \hline \end{array} \\
 \hline \mathcal{C} \\
 \hline
 \end{array} \quad \begin{array}{|c|} \hline \alpha \\ \hline \vdots \\ \hline \lambda \\ \hline J' \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \quad \begin{array}{|c|} \hline \alpha' \\ \hline \vdots \\ \hline \lambda' \\ \hline \mu' \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \\
 \hline \mathcal{C}' \\
 \hline \text{Tr}
 \end{array} \stackrel{(2.20)}{=} \begin{array}{|c|} \hline A \\ \hline \vdots \\ \hline I \\ \hline J \\ \hline K \\ \hline \vdots \\ \hline Z \\ \hline \end{array} \quad \begin{array}{|c|} \hline A' \\ \hline \vdots \\ \hline I' \\ \hline J' \\ \hline K' \\ \hline \vdots \\ \hline Z' \\ \hline \end{array} \\
 \hline \mathcal{C} \\
 \hline \text{Tr} \\
 \hline \begin{array}{|c|} \hline \alpha \\ \hline \vdots \\ \hline \lambda \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \quad \begin{array}{|c|} \hline \alpha' \\ \hline \vdots \\ \hline \lambda' \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \\
 \hline \mathcal{C}' \\
 \hline
 \end{array} \stackrel{(2.20)}{=} \begin{array}{|c|} \hline A \\ \hline \vdots \\ \hline I \\ \hline K \\ \hline \vdots \\ \hline Z \\ \hline \end{array} \quad \begin{array}{|c|} \hline A' \\ \hline \vdots \\ \hline I' \\ \hline J' \\ \hline K' \\ \hline \vdots \\ \hline Z' \\ \hline \end{array} \\
 \hline \mathcal{C} \\
 \hline \begin{array}{|c|} \hline \alpha \\ \hline \vdots \\ \hline \lambda \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \quad \begin{array}{|c|} \hline \alpha' \\ \hline \vdots \\ \hline \lambda' \\ \hline \nu' \\ \hline \vdots \\ \hline \omega' \\ \hline \end{array} \\
 \hline \mathcal{C}' \\
 \hline
 \end{array} \tag{4.6}$$

Finally, the swap of timelines also trivially preserves the non-signalling property since it only reorders the channels rather than altering or connecting them. We have shown that all three allowed second-order operations preserve the non-signalling property of the channels. Therefore supercircuits are maps from non-signalling channels to non-signalling channels. ■

We have characterised the transformations that our model describes. Now we show that our model contains all first-order circuits as supercircuits with a single output timeline.

**Proposition 4.2** *Non-signalling supercircuits with definite causal order (Definition 4.1) with exactly one output timeline are in a 1-to-1 correspondence with first-order circuits.*

**Proof.** We first prove that each such supercircuit has a corresponding circuit. Consider a supercircuit with  $N$  input timelines and 1 output timeline. If  $N = 1$  this is just a single quantum channel which is a valid first-order circuit. Otherwise, we can deterministically construct an equivalent supercircuit with one less input timelines. Since there are more than 1 input timelines, and exactly 1 output timeline, the supercircuit has at least one composition. Take the first composition in the sequence of supermaps, described by the supercircuit. Construct an equivalent supercircuit by substituting the RHS of Equation 4.2 or 4.3 (depending on the composition type) with a single initial channel - the LHS - and hence single initial timeline. We have deterministically reduced the number of timelines by one. Iteratively repeat this operation until the base case  $N = 1$  which we have already shown corresponds to a valid circuit. Hence any supercircuit corresponds to a valid circuit.

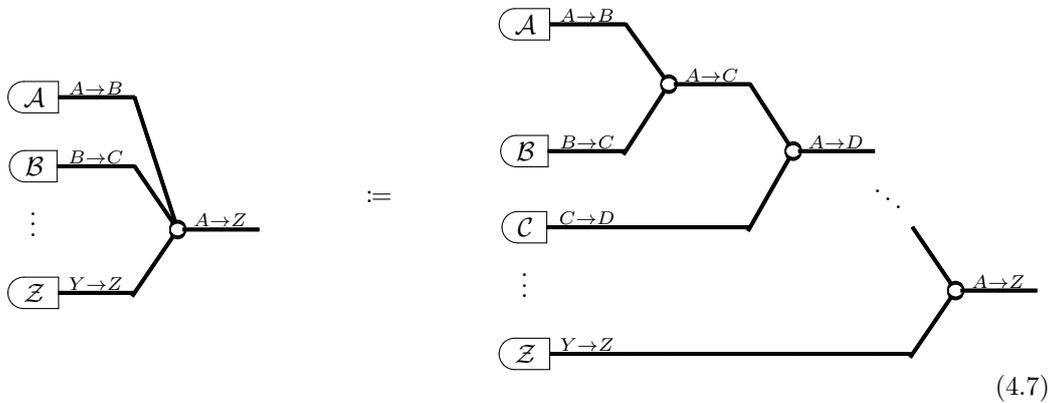
Now we show the opposite - that any circuit can be expressed as a supercircuit with a single output timeline. Any circuit can be seen as a single quantum channel. Quantum channels are represented as states in our model. Hence any circuit is a trivial supercircuit with a single input channel. ■

**Remark 4.2** *The fact that the supercircuits from the proposed basic model converge exactly to first-order circuits when only one timeline is output suggests that the model correctly pinpoints the most basic second-order computation model.*

### 4.3 Derived second-order operations

Now that we have proven the correctness of the basic model, let us consider what second-order operations can be derived from the axiomatised three in Definition 4.1.

Both  $\circ$  and  $\otimes$  are compositions on linear maps. As such, they are associative, in the sense that  $\mathcal{A} \circ (\mathcal{B} \circ \mathcal{C}) = (\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C}$  and  $\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C}) = (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}$ . Therefore we can overload the definitions from Equations 4.2, 4.3 to accept any number of inputs (sequential composition still requires all the types to match). More precisely, we extend the definitions of the sequential composition to:



and the parallel composition to:

(4.8)

Consider the associativity equation  $\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C}) = (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}$ . The same equation is shown using supercircuits in Equation 4.9. The dashed boxes show how this is equivalent to commutativity of the second-order compositions. This is analogously true for sequential composition too. Hence we reach the interesting observation that associativity (of linear maps) on the first level corresponds to commutativity (of compositions) on the second level.

(4.9)

**Remark 4.3** Equations 4.7 and 4.8 are reminiscent of the graphical rules used ZX-calculus, called spider diagrams [13].

Another operation we can construct is the second-order generalisation of the trace. The *discarding* operation for supercircuits erases a timeline, i.e. it acts as a “do not use timeline” command.

**Definition 4.2 (Discarding in supercircuits with definite causal order)**

Consider an arbitrary channel  $\mathcal{C} : A \rightarrow B$ . We define its discarding within a supercircuit

with definite causal order as:

$$\begin{array}{c} \text{C} \end{array} \xrightarrow{A \rightarrow B} \begin{array}{c} \text{---} | \cdot \end{array} := \begin{array}{c} \rho \xrightarrow{1 \rightarrow A} \\ \text{C} \xrightarrow{A \rightarrow B} \\ \text{Tr} \xrightarrow{B \rightarrow 1} \end{array} \quad (4.10)$$

where  $\rho$  is a fixed (but otherwise arbitrary) state of a system of dimension  $A$  and  $\text{Tr} : B \rightarrow 1$  traces out a system of dimension  $B$ . The composition's output timeline is not drawn because it has the trivial type  $1 \rightarrow 1$ .

**Proposition 4.3** *The discarding operation (Definition 4.2) is independent of the state  $\rho$  used. It can be fixed arbitrarily for any system dimension in supercircuits with definite causal order.*

**Proof.** Using the definition of discarding from Equation 4.10 on an arbitrary part of a multipartite non-signalling channel we get:

$$\begin{array}{c} \text{---} A \\ \vdots \\ \text{---} I \\ \vdots \\ \rho \xrightarrow{J} \\ \text{---} K \\ \vdots \\ \text{---} Z \end{array} \begin{array}{c} \text{C} \\ \vdots \end{array} \begin{array}{c} \text{---} A' \\ \vdots \\ \text{---} I' \\ \vdots \\ \text{---} K' \\ \vdots \\ \text{---} Z' \end{array} \xrightarrow{\text{Tr}} \begin{array}{c} \rho \xrightarrow{J} \\ \text{---} K \\ \vdots \\ \text{---} Z \end{array} \begin{array}{c} \text{Tr} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{array} \stackrel{(2.20)}{=} \begin{array}{c} \rho \xrightarrow{J} \\ \text{---} K \\ \vdots \\ \text{---} Z \end{array} \begin{array}{c} \text{Tr} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{array} = \begin{array}{c} \text{---} A \\ \vdots \\ \text{---} I \\ \vdots \\ \text{---} K \\ \vdots \\ \text{---} Z \end{array} \begin{array}{c} \text{C} \\ \vdots \end{array} \begin{array}{c} \text{---} A' \\ \vdots \\ \text{---} I' \\ \vdots \\ \text{---} K' \\ \vdots \\ \text{---} Z' \end{array}$$

Hence when discarding is applied directly to an input channel, the state  $\rho$  gets traced out and its particular value does not matter. This, together with Proposition 4.1, which asserts that supercircuits with definite causal order preserve the non-signalling property, implies that the particular state  $\rho$  does not matter for discarding operations anywhere in such a supercircuit.  $\blacksquare$

In summary, we have defined a simple model of second-order supercircuits. We can think of it as being able to (deterministically) reorder in any way a number of black boxes placed on a table, which represent all the channels to be applied in some first-order circuit. We then wire the boxes either to one another, or externally by leaving unconnected wires as global inputs/outputs. This by itself does not provide us more expressive power than normal circuits. However, it is the first step towards defining a less constrained model which does.

## Chapter 5

# Non-signalling supercircuits with indefinite causal order

In this chapter we extend the basic model of second-order quantum computation to a model that includes the quantum SWITCH and is therefore more expressive than first-order circuits.

### 5.1 Computational model

We now add one more allowed second-order operation - the quantum SWITCH. As discussed in section 3.2, it allows for *superposition of time orders*. We can think of the SWITCH as a controlled swap of timelines, i.e. a *controlled superswap*.

There are two constraints that we impose at this stage. The first one is that we allow the SWITCH only for timelines with matching both input and output types. Without this constraint, the timelines coming out of a SWITCH would be in a *superposition of types*. We discuss that case later in the dissertation. For now we assume that the SWITCH is only defined on timelines with matching both input and output types. Thanks to this assumption, the resulting timelines' types remain the same regardless of the control value.

The second constraint is that the control qubits for all SWITCHes are separate from the ones that the channels act on. We can think of this model as a dynamically programmable quantum computer with indefinite causal order (described by the supercircuit) which is controlled by another quantum computer with definite causal order (which sets the controls of the SWITCHes to desired values).

#### **Definition 5.1 (Non-signalling supercircuits with indefinite causal order)**

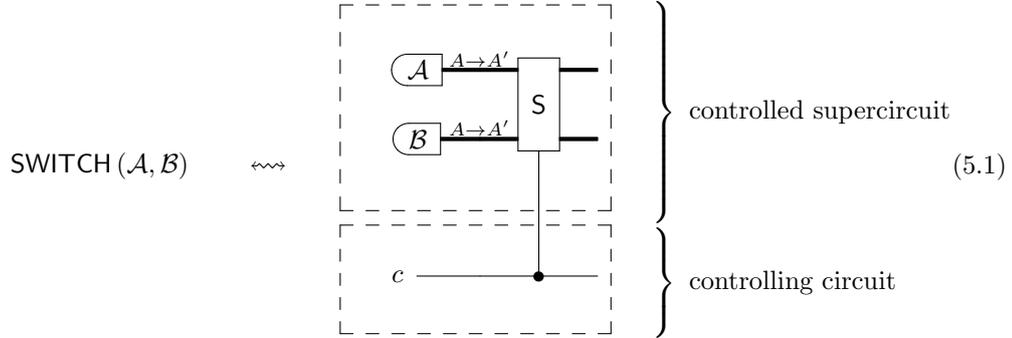
*A non-signalling supercircuit with indefinite causal order describes how a sequence of quantum supermaps acts on an  $N$ -partite non-signalling channel. In order to process such input, we allow the following four supermaps:*

- Sequential composition - *the sequential composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{B} \circ \mathcal{A}$ .*
- Parallel composition - *the parallel composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{A} \otimes \mathcal{B}$ .*
- Swap of timelines - *deterministic reordering in time of channels*
- SWITCH of two timelines with matching types - *swap of timelines controlled by an external qubit.*

This model allows indefinite causal order between first-order operations, i.e. the channels, but only a fixed causal order between supermaps. There are also no allowed operations

between supermaps. Therefore the model itself is one of second-order computation with indefinite causal order.

Graphically, let the second-order SWITCH operation on two channels, controlled by qubit  $c$ , be denoted as:

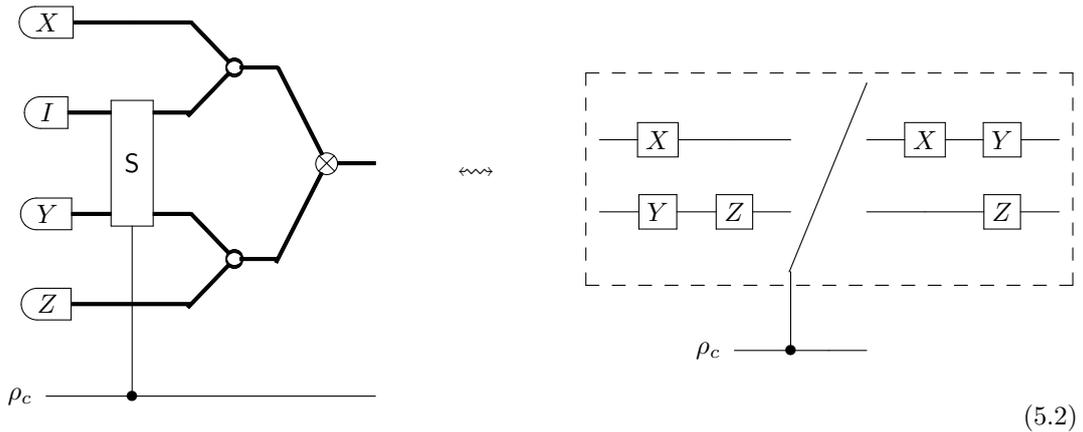


where the controlling circuit could be any normal circuit, as long as the channels of the supercircuit  $\mathcal{A}, \mathcal{B}, \dots$  do not act on any of its qubits. Note that the SWITCH box has two output timelines instead of just one. This allows for the possibility to insert other operations in between  $\mathcal{A}$  and  $\mathcal{B}$ , or apply further SWITCHes to them. As for the other second-order operations, SWITCHes are time-ordered from left to right.

Continuing the analogy of shuffling black boxes on a table, the model described in this section extends the previous one so that we can not only reorder and connect them arbitrarily but also have superpositions of those orders and connectivities.

In supercircuits with indefinite causal order it is possible to change which physical system a channel acts on only by varying a SWITCH's control value. We refer to this flexibility of connecting the operators as *dynamic connectivity*. Example 5.1 below illustrates how it works.

**Example 5.1** *This example shows how dynamic connectivity can be achieved using the SWITCH. Consider the supercircuit below:*



where  $I, X, Y, Z$  are the Pauli single-qubit gates, and types have been left implicit. The notation in the RHS indicates that this is a superposition of two circuits, controlled by a qubit. Depending on the value of the control qubit, the  $Y$  gate can be applied to the first or to the second qubit. We see the supercircuit on the LHS as a more convenient representation, available through our graphical language.

Another example application of the SWITCH has been proposed by Ebler, *et al.* in Ref. [10] as a means to communicate via two completely depolarising channels. In order to

show how it can be represented in our framework, we first present a short summary of the original protocol. We are given a qubit in an unknown state  $\rho$ , and two uses of a completely depolarising channel, which is defined to map any arbitrary state  $\sigma$  into the maximally mixed state:

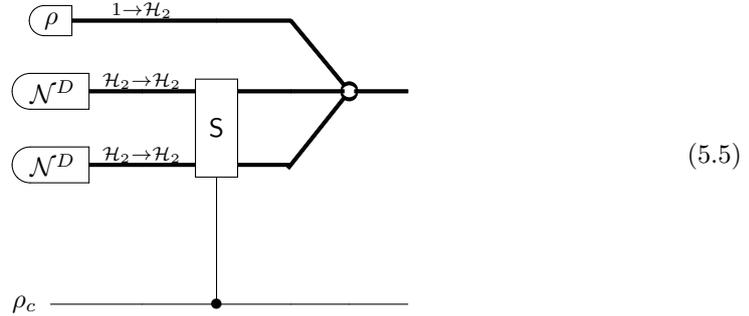
$$\mathcal{N}^D(\sigma) = \frac{I}{\dim(\sigma)} \quad (5.3)$$

The two channels are combined through the quantum SWITCH with control qubit in state  $\rho_c = |+\rangle\langle +|$ , and the result is applied to the system  $\rho$ . Non-zero communication is achieved since the output state depends on the initial value of  $\rho$ . In particular, the receiving party measures the control qubit in the Fourier basis  $\{|\pm\rangle\}$  and obtains

$$\langle \pm |_c \mathcal{S}(\rho_c, \mathcal{N}^D, \mathcal{N}^D)(\rho_c \otimes \rho) |\pm\rangle_c = \frac{I}{2d} \pm \frac{\rho}{d^2} \quad (5.4)$$

where  $d = \dim(\rho)$ . This result seems rather counter-intuitive because combining two completely depolarising channels in any fixed causal order yields simply  $\mathcal{N}^D \circ \mathcal{N}^D = \mathcal{N}^D$ , which does not allow for communication. Therefore it is a useful illustration of an advantage provided by the quantum SWITCH. Example 5.2 below shows how the protocol can be concisely expressed using our graphical language.

**Example 5.2** Consider Ebler’s protocol from Ref. [10]. As summarised above, it describes how communication is made possible by SWITCHing two completely depolarising channels  $\mathcal{N}^D$  and applying the result to a state  $\rho$ . Using our graphical language, we can express in one diagram the procedure, described verbally in [10]:



## 5.2 Characterisation of the transformations

Analogously to Proposition 4.1, we wish to pinpoint a type of channels which the currently considered supercircuits leave invariant. This leads us to the following definition:

### Definition 5.2 (Extended non-signalling channels)

A channel  $\mathcal{C} : ABE \rightarrow A'B'E'$  is called extended non-signalling if for every  $\sigma$  the channel  $\mathcal{N}$ , defined by  $\mathcal{N}(\rho) = \text{Tr}_{E'} \mathcal{C}(\rho \otimes \sigma)$  is non-signalling.

Having this new term at our disposal, we can characterise the current supercircuits as follows.

**Proposition 5.1** Non-signalling supercircuits with indefinite causal order (Definition 5.1) are maps from extended non-signalling channels to extended non-signalling channels.

**Proof.** Consider an arbitrary supercircuit of this type. If there are no SWITCHes in it, it is equivalent to a supercircuit with definite causal order. By Proposition 4.1 then, it preserves the non-signalling and hence the extended non-signalling property.

Otherwise, there is at least one SWITCH in the supercircuit. A SWITCH with control in fixed states  $|0\rangle$  or  $|1\rangle$  corresponds to second-order identity, or the deterministic swap on two

timelines, respectively. Tracing out its control gives us a mixture of two supercircuits. One where the SWITCH is substituted for the identity and one - for the swap. Tracing out all the controls of all, say  $N$ , SWITCHes in this manner, we end up in a mixture of  $2^N$  supercircuits. None of them have any remaining SWITCHes and therefore they are supercircuits with definite causal order. By Proposition 4.1 they are maps from non-signalling channels to non-signalling channels. Hence we have shown that tracing out all SWITCH controls from a supercircuit we obtain a mixture of maps from non-signalling channels to non-signalling channels. Therefore supercircuits with indefinite causal order preserve the extended non-signalling property. ■

**Proposition 5.2** *Non-signalling supercircuits with indefinite causal order (Definition 5.1) with exactly one output timeline correspond to superpositions of first-order circuits. More precisely, tracing out the external control system yields a mixture of first-order circuits.*

**Proof.** We have already seen in the proof of Proposition 5.1 that tracing out the external control system of a supercircuit with indefinite causal order, we obtain a mixture of supercircuits with definite causal order. In the special case that all supercircuits of the mixture have exactly one output timeline it follows directly from Proposition 4.2 that this is also a mixture of first-order circuits. ■

### 5.3 Derived second-order operations

We wish to extend Definition 4.2 of the discarding operation for supercircuits with indefinite causal order. The same definition of preparing a state, applying the channel to it, and then tracing it out, is still applicable. However, as shown below, the value of the state does matter in this case.

**Definition 5.3 (Discarding in supercircuits with indefinite causal order)**

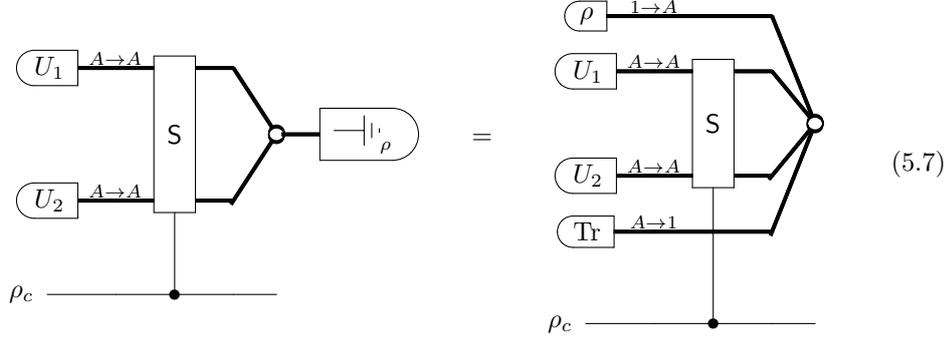
Consider a channel  $C : A \rightarrow B$ . We define its discarding within a supercircuit with indefinite causal order as:

$$\begin{array}{c} \textcircled{C} \end{array} \xrightarrow{A \rightarrow B} \begin{array}{c} \text{---} | \rho \end{array} \quad := \quad \begin{array}{c} \textcircled{\rho} \xrightarrow{1 \rightarrow A} \\ \textcircled{C} \xrightarrow{A \rightarrow B} \\ \textcircled{\text{Tr}} \xrightarrow{B \rightarrow 1} \end{array} \quad , \quad (5.6)$$

where  $\rho$  is a state of a system of dimension  $A$  and  $\text{Tr} : B \rightarrow 1$  traces out a system of dimension  $B$ .

**Proposition 5.3** *The behaviour of the discarding operation (Definition 5.3) in supercircuits with indefinite causal order depends on the value of the state  $\rho$ .*

**Proof.** Consider the following example supercircuit on two unitaries,  $U_1$  and  $U_2$  of type  $A \rightarrow A$ .

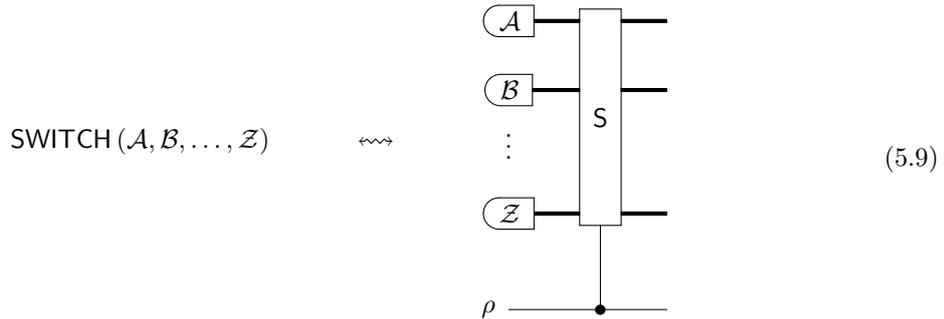


Let  $Q$  be the physical system of dimension  $A$  that is initially prepared in state  $\rho$  and is then acted on by the unitaries. Tracing out  $Q$  we get the following overall state:

$$\begin{aligned}
 \text{Tr}_Q [\mathbf{S}(U_1, U_2)(\rho \otimes \rho_c)] &= \\
 &= \text{Tr}_Q \left[ \left( U_1 U_2 \otimes |0\rangle \langle 0| + U_2 U_1 \otimes |1\rangle \langle 1| \right) (\rho \otimes \rho_c) \left( U_2^\dagger U_1^\dagger \otimes |0\rangle \langle 0| + U_1^\dagger U_2^\dagger \otimes |1\rangle \langle 1| \right) \right] \\
 &= \text{Tr} \left[ U_1 U_2 \rho U_2^\dagger U_1^\dagger \otimes |0\rangle \langle 0| \rho_c |0\rangle \langle 0| + \text{Tr} \left[ U_1 U_2 \rho U_1^\dagger U_2^\dagger \otimes |0\rangle \langle 0| \rho_c |1\rangle \langle 1| \right. \right. \\
 &\quad \left. \left. + \text{Tr} \left[ U_2 U_1 \rho U_2^\dagger U_1^\dagger \otimes |1\rangle \langle 1| \rho_c |0\rangle \langle 0| + \text{Tr} \left[ U_2 U_1 \rho U_1^\dagger U_2^\dagger \otimes |1\rangle \langle 1| \rho_c |1\rangle \langle 1| \right] \right] \right] \\
 &= \text{Tr}[\rho \otimes |0\rangle \langle 0| \rho_c |0\rangle \langle 0| + \text{Tr} \left[ U_1 U_2 \rho U_1^\dagger U_2^\dagger \otimes |0\rangle \langle 0| \rho_c |1\rangle \langle 1| \right. \\
 &\quad \left. + \text{Tr} \left[ U_2 U_1 \rho U_2^\dagger U_1^\dagger \otimes |1\rangle \langle 1| \rho_c |0\rangle \langle 0| + \text{Tr}[\rho \otimes |1\rangle \langle 1| \rho_c |1\rangle \langle 1| \right] \right] \quad (5.8)
 \end{aligned}$$

This state depends on the particular value of  $\rho$ . Hence in supercircuits with indefinite causal order we need to specify the state  $\rho$  used for any particular discarding.  $\blacksquare$

We defined the  $N$ -SWITCH (Definition 3.3) as a supermap that can generate any particular permutation of  $N$  channels by only varying the value of its control system. We denote graphically an  $N$ -SWITCH with control in state  $\rho$  as:



The  $N$ -SWITCH can be decomposed into  $O(N \log_2 N)$  2-SWITCHes [28, 29]. Therefore we think of the  $N$ -SWITCH for any  $N \geq 2$  as a derived operation in the current model of computation. If the complexity of a supercircuit is discussed, the decomposition of all  $N$ -SWITCHes into 2-SWITCHes should be considered, instead of assuming them as a given resource.

## Chapter 6

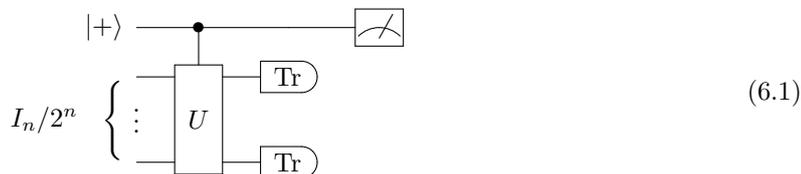
# A novel second-order quantum algorithm

In this chapter we develop an application of the model we have developed earlier in the thesis. We build a novel quantum algorithm that estimates certain properties of a set of quantum channels. We conjecture that our algorithm has an exponential advantage over algorithms with definite causal order, and we provide arguments in favour of this conjecture. That would be the first concrete example of an *exponential* advantage achieved by utilising indefinite causal order. Finding a formal proof for this advantage, however, is still an open problem.

In constructing our algorithm we have taken inspiration from the DQC1 algorithm [14]. There is a similarity between them in the following sense. Both algorithms take quantum black boxes as input, and output an estimate for some quantity of these boxes, obtained by sampling measurement probabilities. In order to describe our algorithm, we begin by considering DQC1 as a simpler precursor.

### 6.1 The DQC1 algorithm

Here we introduce the standard first-order version of the DQC1 protocol (Ref. [14]). For conciseness of notation, we denote the type of an  $n$ -qubit system as  $\mathcal{H}^n := \mathcal{H}_2^{\otimes n}$ . Suppose we are given access to a controlled operation  $\Lambda U$  which implements the black box unitary  $U : \mathcal{H}^n \rightarrow \mathcal{H}^n$  if the control is in state  $|1\rangle$ . The algorithm evaluates the real part of the trace of  $U$  using the following circuit:



The control system  $C$ , which is a single qubit, is initially prepared in state  $\rho_C = |+\rangle\langle+|$ . The application system  $A$  is an  $n$ -qubit system initially prepared in the maximally mixed state  $\rho = I_n/2^n$ . The protocol uses  $C$  as the control for applying  $U$  to the application qubits, after which system  $A$  is traced out. The state of the total system after tracing  $A$  is:

$$\begin{aligned}
 \rho' &= \text{Tr}_A \left[ (|0\rangle\langle 0| \otimes I_n + |1\rangle\langle 1| \otimes U) (\rho_C \otimes \rho) (|0\rangle\langle 0| \otimes I_n + |1\rangle\langle 1| \otimes U^\dagger) \right] \\
 &= \text{Tr}_A \left[ |0\rangle\langle 0| \rho_C |0\rangle\langle 0| \otimes \rho + |0\rangle\langle 0| \rho_C |1\rangle\langle 1| \otimes \rho U^\dagger \right. \\
 &\quad \left. + |1\rangle\langle 1| \rho_C |0\rangle\langle 0| \otimes U \rho + |1\rangle\langle 1| \rho_C |1\rangle\langle 1| \otimes U \rho U^\dagger \right] \\
 &= |0\rangle\langle 0| \rho_C |0\rangle\langle 0| \otimes \rho + |0\rangle\langle 0| \rho_C |1\rangle\langle 1| \otimes \frac{\text{Tr}[U^\dagger]}{2^n} \rho + |1\rangle\langle 1| \rho_C |0\rangle\langle 0| \otimes \frac{\text{Tr}[U]}{2^n} \rho \\
 &\quad + |1\rangle\langle 1| \rho_C |1\rangle\langle 1| \otimes \rho
 \end{aligned} \tag{6.2}$$

Measuring system  $C$  in the  $\{| \pm \rangle\}$  basis then yields the following probabilities

$$\begin{aligned}
 \mathbb{P}(\pm) &= \frac{1 \pm (\text{Tr}[U] + \text{Tr}[U^\dagger]) / 2^{n+1}}{2} \\
 &= \frac{1 \pm \text{Re}(\text{Tr}[U]) / 2^n}{2}
 \end{aligned} \tag{6.3}$$

By repeating the procedure a number of times, an estimate of the measurement probabilities for the two outcomes is obtained. Finally, we use that estimate to evaluate the real part of  $U$ 's trace.

$$\text{Re}(\text{Tr}[U]) = 2^n (2\mathbb{P}(+) - 1) \tag{6.4}$$

**Remark 6.1** *Estimating the  $\mathbb{P}(\pm)$  probabilities is equivalent to estimating the bias of an unfair coin through sampling. It is a standard result in probability theory that in order to have confidence  $\delta$  that we have estimated the bias up to additive error  $\varepsilon$ , we need  $\Theta(\varepsilon^{-2} \log(1/\delta))$  samples [31]. In our case each sample is obtained from one execution of the circuit in Equation 6.1. From Equation 6.3 we see that in order to get a better estimate than simply  $1/2$  we need  $\varepsilon < \text{Re}(\text{Tr}[U]) / 2^n$ .*

*If  $\text{Re}(\text{Tr}[U]) = \Theta(2^n)$ , i.e. it is exponentially large in  $n$ , then the overall number of samples needed is sub-exponential. Consider, however, if it is polynomial, i.e.  $\text{Re}(\text{Tr}[U]) = O(n^\alpha)$  for some  $\alpha = O(1)$ . Then  $\varepsilon = O(2^{-n})$  and an exponentially large (in  $n$ ) number of samples is needed. Hence we remark that the protocol only yields an asymptotic advantage when  $\text{Re}(\text{Tr}[U]) = \Theta(2^n)$ .*

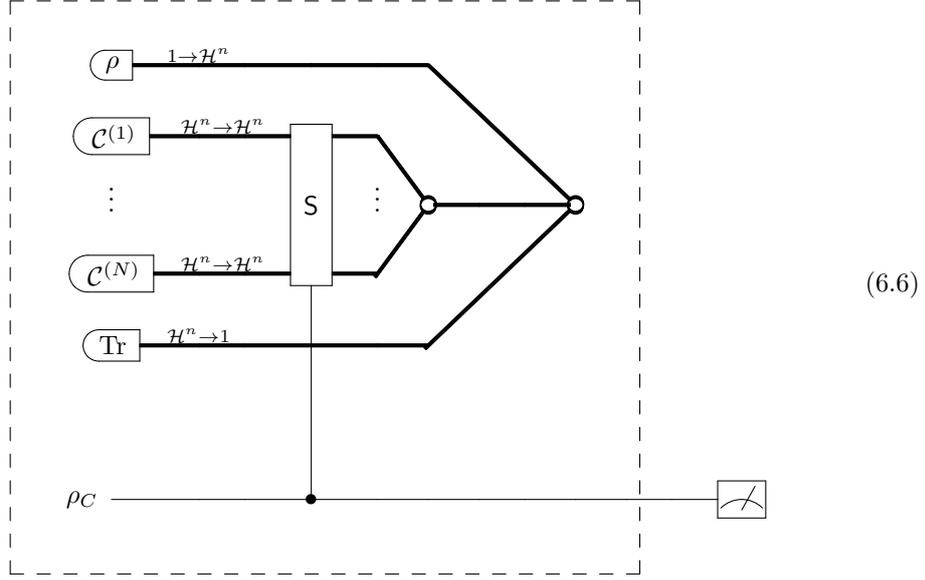
## 6.2 $N$ -channel DQC1 with indefinite causal order

Now we present our algorithm which estimates a property of a set of channels. Consider  $N$  noisy channels  $\mathcal{C}^{(1)}, \dots, \mathcal{C}^{(N)} : \mathcal{H}^n \rightarrow \mathcal{H}^n$  with Kraus decompositions:

$$\mathcal{C}^{(i)}(\rho) = \sum_{j=1}^{d_i} K_j^{(i)} \rho (K_j^{(i)})^\dagger \tag{6.5}$$

where  $d_i$  is the number of Kraus operators of the  $i^{\text{th}}$  channel. Let  $C$  be an  $N!$ -dimensional quantum system with orthonormal basis  $\{|\pi\rangle \mid \pi \in \text{Perm}(N)\}$ . The supercircuit below

(Equation 6.6) describes our  $N$ -channel DQC1 algorithm.



The application and control quantum systems are initially prepared in states

$$\rho = I_n/2^n \quad (6.7)$$

$$\rho_C = \frac{1}{N!} \sum_{\pi, \sigma \in \text{Perm}(N)} |\pi\rangle \langle \sigma| \quad (6.8)$$

respectively. We define the measurement on the control system  $C$  to be a projective measurement described by the set of two measurement operators  $\{P, I_n - P\}$ .  $P$  is an  $n$ -qubit projector for which  $P = P^\dagger = P^2$ , and  $I_n - P$  is its orthogonal projector. This is a valid measurement since it satisfies the completeness relation of Equation 2.3:

$$P^\dagger P + (I_n - P)^\dagger (I_n - P) = P^2 + I_n - P - P + P^2 = I_n \quad (6.9)$$

The point of having only two measurement operators is that then only a single probability needs to be estimated through sampling. Estimating the probability distribution for more than two measurement outcomes would require more samples, and therefore more copies of the input channels. That is why we only want to have a single two-outcome measurement.

The projector  $P$  can be chosen in various ways, each of which yields estimates for different expressions involving the channels' Kraus operators. Therefore we treat it as a degree of freedom of the protocol to choose a suitable projector  $P$ . For the current purpose of describing how the algorithm works, for now we assume that it is chosen to be

$$P = \frac{1}{N!} \sum_{\pi, \sigma \in \text{Perm}(N)} |\pi\rangle \langle \sigma| \quad (6.10)$$

It is easily seen that  $P^\dagger = P$  and  $P^2 = P$ .

The connection between our algorithm and DQC1 can be seen by comparing their respective supercircuit (Equation 6.6) and circuit (Equation 6.1). Both algorithms apply their input channels to a system, prepared in the maximally mixed state, then trace it out, and measure some control system. Our algorithm, however, has a broader scope in that it acts on  $N$  channels, rather than on a single unitary, and utilises the  $N$ -SWITCH supermap to exploit indefinite causal order. In our algorithm the system  $C$  controls the *order of operations*,

rather than *whether* an operation is applied.

Now we consider what the measurement probabilities are, and hence what quantity can be evaluated through estimating them. Using Equation 2.8, we obtain the following.

$$\begin{aligned}
 \mathbb{P}(P) &= \text{Tr} \left[ \frac{1}{N!^2} \sum_{\substack{\alpha, \beta, \pi, \sigma \in \\ \text{Perm}(N)}} |\alpha\rangle \langle \beta| \pi \langle \sigma| \text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \frac{I_n}{2^n} \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \right] \\
 &= \frac{1}{N!^2} \sum_{\substack{\alpha, \beta, \pi, \sigma \in \\ \text{Perm}(N)}} \delta_{\alpha\sigma} \delta_{\beta\pi} \text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \frac{I_n}{2^n} \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \\
 &= \frac{1}{N!^2} \sum_{\pi, \sigma \in \text{Perm}(N)} \text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \frac{I_n}{2^n} \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \quad (6.11)
 \end{aligned}$$

In order to simplify this expression, consider the case when  $\pi = \sigma$ . Let  $\pi^{-1}$  denote the inverse permutation of  $\pi$ , i.e.  $\pi$  read backwards. Then we get the following.

$$\begin{aligned}
 &\text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \frac{I_n}{2^n} \left( \prod_{k=1}^N K_{i_{\pi(k)}}^{(\pi(k))} \right)^\dagger \right] \\
 &= \frac{1}{2^n} \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \text{Tr} \left[ \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \prod_{k=1}^N \left( K_{i_{\pi^{-1}(k)}}^{(\pi^{-1}(k))} \right)^\dagger \right] \\
 &= \frac{1}{2^n} \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \text{Tr} \left[ \prod_{k=1}^N \left( K_{i_{\pi^{-1}(k)}}^{(\pi^{-1}(k))} \right)^\dagger \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right] \\
 &\stackrel{(?)}{=} \frac{1}{2^n} \text{Tr} [I_n] \\
 &= 1 \quad (6.12)
 \end{aligned}$$

where between the second and third lines we have used the trace's cyclic property. Plugging Equation 6.12 into Equation 6.11, we obtain the final expression for the measurement probability.

$$\mathbb{P}(P) = \frac{1}{N!} + \frac{1}{2^n N!^2} \sum_{\substack{\pi, \sigma \in \text{Perm}(N) \\ \pi \neq \sigma}} \text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \quad (6.13)$$

As in the case of the standard DQC1 protocol, by estimating two complementary probabilities,  $\mathbb{P}(P)$  and  $\mathbb{P}(I_n - P) = 1 - \mathbb{P}(P)$ , we evaluate some quantity,  $\tau$ , which depends on the input channels. In the original algorithm  $\tau = \text{Re}(\text{Tr}[U])$ , while in our extended protocol  $\tau$  is a more complicated and abstract object which quantifies the sum of traces of various compositions of the input channels' Kraus operators. Note that the result is independent of the particular Kraus decomposition of the channels.

**Example 6.1** In order to put our algorithm in less abstract terms, consider an example with  $N = 2$  input channels, where the  $N$ -SWITCH is just a 2-SWITCH. Using Equation 6.13 we get the measurement probability for  $P = |+\rangle\langle +|$  to be

$$\mathbb{P}(P) = \frac{1}{2} + \frac{1}{2^{2+n}} \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \left( \text{Tr} \left[ K_{i_1}^{(1)} K_{i_2}^{(2)} K_{i_1}^{(1)\dagger} K_{i_2}^{(2)\dagger} \right] + \text{Tr} \left[ K_{i_2}^{(2)} K_{i_1}^{(1)} K_{i_2}^{(2)\dagger} K_{i_1}^{(1)\dagger} \right] \right) \quad (6.14)$$

Hence the quantity we evaluate as final output is

$$\tau = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \left( \text{Tr} \left[ K_{i_1}^{(1)} K_{i_2}^{(2)} K_{i_1}^{(1)\dagger} K_{i_2}^{(2)\dagger} \right] + \text{Tr} \left[ K_{i_2}^{(2)} K_{i_1}^{(1)} K_{i_2}^{(2)\dagger} K_{i_1}^{(1)\dagger} \right] \right) \quad (6.15)$$

### 6.3 $N$ -channel DQC1 using 2-SWITCHes

A more practical perspective on the  $N$ -SWITCH is if we consider its decomposition into a number of 2-SWITCHes. That way, if we know how to construct the 2-SWITCH in an experiment, we can also construct the  $N$ -SWITCH. Here we develop our algorithm's description to consider the  $N$ -SWITCH as a sequence of 2-SWITCHes.

As discussed in section 3.2, the  $N$ -SWITCH's decomposition into 2-SWITCHes is not unique. Therefore, depending on its particular implementation, its control system may be of different sizes. In order to keep the description general, the only knowledge we assume of the control system  $C$  is that it is a set of  $|C|$  qubits. Similar to before, let all control qubits to be in the  $|+\rangle$  state, i.e.

$$\rho_C = |+\rangle^{|C|} \langle +|^{|C|} \quad (6.16)$$

(The  $\otimes$  symbol in the superscripts is omitted for conciseness of notation.) Also define the measurement projector  $P$  to be

$$P = |+\rangle^{|C|} \langle +|^{|C|} \quad (6.17)$$

Using Equation 2.8, we obtain the following measurement probability.

$$\mathbb{P}(P) = \text{Tr} \left[ \sum_{\vec{x}, \vec{y} \in \{0,1\}^{|C|}} \bigotimes_{i=1}^{|C|} |+\rangle \langle +|_{x_i} \langle x_i|_+ \langle +|_{y_i} \langle y_i| \cdot \text{Tr} \left[ \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \frac{I_n}{2^n} \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \right] \quad (6.18)$$

Let us inspect Equation 6.18 carefully. First, the overall sums over  $\vec{x}, \vec{y}$  enumerate all possible pairs (with repetition) of bitstrings of length  $|C|$ . Each bit of these bitstrings corresponds to a 2-SWITCH control being in state  $|0\rangle$  or  $|1\rangle$ , i.e. not swapping or swapping its two input timelines. The overall  $N$ -SWITCH is composed of  $|C|$  2-SWITCHes and therefore  $|\vec{x}| = |\vec{y}| = |C|$ , i.e. one bit per controlling qubit. The middle part of the expression corresponds to the measurement projector  $P$  acting on the control qubits.

Finally, the inner trace expression corresponds to tracing out the  $n$  application qubits after applying to them the SWITCHed channels. We define  $\pi$  to be the resulting permutation of the  $N$  input channels, if the  $N$ -SWITCH's control qubits had 0/1 values equal to the

bitstring  $\vec{x}$ . Analogously,  $\sigma$  is the resulting permutation if the control values were equal to  $\vec{y}$ . Note that then  $\pi$  is a function (only) of  $\vec{x}$ , and  $\sigma$  - of  $\vec{y}$ . Using this definition, we see that  $\pi$  and  $\sigma$  indeed denote a single permutation, rather than a superposition of permutations, because SWITCHes with controls in 0/1 states are second-order operations with definite causal order - the identity and deterministic swap on timelines.

Having explained all the different bits of Equation 6.18, we now simplify it. First, consider that

$$\forall x_i, y_i \in \{0, 1\} : \text{Tr} [|+\rangle \langle +|_{x_i} \langle x_i| + \rangle \langle +|_{y_i} \langle y_i|] = \left(\frac{1}{\sqrt{2}}\right)^4 = \frac{1}{4} \quad (6.19)$$

As before, all the difference in the final state's coefficients comes from the traced out application qubits. Putting together Equations 6.12 and 6.19 into Equation 6.18, we get the following simplified expression for the probability.

$$\mathbb{P}(P) = \frac{1}{2^{2|C|}} \left( \sum_{\substack{\vec{x}, \vec{y} \in \{0,1\}^{|C|} \\ \text{s.t. } \pi = \sigma}} 1 + \sum_{\substack{\vec{x}, \vec{y} \in \{0,1\}^{|C|} \\ \text{s.t. } \pi \neq \sigma}} \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \frac{1}{2^n} \text{Tr} \left[ \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \right) \quad (6.20)$$

The relation between the control bitstrings  $\vec{x}, \vec{y}$  and the resulting permutations  $\pi, \sigma$  obviously depends on the particular implementation of the  $N$ -SWITCH as a number of 2-SWITCHes. The term

$$\sum_{\substack{\vec{x}, \vec{y} \in \{0,1\}^{|C|} \\ \text{s.t. } \pi = \sigma}} 1 := \chi \quad (6.21)$$

is then a property of the  $N$ -SWITCH's deconstruction into 2-SWITCHes, and does not depend on the input channels  $\{C^{(i)}\}$ . In order to simplify the final expression, then, we let  $\chi$  denote the number of pairs (with repetition) of bitstrings  $\vec{x}, \vec{y}$  which result in equal permutations  $\pi = \sigma$ . Again,  $\chi$  itself does not depend on the input channels but is instead a property of the supercircuit. This gives us the final expression below.

$$\mathbb{P}(P) = \frac{1}{2^{2|C|}} \left( \chi + \sum_{\substack{\vec{x}, \vec{y} \in \{0,1\}^{|C|} \\ \text{s.t. } \pi \neq \sigma}} \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \frac{1}{2^n} \text{Tr} \left[ \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \right) \quad (6.22)$$

Note that the second term in the parentheses - call it  $\tau$  - depends *both* on the input channels and on the  $N$ -SWITCH's implementation.

## 6.4 Measurement modifications to the protocol

As we mention in the initial description of the  $N$ -channel DQC1 protocol, there are many choices for the measurement projector  $P$ . Depending on the final quantity that one wants to evaluate, different values for  $P$  can be chosen. This is a degree of freedom of the proposed

protocol - different quantities of the input channels can be estimating by simply varying the measurement projector, but otherwise leaving the supercircuit the same. We illustrate the point by considering another special case, namely

$$P = |- \rangle^{|C|} \langle -|^{|C|} \quad (6.23)$$

without going into excessive technical detail (because the derivation is analogous). The difference to our previous analysis is that Equation 6.19 becomes

$$\forall x_i, y_i \in \{0, 1\} : \text{Tr} [ |- \rangle \langle -| x_i \rangle \langle x_i| + \rangle \langle +| y_i \rangle \langle y_i| ] = \frac{1}{4} (-1)^{x_i + y_i} \quad (6.24)$$

Then the final expression is

$$\begin{aligned} \mathbb{P}(P) = \frac{1}{2^{2|C|}} & \left( \sum_{\substack{\vec{x}, \vec{y} \in \{0, 1\}^{|C|} \\ \text{s.t. } \pi = \sigma}} (-1)^{s(\vec{x} + \vec{y})} + \right. \\ & \left. + \sum_{\substack{\vec{x}, \vec{y} \in \{0, 1\}^{|C|} \\ \text{s.t. } \pi \neq \sigma}} \sum_{i_1=1}^{d_1} \cdots \sum_{i_N=1}^{d_N} \frac{(-1)^{s(\vec{x} + \vec{y})}}{2^n} \text{Tr} \left[ \left( \prod_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \right) \left( \prod_{k=1}^N K_{i_{\sigma(k)}}^{(\sigma(k))} \right)^\dagger \right] \right) \end{aligned} \quad (6.25)$$

where  $s$  is a parity function, i.e.  $s(\vec{a})$  is 0 if the sum of  $a$ 's elements is even, and 1 otherwise. The final expression looks similar to the one we obtained in Equation 6.20. However, the quantity being evaluated is different. What was an equally weighted sum of traces there is now a linear combination over the same traces with coefficients  $\pm 1$ . This illustrates the point how varying the measurement projector yields different quantities.  $P$  can be an arbitrarily complex operator, rather than the two highly symmetric cases we have shown - all pluses and all minuses. That way the same algorithm can act as an estimator for various quantities that depend on the input channels.

## 6.5 Arguments for the advantage of our algorithm

Having described our proposed  $N$ -channel DQC1 protocol and its variations, now we turn our attention to why we expect it provides computational advantage over algorithms with definite causal order.

The first argument we present is highly informal but is useful as an illustration of the main point. We are given as input access to the channels  $\{\mathcal{C}^{(i)}\}$  but no information about them. The only way we can use this input with fixed causal order is by applying copies of these channels. Their application to some state  $\rho$  is governed by the following equation.

$$\mathcal{C}^{(i)}(\rho) = \sum_{j=1}^{d_i} K_j^{(i)} \rho \left( K_j^{(i)} \right)^\dagger \quad (6.26)$$

The resulting transformation we get by repeatedly applying Equation 6.26 has the symmetry of  $K_j^{(i)}$ s and  $\left( K_j^{(i)} \right)^\dagger$ s around  $\rho$ .

Now consider the final answer we obtained in Example 6.1 (copied here for convenience).

$$\tau = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \left( \text{Tr} \left[ K_{i_1}^{(1)} K_{i_2}^{(2)} K_{i_1}^{(1)\dagger} K_{i_2}^{(2)\dagger} \right] + \text{Tr} \left[ K_{i_2}^{(2)} K_{i_1}^{(1)} K_{i_2}^{(2)\dagger} K_{i_1}^{(1)\dagger} \right] \right) \quad (6.27)$$

The symmetry is not present in the way the Kraus operators are composed. Therefore, constructing  $\tau$  requires some knowledge about the Kraus representation of the channels, which we are not given initially. Obtaining this knowledge is equivalent to performing quantum process tomography (Ref. [32]) on the input channels, which scales poorly with system dimension.

Summarised, the argument is that in order to construct  $\tau$  with fixed causal order, we need information about the input channels that is computationally hard to obtain. As mentioned earlier, this is a highly informal argument that is not to be taken as proof. However, the intuitive insight it provides about where the issue may lay leads us to a more formal argument. We now show a particular example where a specific algorithm with fixed causal order fails to outperform our algorithm. This is not a general proof, but it is a proof about a restricted set of algorithms.

Consider a quantum channel  $\mathcal{C}$  which is a mixture of  $M$  unitaries  $\{U_i\}$ , where  $M$  is arbitrarily large. Its Kraus operators are simply the unitaries with normalisation, i.e.  $K_i = \frac{1}{\sqrt{M}} U_i$ . Suppose we are given access to the environment system  $E$  used to produce the mixed channel  $\mathcal{C}$ . Measuring  $E$ , we fix  $\mathcal{C}$  to be equal to one of the unitaries, and we know which one through the measurement outcome.

Now consider the setting of Example 6.1 where we are given two channels of the type described just above. As discussed in Ref. [5], the SWITCH of unitaries can be simulated with fixed causal order, given two copies of them. Therefore, a possible strategy to simulate the SWITCH is to obtain two copies of the same unitary. Different executions of the supercircuit will correspond to different unitaries from the set  $\{U_i\}$ , so the overall result over many samples correctly estimates  $\tau$  for their mixture - the channel  $\mathcal{C}$ .

How do we get two copies of the same unitary? First, take an instance of  $\mathcal{C}$  and measure its environment  $E$ . Now one copy is available of one of the unitaries, say  $U_1$  without loss of generality. Repeat doing this until again  $U_1$  comes out as a result of the measurement. Let  $X$  be the random variable that counts the number of iterations performed in this way. Each of these subsequent measurements has a probability  $1/M$  of yielding  $U_1$ . Therefore the expected number of repetitions is

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} \left( \frac{M-1}{M} \right)^{i-1} \frac{i}{M} = M \quad (6.28)$$

The point is that this strategy attempts to compute  $\tau$  by establishing correlation between the Kraus operators of  $\mathcal{C}$  through postselection, but the postselection probability for each sample can be arbitrarily small as  $M$  increases. The overall number of copies of  $\mathcal{C}$  needed scales with  $M$ . In our algorithm, however, this is not the case. Therefore, we expect an advantage in this particular example when  $M \gg 1$ .

Taking this argument further, we can consider  $N$  channels, each a mixture of  $M$  unitaries. In order to simulate the  $N$ -SWITCH now we need  $N$  copies of the same unitary. The probability for obtaining the same measurement outcome (and hence the same unitary)  $N$  consecutive times is  $(1/M)^N$ . In this case our algorithm's expected advantage becomes exponential.

Formal proof of advantage of our protocol over *any* fixed causal order algorithm is beyond the scope of this thesis. However, the arguments provided in this section illustrate why we expect that this advantage exists.

# Chapter 7

## Towards higher order computation with signalling channels

In this chapter we extend our model of second-order computation in order to allow signalling channels as inputs.

### 7.1 Computational model on signalling channels

In extending our model to signalling channels, we need to introduce extra rules to make sure that no illegal transformations occur. Informally, an illegal transformation is a transformation that reorders the input channels in time in a way that contradicts their causal structure. For instance, if a channel  $\mathcal{A}$  signals to a channel  $\mathcal{B}$ , a supercircuit cannot compose  $\mathcal{B}$  before  $\mathcal{A}$ .

In order to make these statements more precise, from now on we consider the causal structure as a poset (partially ordered set) that defines which channels signal to which [33]. Allowed transformations are then order-preserving (monotone) morphisms on that poset. Hence we define the *causal swap* as a supermap that swaps two timelines only if the resulting morphism on the causal poset is monotone.

From now on we use the following notation for posets. For a poset  $\mathcal{P}$ , let  $S(\mathcal{P})$  denote its underlying set, and

$$\text{IN}(x) := \{y \mid y \in S(\mathcal{P}), y \prec_{\mathcal{P}} x\} \quad (7.1)$$

$$\text{OUT}(x) := \{y \mid y \in S(\mathcal{P}), x \prec_{\mathcal{P}} y\} \quad (7.2)$$

Denote the relation between two elements  $x, y \in S(\mathcal{P})$  by  $\mathcal{P}(x, y)$ .

#### **Definition 7.1 (Signalling supercircuits with definite causal order)**

*A signalling supercircuit with definite causal order describes how a sequence of quantum supermaps acts on an  $N$ -partite signalling channel. In order to process such input, we allow the following three supermaps:*

- Sequential composition of adjacent timelines - the sequential composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{B} \circ \mathcal{A}$ . Composition cannot be applied to some timelines if there are other timelines between them in the supercircuit.
- Parallel composition of adjacent timelines - the parallel composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{A} \otimes \mathcal{B}$ . Composition cannot be applied to some timelines if there are other timelines between them in the supercircuit.

- Causal swap of timelines - deterministic reordering in time of channels such that the resulting morphism on the causal poset is monotone.

The same categorisation as for the basic model (Definition 4.1) holds. There are no third, or higher, order operations, hence the model is of second-order computation. Also, there is no non-deterministic reordering, therefore it is a model of definite causal order.

In order to accommodate for signalling channels in our graphical language, we draw signalling dependencies as arrows between timelines. Consider a channel  $\mathcal{A}$  which signals to another channel  $\mathcal{B}$ . Extending our notation from Equation 4.1, we depict this as:

$$(7.3)$$

The arrows can only point from top to bottom, because of our convention that higher timelines are executed before (or simultaneously with) lower timelines. A supercircuit with an arrow pointing up does not describe a valid physical transformation.

Compositions preserve arrows in the natural way, defined by Equations 7.4 - 7.7 below.

$$\text{IN}(\mathcal{B} \circ \mathcal{A}) = \text{IN}(\mathcal{A}) \cup \text{IN}(\mathcal{B}) \quad (7.4)$$

$$\text{OUT}(\mathcal{B} \circ \mathcal{A}) = \text{OUT}(\mathcal{A}) \cup \text{OUT}(\mathcal{B}) \quad (7.5)$$

$$\text{IN}(\mathcal{A} \otimes \mathcal{B}) = \text{IN}(\mathcal{A}) \cup \text{IN}(\mathcal{B}) \quad (7.6)$$

$$\text{OUT}(\mathcal{A} \otimes \mathcal{B}) = \text{OUT}(\mathcal{A}) \cup \text{OUT}(\mathcal{B}) \quad (7.7)$$

Equation 7.8 below is an example application of Equation 7.5 in the graphical language.

$$(7.8)$$

There are two differences in the definition, compared to non-signalling supercircuits with definite causal order, which stem from the fact that now we consider signalling channels. First, we substituted the swap of channels for the causal swap of channels in order to only describe physical transformations. In the graphical language, this corresponds to never inverting an arrow to point upwards.

The second additional constraint is that compositions can only be applied on adjacent timelines. Without it, the following example issue could occur. Consider  $\mathcal{A}$  signals to  $\mathcal{B}$ , and  $\mathcal{B}$  to  $\mathcal{C}$ . Composing  $\mathcal{A}$  and  $\mathcal{C}$  sequentially, the only ordering consistent with the causal structure is that the resulting channel  $\mathcal{C} \circ \mathcal{A}$  needs to happen simultaneously with  $\mathcal{B}$ .

$$(7.9)$$

This example, also illustrated in the supercircuit above (Equation 7.9), results in outputting an arbitrary bipartite channel that allows signalling in both directions. For convenience, we avoid this scenario by imposing the constraint that only adjacent timelines can be composed.

This constraint does not lead to any reduction of the expressiveness of the model. Any valid composition can still be achieved by first swapping timelines to make them adjacent.

## 7.2 Characterisation of the transformations

As for the previous two models we defined, we now consider the transformations the current model describes.

**Proposition 7.1** *Signalling supercircuits with definite causal order (Definition 7.1) are maps from signalling channels to signalling channels.*

**Proof.** In order to prove the proposition, we need to show that all allowed supermaps in the model preserve the signalling property of channels. That is, we need to show that for all supermaps the resulting morphism on the causal poset is monotone. Let us examine the three allowed supermaps individually. Consider a supercircuit acting on a set  $S(\mathcal{P})$  of channels with causal poset  $\mathcal{P}$ .

First, consider sequential composition on two channels  $\mathcal{A}, \mathcal{B} \in S(\mathcal{P})$ . Let its action on the causal poset be  $f : \mathcal{P} \mapsto f(\mathcal{P})$ .  $f$  is a monotone morphism if

$$\forall \mathcal{X}, \mathcal{Y} \in S(\mathcal{P}) : \mathcal{X} \preceq_{\mathcal{P}} \mathcal{Y} \implies f(\mathcal{X}) \preceq_{f(\mathcal{P})} f(\mathcal{Y}) \quad (7.10)$$

Now we consider all cases for  $\mathcal{X}$  and  $\mathcal{Y}$ .

1. If  $\mathcal{X} = \mathcal{A}$  and  $\mathcal{Y} = \mathcal{B}$ . Then  $f(\mathcal{X}) = f(\mathcal{Y}) = \mathcal{B} \circ \mathcal{A}$ . Hence  $f(\mathcal{X}) \preceq_{f(\mathcal{P})} f(\mathcal{Y})$ .
2. If  $\mathcal{X} \in \{\mathcal{A} \cup \mathcal{B}\}$  and  $\mathcal{Y} \notin \{\mathcal{A} \cup \mathcal{B}\}$ . Then  $\mathcal{X} \preceq_{\mathcal{P}} \mathcal{Y} \stackrel{\text{Eq. (7.5)}}{\implies} f(\mathcal{X}) \preceq_{f(\mathcal{P})} f(\mathcal{Y})$ .
3. If  $\mathcal{X} \notin \{\mathcal{A} \cup \mathcal{B}\}$  and  $\mathcal{Y} \in \{\mathcal{A} \cup \mathcal{B}\}$ . Then  $\mathcal{X} \preceq_{\mathcal{P}} \mathcal{Y} \stackrel{\text{Eq. (7.4)}}{\implies} f(\mathcal{X}) \preceq_{f(\mathcal{P})} f(\mathcal{Y})$ .
4. If  $\mathcal{X}, \mathcal{Y} \notin \{\mathcal{A} \cup \mathcal{B}\}$ . Then  $f(\mathcal{X}) = \mathcal{X}$ ,  $f(\mathcal{Y}) = \mathcal{Y}$  and  $f(\mathcal{P}(\mathcal{X}, \mathcal{Y})) = \mathcal{P}(\mathcal{X}, \mathcal{Y})$ . Hence  $\mathcal{X} \preceq_{\mathcal{P}} \mathcal{Y} \implies f(\mathcal{X}) \preceq_{f(\mathcal{P})} f(\mathcal{Y})$ .

For all cases Equation 7.10 holds. Therefore, sequential composition's action on the causal poset is a monotone morphism. The exact same argument holds for parallel composition, only using Equations 7.6 and 7.7 instead of 7.4 and 7.5. Finally, the causal swap is defined as a monotone morphism on the causal poset. We have shown that all allowed supermaps in the model act as a monotone morphism on the causal poset. Therefore, signalling supercircuits with definite causal order send signalling into signalling channels. ■

The second result is the same as for non-signalling supercircuits with definite causal order (Proposition 4.2).

**Proposition 7.2** *Signalling supercircuits with definite causal order (Definition 7.1) with exactly one output timeline are in a 1-to-1 correspondence with first-order circuits.*

**Proof.** First, consider a supercircuit that outputs a single timeline. By definition, an  $N$ -partite channel is described by  $N$  timelines in our framework. Hence, the output of this supercircuit is a 1-partite channel which by Definition 2.1 is a valid non-signalling channel. By Proposition 4.2 then it is exactly a first-order circuit.

Now we show the opposite - that any circuit can be expressed as a supercircuit with a single output timeline. Any circuit can be seen as a single quantum channel. Quantum channels are represented as states in our model. Hence any circuit is a trivial supercircuit with a single input channel. This part of the proof is the same as in the proof of Proposition 4.2 because a 1-partite channel can be seen both as a signalling and as a non-signalling channel. ■

### 7.3 No-go theorem for naïve SWITCHing of signalling channels

Following the line of development of chapter 5, at this point we should introduce the SWITCH as a resource. However, this turns out to be a bigger challenge when considering signalling channels. Here we present a naïve definition of signalling supercircuits with indefinite causal order, and proceed to prove that its simplicity does not allow the decomposition of the  $N$ -SWITCH. This result indicates that it is not a good computational model, since the set of assumed available operations is infinite, rather than finite.

In order to define the computational model, we first define a new supermap - the causal  $N$ -SWITCH. In the definition, we use the concept of a linear extensions of a poset. By definition, a linear extension of a partially ordered set is a total order on the set, which respects the partial order [34, 35]. We denote the set of linear extensions of a poset  $\mathcal{P}$  with  $\text{LE}(\mathcal{P})$ . Informally, we define a causal  $N$ -SWITCH to be a supermap that takes  $N$  channels as input and outputs a controlled superposition (possibly with some coefficients equal to zero) of all their allowed permutations, defined by the set of all linear extension of the causal poset. The formal definition is the following.

**Definition 7.2 (Causal  $N$ -SWITCH)**

Consider a set  $\{\mathcal{C}^{(i)}\}$  of  $N$  quantum channels with Kraus decompositions

$$\mathcal{C}^{(i)}(\rho) = \sum_{j=1}^{d_i} K_j^{(i)} \rho \left( K_j^{(i)} \right)^\dagger, \quad (7.11)$$

where  $d_i$  is the number of Kraus operators of the  $i^{\text{th}}$  channel. Consider an  $N!$ -dimensional quantum system  $C$  with orthonormal basis  $\{|\pi\rangle \mid \pi \in \text{Perm}(N)\}$ . The causal  $N$ -SWITCH is a supermap that, given the set of channels  $\{\mathcal{C}^{(i)}\}$  as input and using  $C$  as control, outputs a quantum map with Kraus operators  $W_{i_1, \dots, i_N}$  given by

$$W_{i_1, \dots, i_N} = \sum_{\pi \in \text{LE}(\mathcal{P})} |\pi\rangle \langle \pi| \otimes \bigotimes_{j=1}^N K_{i_{\pi(j)}}^{(\pi(j))} \quad (7.12)$$

The only difference in the current definition of the causal  $N$ -SWITCH to the  $N$ -SWITCH (Definition 3.3) is that now we sum over the linear extensions  $\text{LE}(\mathcal{P})$ , rather than over all permutations  $\text{Perm}(S(\mathcal{P}))$ . A causal SWITCH is then an instance of the causal  $N$ -SWITCH for  $N = 2$ . Using these new terms, we can define the computational model as follows.

**Definition 7.3 (Naïve signalling supercircuits with indefinite causal order)**

A signalling supercircuit with indefinite causal order describes how a sequence of quantum supermaps acts on an  $N$ -partite signalling channel. In order to process such input, we allow the following four supermaps:

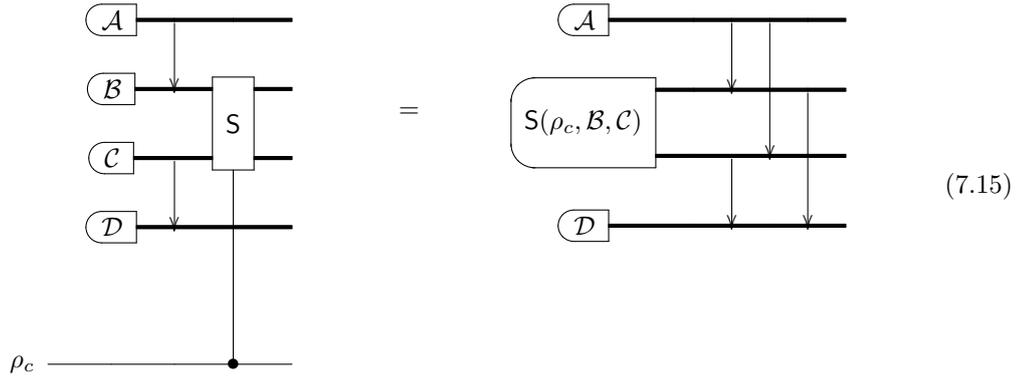
- Sequential composition of adjacent timelines - the sequential composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{B} \circ \mathcal{A}$ . Composition cannot be applied to some timelines if there are other timelines between them in the supercircuit.
- Parallel composition of adjacent timelines - the parallel composition of two channels  $\mathcal{A}$  and  $\mathcal{B}$  is the channel  $\mathcal{A} \otimes \mathcal{B}$ . Composition cannot be applied to some timelines if there are other timelines between them in the supercircuit.
- Causal swap of timelines - deterministic reordering in time of channels such that the resulting morphism on the causal poset is monotone.
- Causal SWITCH of adjacent timelines.

The action of the causal SWITCH on the causal poset is the following. Consider a causal SWITCH with input channels  $\mathcal{A}$  and  $\mathcal{B}$  and output channels  $\mathcal{A}'$  and  $\mathcal{B}'$ . Then

$$\text{IN}(\mathcal{A}') = \text{IN}(\mathcal{B}') = \text{IN}(\mathcal{A}) \cup \text{IN}(\mathcal{B}) \quad (7.13)$$

$$\text{OUT}(\mathcal{A}') = \text{OUT}(\mathcal{B}') = \text{OUT}(\mathcal{A}) \cup \text{OUT}(\mathcal{B}) \quad (7.14)$$

Equations 7.13, 7.14 above define the property we call naïve in the supercircuit definition name. It is illustrated in Equation 7.15 below.



We have defined the causal SWITCH's action on causal relations (Equations 7.13 and 7.14) in that way for the following reason. Consider the supercircuit in Equation 7.15. When we SWITCH  $\mathcal{B}$  and  $\mathcal{C}$ , we get a superposition of the two orderings -  $\mathcal{A}\mathcal{B}\mathcal{C}\mathcal{D}$  and  $\mathcal{A}\mathcal{C}\mathcal{B}\mathcal{D}$ . Because they are in a superposition, it makes sense to allow only such further operations, which do not contradict the causal restrictions of either of these two orders. Hence we overlay the arrows of the two cases - as if  $\rho_c = |0\rangle\langle 0|$  and as if  $\rho_c = |1\rangle\langle 1|$ . This argument leads us to Equations 7.13 and 7.14.

In our model on non-signalling channels we saw that the  $N$ -SWITCH can be decomposed into a (polynomial) number of 2-SWITCHes. This is convenient because only a single operation needs to be assumed as an available resource in the model. This, however, does not hold for the currently discussed model on signalling channels. Now we prove that, unlike in the case of non-signalling supercircuits, the causal  $N$ -SWITCH in signalling supercircuits does not decompose into causal 2-SWITCHes.

**Proposition 7.3 (No-go theorem for causal  $N$ -SWITCH decomposition in naïve signalling supercircuits)** *The causal  $N$ -SWITCH in a naïve signalling supercircuit with indefinite causal order (Definition 7.3) cannot be decomposed into any number of causal SWITCHes on less than  $N$  channels. Therefore, in order for arbitrary causal  $N$ -SWITCHing to be available, all causal  $N$ -SWITCHes for any  $N \geq 2$  need to be assumed as available resources in the model.*

**Proof.** The proof proceeds by induction. The first step is to show that the causal 3-SWITCH cannot be decomposed into causal 2-SWITCHes. Consider the following input to a supercircuit with  $S(\mathcal{P}) = \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$  and  $\mathcal{A} \prec_{\mathcal{P}} \mathcal{B}$ .



A causal 2-SWITCH applied to  $\mathcal{A}$  and  $\mathcal{B}$  has no effect because their two-element poset forms a total order (i.e.  $\mathcal{A} \prec_{\mathcal{P}} \mathcal{B}$ ). A causal 2-SWITCH applied to  $\mathcal{A}$  and  $\mathcal{C}$  also has no effect because  $\mathcal{C} \prec \mathcal{B} \prec \mathcal{A} \notin \text{LE}(\mathcal{P})$ . Acting on  $\mathcal{B}$  and  $\mathcal{C}$  with a causal 2-SWITCH generates two possible linear extensions ( $\mathcal{A} \prec \mathcal{B} \prec \mathcal{C}$  and  $\mathcal{A} \prec \mathcal{C} \prec \mathcal{B}$ ), but not all, since  $\mathcal{C} \prec \mathcal{A} \prec \mathcal{B}$  is not possible. However, at that point the full three-element poset is a total order, and no further causal SWITCHing is possible. Therefore, no way of applying causal 2-SWITCHes leads to the desired output of a causal 3-SWITCH. Hence it is non-decomposable.

For the induction step, suppose all causal  $N$ -SWITCHes for  $N \leq N^*$  are assumed as freely available resources in the model. We need to show that the causal  $(N^* + 1)$ -SWITCH cannot be constructed from any combination of theirs. Consider  $(N^* + 1)$  channels input into a supercircuit where the causal structure of the first  $N^*$  channels is a total order, and the  $(N^* + 1)$ st channel has no causal restrictions. That is, their causal poset  $\mathcal{P}$  is given by:

$$\underbrace{\bullet}_{c_1} \longrightarrow \underbrace{\bullet}_{c_2} \longrightarrow \cdots \longrightarrow \underbrace{\bullet}_{c_{N^*}} \quad \underbrace{\bullet}_{c_{N^*+1}} \quad (7.17)$$

A causal  $(N^* + 1)$ -SWITCH needs to be able to send the  $(N^* + 1)$ st channel to any of the  $(N^* + 1)$  timeline positions, since it has no arrows forbidding this. Consider applying any of the available causal  $N$ -SWITCHes ( $N \leq N^*$ ). If none of the  $N$  channels is the  $(N^* + 1)$ st one, the SWITCH achieves nothing, because its inputs are in a total order. Otherwise, one of its inputs is the  $(N^* + 1)$ st channel. Because  $N < N^* + 1$ , we know by the pigeon-hole principle that this single operation cannot send this channel to all of the timeline positions. However, after it has been applied, the causal structure of all  $(N^* + 1)$  channels is a total order. Hence, no further causal SWITCHing is possible. The goal of sending the  $(N^* + 1)$ st channel to an arbitrary timeline position has not been achieved, though. Therefore it is impossible to achieve using only causal  $N$ -SWITCHes with  $N \leq N^*$ . The causal  $(N^* + 1)$ -SWITCH cannot be decomposed into causal  $N$ -SWITCHes with  $N \leq N^*$ . By induction, the causal  $N$ -SWITCH cannot be decomposed into causal SWITCHes on less than  $N$  channels. ■

Significant literature on the topic of linear extensions of posets exists and may be useful for further work on the causal  $N$ -SWITCH. A particular (classical) result worth mentioning is the proof in Ref. [34] that the problem of counting the number of linear extensions is #P-Complete. This indicates why we expect it to be hard to extend the  $N$ -SWITCH into a causal  $N$ -SWITCH - just counting the number of outputs is #P-Complete.

Another important result from Ref. [35] is that it is possible to generate all linear extensions in constant amortized time, i.e. constant in the size of the set of all linear extensions. The paper argues that this is as fast as any algorithm can do, i.e. constant amortized, which however is still exponential in general, because the set of linear extensions can be exponentially large. It is classically hard both to count and to generate linear extensions, which is the underlying operation of the causal  $N$ -SWITCH. Hence we expect its implementation to be a hard problem.

## 7.4 The causal $N$ -SWITCH as a direct sum

We have shown that the naïve way of defining a computational model with SWITCHing of signalling channels fails because of the  $N$ -SWITCH's impossibility of decomposition. Other models, however, could exist, where this is not an issue. Here we discuss such potential models.

The problem with our naïve approach arose because SWITCHing channels produced overly restrictive causal structures. In reality, the SWITCH's output can be considered as a *direct sum* of the orderings. Each branch of this direct sum corresponds to a particular order, together with its causal structure. Hence, subsequent operations in the supercircuit could be valid if they provide a valid operation for each branch of the direct sum output. In

comparison, the naïve approach requires any operation after the SWITCH to be compatible with all branches.

Splitting cases between different branches of a direct sum allows for relaxation of some constraints. A particular constraint we have imposed on SWITCHes throughout this thesis is the requirement that the types of its input channels must match. This was needed because otherwise the output would be in a superposition of types, which we had no means to deal with. Considering the output of a SWITCH as a direct sum, however, allows us to drop that requirement. Each branch of the direct sum would correspond to a single well-defined type for each timeline.

The reason for our current discussion of expressing the superposition created by a causal  $N$ -SWITCH as a direct sum is that there exists literature which may be of use for further development. In particular, various quantum lambda calculi exist and some of them provide semantic rules for the direct sum, which may be used to discuss the causal  $N$ -SWITCH within their framework. For instance Sabry, *et. al.*'s lambda calculus (Ref. [36]) includes superpositions as direct sums, but only for pure states and processes. We mention the relation of quantum lambda calculus to our problem of defining the causal  $N$ -SWITCH as a possible research direction.

# Chapter 8

## Conclusions

### 8.1 Summary of results

This thesis focuses on quantum computation with indefinite causal order. We have developed a series of higher-order computational models with an increasing level of generality. The first step in this sequence was to rigorously define a basic computational model for second-order quantum computation. We developed a graphical language for the model so that we could visualise computation using supercircuits. In order to better understand our basic model, we also characterised its transformations as maps from non-signalling channels to non-signalling channels.

Having established the basic model and its graphical representation, we then moved on to enriching them with the desired indefinite causal order by adding the quantum SWITCH. We imposed two constraints on the SWITCH in our model. First, that the SWITCH's control system needs to be external, i.e. not acted on by the channels input to the supercircuit. Second, that only channels with matching input and output types can be SWITCHed.

In order to classify the type of transformations of non-signalling supercircuits with indefinite causal order, we introduced the new concept of extended non-signalling channels. With that concept, we could characterise the transformations as maps that preserve a certain invariant, in that case the extended non-signalling property. We then extended the graphical language to provide means for depicting the SWITCH in an intuitive way as a controlled swap of timelines. We then considered the  $N$ -SWITCH as a derived operation that overloads the SWITCH, and included it into the graphical language.

After defining our model, we set out to explore its applications, using the graphical language as a tool. We constructed a novel quantum algorithm, inspired by the DQC1 algorithm. Given a set of quantum channels as input, our algorithm estimates a quantity that depends on the traces of various compositions of their Kraus operators. We conjectured that the algorithm offers an exponential advantage over all algorithms that use a fixed causal order. If this conjecture is proven correct, our algorithm would be the first concrete example where indefinite causal order provides *exponential* computational advantage. In comparison, the strongest advantage of that type currently known is only polynomial (e.g. Araújo, *et al.*'s problem in Ref. [9], discussed earlier). Showing our potential exponential advantage, however, requires further work to extend our arguments for the advantage into a rigorous proof.

Finally, we extended our computational model to describe signalling supercircuits with definite causal order. Trying to take the conclusive step to signalling supercircuits with indefinite causal order, however, turned out to be a significantly harder task. We introduced a simple extension of the computational model including the causal  $N$ -SWITCH, namely a quantum  $N$ -SWITCH that respects the causal structure of its input channels. We then proved that in this model the causal  $N$ -SWITCH cannot be decomposed into causal 2-SWITCHes.

A brief discussion followed on how to define alternative computational models that allow the decomposition of the causal  $N$ -SWITCH into causal 2-SWITCHes.

## 8.2 Discussion on further work

Higher-order quantum computation is still a new field that has a lot of potential development. In this thesis we have worked towards formalising computation by defining our computational model and the related graphical language. One necessary future development is to find a way to go around the no-go theorem and incorporate signalling channels in supercircuits with indefinite causal order. As discussed in the main text, one potential direction to achieve that is to consider the causal  $N$ -SWITCH's output as a direct sum, and relate the work to a quantum lambda calculus which provides semantic rules for the direct sum.

Another restriction that also needs to be relaxed is the constraint that all SWITCHes are controlled externally, by a system acted on only by operations with fixed causal order. In principle, however, there is no reason why the control systems could not also be acted on by the supercircuit's channels. In the analogy of our model describing a quantum computer, dynamically programmable by another quantum computer this becomes another level of abstraction where the two computers dynamically program each other. Defining well this kind of feedback machine where SWITCHes are controlled internally is another possible further research topic that would enrich our model.

Our models, even with the generalisations described here, would still be highly restricted models of second-order computation, because only a few second-order maps are allowed. Defining a computational model that utilises fully the second order of the hierarchy is a highly important question whose advantages we cannot currently predict. Looking even further, this is only the first step to understanding arbitrary higher-order computation. Once a more firm grasp has been established of the second order, a similar treatment should be developed for higher orders. Having this broader roadmap in view, we wish the current thesis to set one of the cornerstones for the general treatment of higher-order computation.

Another avenue for further research lies in utilising our framework itself, rather than developing it. In the same way it allowed us to construct a new second-order algorithm, it may lead to other novel quantum algorithms being more easily accessible through the graphical language. If an algorithm of practical importance is discovered that way, the available experimental research on indefinite causal order would allow it to be verified, and even potentially exploited in practice.

We see the two main types of further work - developing the model, and using it for applications - as being in a self-reinforcing cycle. The more advanced the computational model is, the more algorithms could be discovered that fit in it. The more algorithms discovered, the better understanding we would get of how the model itself works, and hence how to enrich it further.

# Bibliography

- [1] P. W. Shor, “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer,” *SIAM J. Comput.* **26** no. 5, (Oct, 1997) 1484–1509.
- [2] L. K. Grover, “A Fast Quantum Mechanical Algorithm for Database Search,” in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pp. 212–219. ACM, New York, NY, USA, 1996.
- [3] D. Deutsch, “Quantum Computational Networks,” *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences* **425** (1989) 73–90.
- [4] L. Hardy, *Quantum Gravity Computers: On the Theory of Computation with Indefinite Causal Structure*, pp. 379–401. Springer Netherlands, Dordrecht, 2009.
- [5] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, “Quantum computations without definite causal structure,” *Phys. Rev. A* **88** no. 2, (Aug, 2013) 22318.
- [6] O. Oreshkov, F. Costa, and Č. Brukner, “Quantum correlations with no causal order,” *Nature Communications* **3** (Oct, 2012) 1092.
- [7] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Transforming quantum operations: Quantum supermaps,” *EPL (Europhysics Letters)* **83** no. 3, (2008) 30004.
- [8] G. Chiribella, “Perfect discrimination of no-signalling channels via quantum superposition of causal structures,” *Phys. Rev. A* **86** no. 4, (Oct, 2012) 40301.
- [9] M. Araújo, F. Costa, and Č. Brukner, “Computational Advantage from Quantum-Controlled Ordering of Gates,” *Physical review letters* **113** (2014) 250402.
- [10] D. Ebler, S. Salek, and G. Chiribella, “Enhanced Communication with the Assistance of Indefinite Causal Order,” *Phys. Rev. Lett.* **120** no. 12, (Mar, 2018) 120502.
- [11] A. Kissinger and S. Uijlen, “Picturing Indefinite Causal Structure,” in *Proceedings 13th International Conference on Quantum Physics and Logic, {QPL} 2016, Glasgow, Scotland, 6-10 June 2016.*, R. Duncan and C. Heunen, eds., vol. 236 of *EPTCS*, pp. 87–94. 2016.
- [12] A. Kissinger and S. Uijlen, “A categorical semantics for causal structure,” 2017.
- [13] B. Coecke and A. Kissinger, *Picturing Quantum Processes A First Course in Quantum Theory and Diagrammatic Reasoning*. 2017.
- [14] E. Knill and R. Laflamme, “Power of One Bit of Quantum Information,” *Phys. Rev. Lett.* **81** no. 25, (Dec, 1998) 5672–5675.
- [15] L. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. Hamel, L. Rozema, Č. Brukner, and P. Walther, “Experimental Superposition of Orders of Quantum Gates,” *Nature Communications* **6** (2014) .

- [16] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, “Experimental verification of an indefinite causal order,” *Science Advances* **3** no. 3, (2017) .
- [17] R. Shankar, *Principles of quantum mechanics*. Plenum Press, 1994.
- [18] J. von Neumann, “Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1927** (1927) 245–272.
- [19] K. Kraus, A. Böhm, J. Dollard, and W. Wootters, eds., *States, Effects, and Operations Fundamental Notions of Quantum Theory*, vol. 190 of *Lecture Notes in Physics*, Berlin Springer Verlag. 1983.
- [20] A. Jamiołkowski, “An effective method of investigation of positive maps on the set of positive definite operators,” *Reports on Mathematical Physics* **5** no. 3, (1974) 415–424.
- [21] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra and its Applications* **10** no. 3, (1975) 285–290.
- [22] M. Jiang, S. Luo, and S. Fu, “Channel-state duality,” *Phys. Rev. A* **87** no. 2, (Feb, 2013) 22310.
- [23] A. Royer, “Wigner function in Liouville space: A canonical formalism,” *Phys. Rev. A* **43** no. 1, (Jan, 1991) 44–56.
- [24] G. M. D’Ariano, P. L. Presti, and M. F. Sacchi, “Bell measurements and observables,” *Physics Letters A* **272** no. 1, (2000) 32–38.
- [25] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, “Properties of quantum nonsignaling boxes,” *Phys. Rev. A* **74** no. 1, (Jul, 2006) 12305.
- [26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
- [27] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Theoretical framework for quantum networks,” *Phys. Rev. A* **80** no. 2, (Aug, 2009) 22339.
- [28] T. Colnaghi, G. M. D’Ariano, S. Facchini, and P. Perinotti, “Quantum computation with programmable connections between gates,” *Physics Letters, Section A: General, Atomic and Solid State Physics* **376** no. 45, (2012) 2940–2943.
- [29] S. Facchini and S. Perdrix, “Quantum circuits for the unitary permutation problem,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9076, pp. 324–331. 2015. [arXiv:1405.5205](https://arxiv.org/abs/1405.5205).
- [30] T. M. Rambo, J. B. Altepeter, P. Kumar, and G. M. D’Ariano, “Functional quantum computing: An optical approach,” *Phys. Rev. A* **93** no. 5, (May, 2016) 52321.
- [31] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, New York, NY, USA, 2005.
- [32] I. L. Chuang and M. A. Nielsen, “Prescription for experimental determination of the dynamics of a quantum black box,” *Journal of Modern Optics* **44** no. 11-12, (1997) 2455–2467.
- [33] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, “Space-time as a causal set,” *Phys. Rev. Lett.* **59** no. 5, (Aug, 1987) 521–524.

- [34] G. Brightwell and P. Winkler, “Counting Linear Extensions is #P-complete,” in *Proceedings of the Twenty-third Annual ACM Symposium on Theory of Computing*, STOC '91, pp. 175–181. ACM, New York, NY, USA, 1991.
- [35] G. Pruesse and F. Ruskey, “Generating Linear Extensions Fast,” *SIAM Journal on Computing* **23** no. 2, (1994) 373–386.
- [36] A. Sabry, B. Valiron, and J. Kaizer Vizzotto, “From Symmetric Pattern-Matching to Quantum Control,” 2018.