HAL
open science

Multipartite causal relations in quantum theory

Julian Wechs

To cite this version:

Julian Wechs. Multipartite causal relations in quantum theory. Quantum Physics [quant-ph]. Université Grenoble Alpes [2020-..], 2020. English. NNT : 2020GRALY008 . tel-02949496

HAL Id: tel-02949496

https://theses.hal.science/tel-02949496v1

Submitted on 25 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique Théorique

Arrêté ministériel : 25 mai 2016

Présentée par

Julian WECHS

Thèse dirigée par **Cyril BRANCIARD**, Chargé de Recherche, Université Grenoble Alpes
et codirigée par **Alexia AUFFEVES**, Université Grenoble Alpes
et **Mehdi MHALLA**, CNRS
préparée au sein du **Laboratoire Institut Néel**
dans **l'École Doctorale Physique**

Relations causales multipartites en théorie quantique

Multipartite causal relations in quantum theory

Thèse soutenue publiquement le **9 mars 2020**,
devant le jury composé de :

Monsieur CASLAV BRUKNER
PROFESSEUR, UNIVERSITE DE VIENNE - AUTRICHE, Rapporteur
Monsieur PABLO ARRIGHI
PROFESSEUR, UNIVERSITE AIX-MARSEILLE, Président , Rapporteur
Monsieur NICOLAS BRUNNER
PROFESSEUR ASSOCIE, UNIVERSITE DE GENEVE - SUISSE, Examinateur
Monsieur SIMON PERDRIX
CHARGE DE RECHERCHE HDR, CNRS DELEGATION CENTRE-EST, Examinateur
Monsieur AURELIEN DREZET
CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALPES, Examinateur

Abstract

In recent years, the investigation of causal relations in quantum theory has attracted a lot of interest. In particular, it has been found that it is possible to conceive of scenarios where some parties perform operations that are compatible with quantum theory locally, but that cannot be embedded into a global background causal structure. Such indefinite causal structures are of interest from a fundamental point of view, but also from the perspective of quantum information processing, since they do not fit into the usual paradigm of *quantum circuits*, which assumes a definite causal order. The main aim of this thesis is to study indefinite quantum causal relations involving more than two parties. Compared to the bipartite case, there are many new aspects and complications that arise in multipartite situations, which need to be clarified in order to fundamentally understand quantum causal structures, and to shed light on their implications and potential usefulness for quantum information processing.

A suitable mathematical framework for the investigation of quantum causal relations is the *process matrix formalism*. In this framework, the notion of *causal nonseparability* was introduced in order to qualify scenarios that are incompatible with a definite causal order. In Chapter 2, we study how to generalise this concept to the multipartite case, how to characterise multipartite causally (non)separable quantum processes, and how to certify multipartite causal nonseparability. Another important topic is to determine which quantum scenarios with indefinite causal order are physically implementable, and how they can be realised concretely. In Chapter 3, we introduce two new classes of physically realisable multipartite quantum processes, and characterise them in terms of their process matrix descriptions. In particular, we define the class of *quantum circuits with quantum control of causal order*. The simplest example of such a circuit is the *quantum switch*, a protocol in which the order between two operations is controlled by a qubit in a superposition state, and which defines a causally nonseparable process. The class we introduce also contains more general examples of causally nonseparable processes with new features. We then show how the process matrix characterisation of this class allows us to search for new quantum information processing applications of causal nonseparability. In Chapter 4, we investigate a particular quantum communication effect in a scenario involving coherent control between two quantum channels. This leads us to a more general analysis of the notion of a coherently controlled channel, which involves certain subtilities. In Chapter 5, we turn to another problem that arises in multipartite scenarios, namely whether a given phenomenon is *genuinely multipartite* or not. More particularly, we study *genuinely multipartite (non)causal correlations*. In Chapter 6, we show that anomalous weak values are possible without post-selection. In Chapter 7, we show that certain tripartite process matrices, namely those that are *unitarily extensible*, have a realisation on so-called *time-delocalised subsystems*, i.e., quantum subsystems that are not associated with a definite time. The class of unitarily extensible tripartite process matrices is larger than the class of quantum circuits with quantum control, and in particular contains processes that violate so-called *causal inequalities*.

Résumé

Ces dernières années, une grande attention a été portée à l'étude des relations causales en théorie quantique. Plus particulièrement, il a été montré qu'il est possible de concevoir des scénarios dans lesquelles des parties réalisent des opérations qui sont compatibles avec la théorie quantique, mais qui ne peuvent pas être intégrées dans une structure causale globale. De tels ordres causaux indéfinis sont intéressants d'un point de vue fondamental, mais aussi sous l'angle de l'informatique quantique, étant donné qu'ils sortent du paradigme habituel des *circuits quantiques*, dans lequel on présuppose un ordre causal bien défini. L'objectif principal de cette thèse est d'étudier des relations causales indéfinies dans des scénarios comportant plus de deux parties. Comparées au cas bipartite, les situations multipartites font apparaître des aspects et problèmes nouveaux qui nécessitent d'être clarifiés afin de comprendre fondamentalement les structures causales quantiques indéfinies, et de mettre en évidence leurs implications et leur utilité potentielle pour l'informatique quantique.

Une approche particulière pour étudier des relations causales quantiques est le formalisme des *matrices de processus*. Dans ce formalisme, le concept de la *non-séparabilité causale* a été introduit afin de qualifier des scénarios qui ne sont pas compatibles avec un ordre causal. Dans le chapitre 2, nous étudions comment généraliser ce concept au cas multipartite, comment caractériser des processus multipartites causalement (non)-séparables, et comment certifier la non-séparabilité causale multipartite. Un autre sujet important est de déterminer quels scénarios quantiques avec un ordre causal indéfini sont physiquement réalisables, et comment ils peuvent être réalisés concrètement. Dans le chapitre 3, nous introduisons deux nouvelles classes de processus quantiques multipartites qui sont réalisables en pratique et nous caractérisons les matrices de processus correspondantes. En particulier, nous définissons la classe des *circuits quantiques avec un ordre causal contrôlé de manière quantique*. L'exemple le plus simple d'un tel circuit est le *quantum switch* : un protocole dans lequel l'ordre entre deux opérations est contrôlé par un qubit dans un état de superposition, et qui définit un processus causalement non-séparable. La classe que nous introduisons contient des exemples plus généraux de processus causalement non-séparables avec de nouvelles propriétés. Nous montrons ensuite comment la caractérisation des processus de cette classe nous permet d'étudier de nouvelles applications de la non-séparabilité causale. Dans le chapitre 4, nous étudions un effet particulier de communication quantique dans un scénario avec un contrôle cohérent entre deux canaux quantiques. Ceci nous conduit à une analyse plus générale de la notion de canal quantique contrôlé de façon cohérente, qui implique certaines subtilités. Dans le chapitre 5, nous abordons un autre problème inhérent aux scénarios multipartites, qui est de savoir si un phénomène donné est véritablement multipartite (« genuinely multipartite ») ou non. Plus particulièrement, nous étudions des *corrélations (non)-causales* « véritablement multipartites ». Dans le chapitre 6, nous mettons en évidence que des valeurs faibles anormales sont possibles sans post-sélection. Enfin, dans le chapitre 7, nous montrons qu'une certaine classe de matrices de processus tripartites, à savoir celles qui sont *unitairement extensibles*, ont une réalisation sur des sous-systèmes dits

temporellement délocalisés, c'est-à-dire des sous-systèmes quantiques qui ne sont pas associés à un temps bien défini. Cette classe est plus grande que la classe des circuits quantiques avec un ordre causal contrôlé de manière quantique. Un point intéressant est qu'elle contient des processus qui violent des *inégalités causales*.

Acknowledgements

I would like to thank all the people who contributed directly or indirectly to the completion of this work. First, I want to thank my supervisor Cyril Branciard and my co-supervisors Mehdi Mhalla and Alexia Auffèves for the opportunity to spend the last three years studying such an interesting topic, the competent guidance at all times, and the great working environment with many possibilities to develop collaborations and to participate in conferences and summer schools. I also would like to thank all the other people I have worked with and learned from during this thesis, in particular Alastair A. Abbott, Nicolas Brunner, Fabio Costa, Hippolyte Dourdent, Dominic Horsman, Ognyan Oreshkov, Marco Túlio Quintino and Ralph Silva.

I would also like to express my gratitude to all the people from Institut Néel who made daily life there a very pleasant experience. In particular, I will remember the great fellowship with my "F218" office mates and the other "PLUMs". Thanks Marco, Hippolyte, Laurie, Alexandra, Pierre, Smail, Guilliam, Noah, Juliette, Patrice, Maria, Morgane, Richard and all the others for the great time spent at work, and during our various activities outside the lab. I also want to thank all other friends that I met during my time in Grenoble, in particular Amandine, Safi and Kaveh. A special thanks goes to Safi, who finished his thesis at the same time as myself and who always had a motivating word. Thanks also to my friends from university Markus, Reimund and Simon. I am happy that we still see each other regularly despite the geographical distance.

Last but not least, I want to thank my family, and my girlfriend Marie for her continuous and ongoing invaluable support.

Contents

Introduction **1**

1 Preliminaries **6**
1.1 The process matrix formalism . . . 7
1.1.1 Local quantum events and the Choi-Jamiołkowski isomorphism . . . 7
1.1.2 Process matrices . . . 9
1.1.3 Examples . . . 11
1.2 Causal (non)separability . . . 12
1.2.1 Process matrices compatible with a definite causal order . . . 12
1.2.2 Causally (non)separable process matrices . . . 13
1.2.3 Witnesses of causal nonseparability . . . 14
1.3 Noncausal correlations . . . 17
1.4 Towards multipartite scenarios . . . 19
1.4.1 Multipartite process matrices . . . 19
1.4.2 Multipartite causal (non)separability and (non)causal correlations . . . 20
1.5 A tripartite process: The quantum switch . . . 22
1.6 Information processing advantages using indefinite causal structures . . . 24

2 On the definition and characterisation of multipartite causal (non)separability **26**
2.1 Defining multipartite causal (non)separability . . . 27
2.1.1 Araújo *et al.*'s definition . . . 27
2.1.2 Oreshkov and Giarmatzi's definitions . . . 27
2.1.3 Comparison . . . 29
2.1.4 Our choice of definition . . . 31
2.2 Characterising multipartite causal (non)separability . . . 32
2.2.1 Valid process matrices and compatibility with a fixed causal order . . . 33
2.2.2 Bipartite and tripartite causally (non)separable process matrices . . . 35
2.2.3 General multipartite causally (non)separable process matrices . . . 36
2.2.3.1 Necessary condition . . . 37
2.2.3.2 Sufficient condition . . . 39
2.2.4 Multipartite witnesses of causal nonseparability . . . 41
2.2.5 Examples . . . 42
2.3 Discussion . . . 45

3 Quantum circuits with quantum versus classical control of causal order **47**
3.1 The link product . . . 48

3.2 Process matrices as quantum supermaps . . . 50
3.3 Quantum circuits with fixed causal order . . . 52
3.3.1 Description . . . 52
3.3.2 Characterisation . . . 54
3.3.3 Examples . . . 57
3.4 Quantum circuits with classical control of causal order . . . 57
3.4.1 Description . . . 57
3.4.2 Characterisation . . . 60
3.4.3 Example . . . 63
3.5 Quantum circuits with quantum control of causal order . . . 64
3.5.1 Revisiting the description of quantum circuits with classical control of causal order . . . 64
3.5.2 Turning the classical control into a coherent control of causal order . . . 68
3.5.3 Characterisation . . . 71
3.5.4 Examples . . . 74
3.5.5 Correlations generated by quantum circuits with quantum control of causal order . . . 75
3.6 Quantum superinstruments . . . 75
3.6.1 Quantum superinstruments with fixed causal order . . . 76
3.6.2 Quantum superinstruments with classical control of causal order . . . 77
3.6.3 Quantum superinstruments with quantum control of causal order . . . 79
3.7 Applications . . . 82
3.8 Discussion . . . 84

4 Communication through coherent control of quantum channels 86
4.1 Communication through the "depolarising quantum switch" . . . 87
4.2 Communication through coherently-controlled depolarising channels . . . 88
4.3 Dependence on channel implementation . . . 90
4.4 Distinguishing different implementations of coherently-controlled channels . . . 93
4.5 Channel implementation independence for the full quantum switch and other quantum processes . . . 93
4.6 Holevo information of the coherently controlled depolarising channels . . . 95
4.7 Discussion . . . 96

5 Genuinely multipartite noncausality 98
5.1 Genuinely N-partite noncausal correlations . . . 99
5.2 Analysis of the tripartite "lazy scenario" . . . 102
5.2.1 Characterisation of the polytope of 2-causal correlations . . . 103
5.2.2 Violations of 2-causal inequalities by process matrix correlations . . . 106
5.3 Discussion . . . 108

6 Anomalous weak values without post-selection 110
6.1 The two-time state formalism . . . 111
6.2 Weak values and measurements . . . 112
6.3 An anomalous weak value without post-selection . . . 115
6.4 Analysis for arbitrary observables . . . 119
6.5 Discussion . . . 122

7 Noncausal processes on time-delocalised subsystems **123**
7.1 Time-delocalised quantum subsystems 124
7.2 Quantum circuits with quantum control of causal order on time-delocalised subsystems 125
7.3 Bipartite unitarily extensible process matrices on time-delocalised subsystems . . 128
7.4 Generalisation to tripartite unitarily extensible processes 129
7.5 Example: A noncausal process on time-delocalised subsystems 131
7.6 Discussion 133

Conclusion **134**

Résumé en Français **136**

A Quantum circuits with quantum control of causal order cannot violate causal inequalities **142**

Abbreviations

CJ	Choi-Jamiołkowski
CP	completely positive
CPTP	completely positive trace-preserving
CS	causal separability, causally separable
ECS	extensible causal separability, extensibly causally separable
GYNI	guess your neighbour's input
LGYNI	lazy guess your neighbour's input
OG	Oreshkov and Giarmatzi
POVM	positive operator-valued measure
QC-CC	quantum circuit(s) with classical control of causal order
QC-FO	quantum circuit(s) with fixed causal order
QC-QC	quantum circuit(s) with quantum control of causal order
SDP	semidefinite programming

Introduction

The concept of causality is essential to physics, and to our perception of the world in general. Our usual understanding is that events are embedded into a definite causal order. That is, they are ordered according to some time parameter, such that the events in the past can influence the events in the future, but not vice versa. We are all familiar with this idea from everyday experience. For example, the rising of the sun causes the rooster to crow, but whether the rooster crows or not does not have any influence on the sun.

One may now wonder whether this concept is really fundamental, or whether one could think of some scenario which is not compatible with such an underlying background causal structure. This question is particularly interesting when we turn to the quantum world, where many things are strange and counterintuitive—for instance, physical quantities can be subject to uncertainty and in superposition states, which leads to phenomena fundamentally different from classical physics. This raises the question of whether the causal structure between events can itself be subject to quantum effects in a similar way.

On the one hand, these issues are relevant for the foundations of physics. Questions of that kind arise, for instance, at the interface of quantum theory and gravity, where situations like a "quantum superposition of spacetime geometries" that lead to indefinite causal relations could appear [1]. One of the motivations for studying indefinite causal structures is therefore the hope for new conceptual insights into the relationship between general relativity and quantum theory, which could lead to progress in the attempts of combining both to a quantum theory of gravity [2].

On the other hand, these questions have implications for quantum information theory. The standard paradigm in quantum computation is that of *quantum circuits*, which assumes a definite causal structure. In this framework, one considers sequences of transformations ("quantum gates") that are applied to quantum systems in a causally ordered way—a certain quantum gate is applied either before or after the others. An interesting perspective, motivating the investigation of quantum causal relations from a more applied point of view, is whether relaxing this assumption of a predefined causal order could lead to new resources for quantum information processing, and allow for new protocols or computational advantages to be realised [3].

Due to their importance from these different perspectives, there has recently been a lot of interest from the quantum foundations and quantum information community in these questions revolving

around the role of time and causality in quantum theory, and a variety of different frameworks and approaches have been proposed to study them [4–10]. In this thesis, we mainly use the *process matrix formalism*, which was introduced in Ref. [4]. In this framework, one assumes multiple parties that perform quantum operations, but without any *a priori* causal order into which they are embedded, and one considers the most general correlations that the parties can establish in such a situation. Through this "top-down" approach, one recovers standard quantum scenarios, for instance quantum states shared between the parties, or causally ordered quantum circuits in which the parties are connected. However, one also finds scenarios that are incompatible with any definite causal order [4, 11, 12].

The central object of this formalism that relates the parties, and encodes the causal relations between them, is the *process matrix*, which can be understood as a generalisation of a multipartite density matrix that allows for the description of signalling and non-signalling scenarios in a unified way. Process matrices that are incompatible with any definite causal order are called *causally nonseparable* [4]. This terminology suggests that there is some analogy with the concept of entanglement (as opposed to separability) for quantum states. Indeed, this analogy exists, and it extends further. Some causally nonseparable process matrices are able to generate *noncausal correlations*, and to violate *causal inequalities* [4, 12], which are device-independent constraints similar in concept to Bell inequalities. Furthermore, one can construct *witnesses of causal nonseparability* [11, 13], in analogy to entanglement witnesses.

This mathematical and conceptual apparatus to analyse scenarios with indefinite causal order has been developed mainly for the case of two parties. Some initial multipartite generalisations have been studied [11, 14, 15], but many open questions remain. A central goal of this thesis is to gain a comprehensive understanding of multipartite causal relations in quantum mechanics, and to further develop the theoretical tools needed for their investigation.

Another open question of crucial importance is which process matrices with indefinite causal order have an operational interpretation or physical realisation. Up to the present, it is not known whether all process matrices can be implemented physically. An example of a physically realisable situation that can be described in terms of a causally nonseparable process matrix is the *quantum switch* [3], a protocol in which the order between two quantum operations is controlled by a quantum bit that is itself in a superposition. This leads to a "superposition of causal orders", in which one cannot say which operation is applied first. Such protocols have indeed been realised in practical experiments [16–22], and their incompatibility with a definite causal order has been certified experimentally. The quantum switch has also been shown to enable various computational and communication advantages over causally ordered configurations [23–29]. Another aim of this thesis is to find new examples of physically realisable processes with an indefinite causal order, beyond the quantum switch, and to study their potential usefulness for quantum information processing tasks.

It is assumed that the reader is familiar with the basic formalism of quantum theory, including the treatment of open quantum systems (as presented, for instance, in Chapters 2 and 8 of [30] or

part II of [31]). In Chapter 1, we introduce the process matrix framework, as well as the notions of causal (non)separability and (non)causal correlations in the bipartite case. We also outline some of the generalisations to multiple parties that have been carried out in the past, and we discuss the example of the quantum switch, as well as some information processing applications of indefinite causal orders that have been identified in earlier works.

The two core parts of this thesis are presented in Chapter 2 and Chapter 3.

In Chapter 2, we present the results of Ref. [32], where we investigate the concept of causal (non)separability in the multipartite case. The first step is to define multipartite causal (non)separability, which is already not so clear. We compare several different definitions that have been proposed previously in the literature, and we identify an incompatibility between them. We resolve this issue by proposing our own definition, which we argue is the most natural one for multipartite scenarios. We then derive conditions to characterise multipartite causally (non)separability in practice, and we generalise the technique of witnesses of causal nonseparability.

In Chapter 3 (which will be published as part of an upcoming paper Ref. [33]), we then explore new examples of physically realisable quantum processes with an indefinite causal order. Through a constructive, "bottom-up approach" we first define a class of *quantum circuits with classical control of causal order*, which are still compatible with a definite causal order, established dynamically during the protocol. Building on this, we then introduce *quantum circuits with quantum control of causal order*, which generalise the quantum switch, and which define causally nonseparable processes in general. Extending similar results for quantum circuits with a fixed causal order ("quantum combs"), we characterise the sets of process matrices that these protocols correspond to. We furthermore generalise our analysis to probabilistic, post-selected circuits, which define *quantum superinstruments*, and we show how our characterisations allow for a systematic study of possible quantum information processing applications through semidefinite optimisation.

Three additional projects are presented in Chapters 4–6.

In Chapter 4, we present the results from Ref. [34]. We study a particularly interesting quantum communication effect identified in a previous work [29], where it has been shown that the quantum switch allows for information to be transmitted through certain noisy quantum channels. We show this feature to be present also in an even simpler scenario where one coherently controls between applying one channel or the other (rather than between the two possible orders). Based on this result, we argue that this effect should be attributed to the presence of this coherent control, rather than indefinite causal order. Our analysis of this example leads us to a closer inspection of the concept of a "coherently controlled quantum operation", which we find to be ill-defined in general.

A further question that arises in multipartite scenarios is whether a given phenomenon is "genuinely multipartite", in the sense that it cannot be reduced to a phenomenon involving only

a subset of parties. In Chapter 5 (which is based on Ref. [35]) we study how to define and characterise *genuinely multipartite noncausal correlations*, which cannot arise in multipartite scenarios where there exists no definite causal order between all parties, but still some "coarse-grained" causal order between certain subsets of parties.

In Chapter 6, we present a further result (presented in Ref. [36]) which was obtained when studying certain questions about the relation between the process matrix formalism and the framework of *multiple-time states*. We show that it is possible to obtain an anomalous weak value, a counterintuitive phenomenon that can arise in weak measurements, in scenarios where there is no post-selection.

In the final Chapter 7, we present another project in progress. The above-mentioned experiments realising the quantum switch have been accompanied by some debate about their interpretation, in particular whether they can really be considered "realisations" of processes with an indefinite causal order, or whether they are just "simulations" of such processes [37–39]. In Ref. [38], it has been argued that these experiments can indeed be considered actual realisations, with the parties acting on *time-delocalised subsystems*, which are concrete, physical quantum systems associated with multiple times. In Chapter 7, we first show how this argument applies to the quantum circuits with quantum control of order that we introduced. Based on our results from Chapter 3, we then extend a proof from Ref. [38] to the tripartite case, which shows that there are even more general processes that have a realisation on time-delocalised subsystems, including certain processes violating causal inequalities.

The format of this thesis is paper-based. It comprises literal excerpts from the following works that have already been published or submitted for publication:

- *Julian Wechs, Alastair A. Abbott, and Cyril Branciard.* On the definition and characterisation of multipartite causal (non)separability, *New J. Phys. 21 013027 (2019)* (Chapter 2).
- *Alastair A. Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, and Cyril Branciard.* Communication through coherent control of quantum channels, *arXiv:1810.09826 [quant-ph]* (Chapter 4).
- *Alastair A. Abbott, Julian Wechs, Fabio Costa, and Cyril Branciard.* Genuinely multipartite noncausality, *Quantum 1, 39 (2017)* (Chapter 5).
- *Alastair A. Abbott, Ralph Silva, Julian Wechs, Nicolas Brunner, and Cyril Branciard.* Anomalous Weak Values Without Post-Selection, Quantum 3, 194 (2019) (Chapter 6).

These manuscripts have been reworked to a varying extent, in order to embed them into a coherent whole. That is, some modifications have been made in order to avoid repetitions, to make notations consistent, and to add relevant details or explanations where appropriate. A note is included at the beginning of each of these chapters, describing the precise correspondence between its content and the respective publication.

The results of Chapter 3 will be presented as part of an upcoming manuscript:

- *Julian Wechs, Hippolyte Dourdent, Alastair A. Abbott, and Cyril Branciard.* Quantum circuits with quantum versus classical control of causal order, *in preparation (2020).*

A manuscript based on the results of Chapter 7 is in preparation as well.

Chapter 1

Preliminaries

In this first chapter, the basic concepts and mathematical tools that are used in this work to study quantum scenarios without definite causal order are introduced. Since these ideas are most easily understood in the bipartite case, we start by considering scenarios with two parties. We present the process matrix framework, as well as the notion of causal (non)separability, which qualifies scenarios that are incompatible with a definite causal order. We also explain the concept of (non)causal correlations, which characterises the (in)compatibility with a global causal order in a device-independent way. We then build the ground for the investigation of multipartite scenarios, by outlining how the framework itself, as well as the notion of causal correlations, generalise to multiple parties. Furthermore, the example of the quantum switch, a particular tripartite causally nonseparable process with a physical interpretation, is discussed. We conclude the first chapter by summarising some of the quantum information processing tasks and advantages that have been studied in the context of indefinite causal orders.

1.1 The process matrix formalism

The mathematical framework predominantly used in this thesis is the *process matrix formalism*, a framework introduced in Ref. [4] to study scenarios without a definite, global causal order between quantum operations. In the process matrix formalism, quantum events are assumed to take place locally, but the causal order between them is not specified *a priori*. The physical resource relating the local events is described by a *process matrix*, which, broadly speaking, is a generalisation of a multipartite density matrix allowing also for the description of signalling scenarios, such as quantum channels.

1.1.1 Local quantum events and the Choi-Jamiołkowski isomorphism

In the bipartite scenario, we consider two parties, who we shall call Alice (A) and Bob (B), and who are each associated with closed laboratories. That is, the parties perform an experiment during which their interactions with the "outside world" (and hence with each other) are restricted to opening their laboratories only once to let an incoming physical system enter, and once to send out an outgoing system. Alice and Bob may choose local operations to perform within their laboratories, possibly depending on some external (classical) input x or y for A and B, and producing (classical) measurement outcomes a and b, respectively. The correlations established between the parties after repeating the experiment many times are described by the conditional probability distribution $P(a, b|x, y)$.

While no assumption is made about the global causal order between the parties, we assume that the local operations performed inside the laboratories are described by standard quantum theory. We can therefore assign some "incoming" and "outgoing" Hilbert spaces to the parties, which we denote $\mathcal{H}^{A_I}, \mathcal{H}^{A_O}$ (for Alice) and $\mathcal{H}^{B_I}, \mathcal{H}^{B_O}$ (for Bob), of dimensions $d_{A_I}, d_{A_O}, d_{B_I}$ and d_{B_O}, respectively. For convenience we also define $\mathcal{H}^{A_{IO}} := \mathcal{H}^{A_I} \otimes \mathcal{H}^{A_O}$, $\mathcal{H}^{B_{IO}} := \mathcal{H}^{B_I} \otimes \mathcal{H}^{B_O}$, $d_{A_{IO}} := d_{A_I} d_{A_O}$ and $d_{B_{IO}} := d_{B_I} d_{B_O}$. The space of linear operators over a Hilbert space $\mathcal{H}^X$ will be denoted by $\mathcal{L}(\mathcal{H}^X)$. In this thesis, we will only consider finite-dimensional Hilbert spaces; for a generalisation of the framework to infinite-dimensional systems, see Ref. [40].

According to quantum theory, Alice and Bob's local operations can most generally be described as quantum instruments [41]. A quantum instrument is a quantum evolution that admits multiple classical outputs ("*measurement outcomes*") while also keeping track of the output quantum state corresponding to each outcome. Mathematically, a quantum instrument is described by a collection of *completely positive* (CP) maps (associated to the different classical outputs), which sum up to a *completely positive trace-preserving* (CPTP) map. Alice's instrument, for instance, is thus described as a collection $\{\mathcal{M}^A_{a|x}\}_a$, where each $\mathcal{M}^A_{a|x} : \mathcal{L}(\mathcal{H}^{A_I}) \to \mathcal{L}(\mathcal{H}^{A_O})$ is a CP map, and with $\mathrm{Tr}(\sum_a \mathcal{M}^A_{a|x}(\rho^{A_I})) = \mathrm{Tr}(\rho^{A_I})\ \forall \rho^{A_I} \in \mathcal{L}(\mathcal{H}^{A_I})$ (and with the subscript x indicating the classical input according to which Alice's instrument is chosen).

Deterministic transformations, such as unitary evolutions, or, more generally, CPTP maps, are a special case of quantum instruments. They correspond to instruments with only one outcome that occurs with certainty. Likewise, projective measurements, or, more generally, positive operator-valued measures (POVMs) are a special case of quantum instruments, which is recovered when the output space $\mathcal{H}^{A_O}$ is one-dimensional, and the output is just a probability.

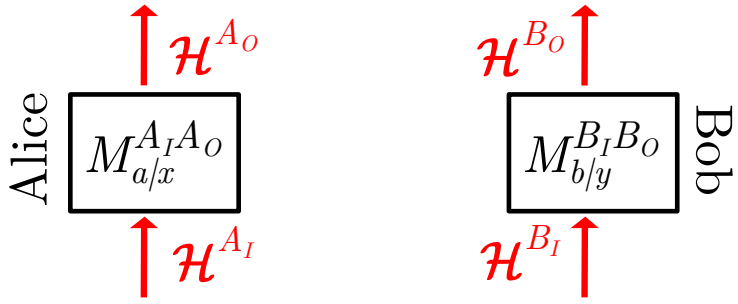

FIGURE 1.1: Alice and Bob perform local operations, which are compatible with quantum theory, but not embedded into any global causal order.

A convenient way to represent the parties' operations as matrices is via the Choi-Jamiołkowski (CJ) isomorphism [42, 43]. In order to define it[1], we choose for each Hilbert space $\mathcal{H}^X$ under consideration a fixed orthonormal, so-called "computational" basis $\{|i\rangle^X\}_i$. We then define the CJ representation of a linear map $\mathcal{M} : \mathcal{L}(\mathcal{H}^X) \to \mathcal{L}(\mathcal{H}^Y)$ as

$$M := [\mathcal{I} \otimes \mathcal{M}(|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|)] = \Big[\sum_{i,i'} |i\rangle\langle i'| \otimes \mathcal{M}(|i\rangle\langle i'|)\Big] \in \mathcal{L}(\mathcal{H}^X \otimes \mathcal{H}^Y), \tag{1.1}$$

where $\mathcal{I}$ is the identity channel, and $|\mathbb{1}\rangle\rangle := \sum_i |i\rangle^X \otimes |i\rangle^X \in \mathcal{H}^X \otimes \mathcal{H}^X$. From its CJ representation M it is easy to recover the map $\mathcal{M}$, using

$$\mathcal{M}(\rho) = \mathrm{Tr}_X[M \cdot (\rho^T \otimes \mathbb{1}^Y)], \tag{1.2}$$

where T denotes transposition in the computational basis, Tr_X denotes the partial trace over the system X, and $\mathbb{1}^Y$ denotes the identity operator in the space $\mathcal{L}(\mathcal{H}^Y)$ (in general, superscripts on operators, which may be omitted when clear enough, denote the system(s) they apply to).

When one deals with pure operations of the form $\mathcal{M}(\rho) = V\rho V^\dagger$, with $V : \mathcal{H}^X \to \mathcal{H}^Y$, it is often convenient to work on the level of Hilbert space vectors. In that case, the CJ representation is a rank-one matrix $M = |V\rangle\rangle\langle\langle V|$, with

$$|V\rangle\rangle := \mathbb{1} \otimes V\,|\mathbb{1}\rangle\rangle = \sum_i |i\rangle \otimes V\,|i\rangle \;\in \mathcal{H}^X \otimes \mathcal{H}^Y, \tag{1.3}$$

[1]Note that different versions of the CJ isomorphism exist in the literature. In particular, the transposition that appears in the "generalised Born rule" (1.5) is sometimes absorbed into the definition [4, 11, 15].

We refer to Eq. (1.3) as the *pure CJ representation* of V. The operator $V : \mathcal{H}^X \to \mathcal{H}^Y$ can be recovered from its pure CJ representation $|V\rangle\rangle \in \mathcal{H}^X \otimes \mathcal{H}^Y$ via

$$V = \sum_{i,j} \langle i, j\, |V\rangle\rangle\, |j\rangle\langle i|, \tag{1.4}$$

where $\{|j\rangle\}_j$ is the computational basis of $\mathcal{H}^Y$, and where $\langle i, j| = \langle i| \otimes \langle j|$.

A linear map $\mathcal{M} : \mathcal{L}(\mathcal{H}^X) \to \mathcal{L}(\mathcal{H}^Y)$ is completely positive if and only if its CJ representation is positive semidefinite, and it is trace preserving if and only if its CJ representation satisfies $\mathrm{Tr}_Y\, M = \mathbb{1}^X$.

The CJ isomorphism thus allows us to represent the CP maps performed by Alice and Bob as positive semidefinite matrices $M^{A_{IO}}_{a|x}, M^{B_{IO}}_{b|y}$, and the corresponding CPTP maps as positive semidefinite matrices $M^{A_{IO}}_x := \sum_a M^{A_{IO}}_{a|x}$, $M^{B_{IO}}_y := \sum_b M^{B_{IO}}_{b|y}$ that satisfy $\mathrm{Tr}_{A_O}\, M^{A_{IO}}_x = \mathbb{1}^{A_I}$ and $\mathrm{Tr}_{B_O}\, M^{B_{IO}}_y = \mathbb{1}^{B_I}$.

1.1.2 Process matrices

We now want to specify the most general correlations that Alice and Bob can establish in the scenario described above, where they both reside in closed laboratories and perform local quantum operations. As shown in Ref. [4], requiring compatibility with quantum mechanics locally and assuming the non-contextuality[2] of the probabilities imply that the probabilities $P(a, b|x, y)$ must be bilinear in the CP maps associated with the operations of A and B—or, equivalently, bilinear in their CJ representations. (Throughout this work, we will often refer to CP maps by their equivalent CJ representation and *vice versa*.) It follows that the correlations can be obtained as

$$P(a, b|x, y) = \mathrm{Tr}\left[(M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{b|y})^T \cdot W\right] \tag{1.5}$$

where Tr now denotes the full trace, and W is a Hermitian operator $\in \mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}})$ [4].

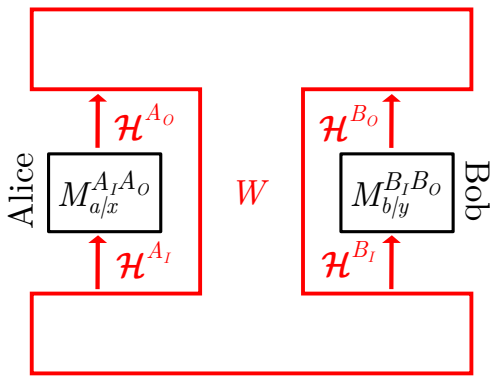

FIGURE 1.2: The process matrix relates the local operations of Alice and Bob through Eq. (1.5).

[2]That is, we assume that the probability $P(a, b|x, y)$ depends only on the CP maps $\mathcal{M}^A_{a|x}$ and $\mathcal{M}^B_{b|y}$, and not on the instruments $\{\mathcal{M}^A_{a|x}\}_a$ and $\{\mathcal{M}^B_{b|y}\}_b$ they belong to [4].

The framework also allows the parties to extend their local operations onto some extra incoming Hilbert spaces, and to share an additional, (possibly entangled) ancillary quantum state [4]. A Hermitian operator such that Eq. (1.5) yields valid (i.e., nonnegative and normalised) probabilities for any quantum instruments that Alice and Bob can perform, even when the parties share arbitrary ancillary states, is referred to as a (bipartite) *process matrix*.

Definition 1.1. A *bipartite process matrix* is a Hermitian matrix $W \in \mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}})$ such that for any extra incoming Hilbert spaces $\mathcal{H}^{A_{I'}}, \mathcal{H}^{B_{I'}}$ of arbitrary (finite) dimension, any quantum state $\rho \in \mathcal{L}(\mathcal{H}^{A_{I'}} \otimes \mathcal{H}^{B_{I'}})$, and any quantum instruments $\{M^{A_{II'O}}_{a|x}\}_a$ and $\{M^{B_{II'O}}_{b|y}\}_b$,

$$\begin{aligned} \operatorname{Tr}\left[(M^{A_{II'O}}_{a|x})^T \otimes (M^{B_{II'O}}_{b|y})^T \cdot W \otimes \rho\right] \geq 0 \quad \forall\, a, b \quad \text{and} \\ \sum_{a,b} \operatorname{Tr}\left[(M^{A_{II'O}}_{a|x})^T \otimes (M^{B_{II'O}}_{b|y})^T \cdot W \otimes \rho\right] = 1. \end{aligned} \tag{1.6}$$

It has been shown in Refs. [4, 11] that this condition is equivalent to W satisfying the following constraints, which thus characterise the set of bipartite process matrices:

$$W \geq 0, \quad W \in \mathcal{L}^{\{A,B\}}, \quad \text{and} \quad \operatorname{Tr} W = d_{A_O} d_{B_O}, \tag{1.7}$$

where $\mathcal{L}^{\{A,B\}}$ is a particular real vector space of Hermitian matrices in $\mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}})$ [4, 11]. To specify what this subspace is, we introduce the following "trace and replace" notation (used first in Ref. [11]):

$${}_X W := (\operatorname{Tr}_X W) \otimes \frac{\mathbb{1}^X}{d_X}, \tag{1.8}$$

with d_X the dimension of the Hilbert space of system X. We furthermore define

$${}_{[\sum_X \alpha_X X]} W := \sum_X \alpha_X \, {}_X W, \tag{1.9}$$

with $\alpha_X \in \{1, -1\}$ (and with ${}_1 W := W$).

The subspace $\mathcal{L}^{\{A,B\}}$ is then the subspace of Hermitian matrices W that satisfy the linear conditions

$$\begin{aligned} {}_{[1-A_O]B_{IO}} W &= 0 \\ {}_{[1-B_O]A_{IO}} W &= 0 \\ {}_{[1-A_O][1-B_O]} W &= 0 \end{aligned} \tag{1.10}$$

The first line in Eq. (1.10), for instance, means that tracing out Bob's systems (and replacing them by the identity) yields a matrix with the identity on Alice's output system A_O.

The condition $W \in \mathcal{L}^{\{A,B\}}$ follows from the fact that the probabilities should sum up to one (i.e., Eq. (1.5) should yield one whenever the local operations of Alice and Bob are CPTP maps). The positivity of W in Eq. (1.7) follows from the fact that the probabilities in Eq. (1.5) should be positive, even when allowing for shared input ancillas [4, 11].

A process matrix can be understood as a generalisation of a multipartite density matrix that allows also for the description of scenarios with signalling, and Eq. (1.5) can be understood as a "generalised Born rule". In the next subsection, we will illustrate this through two simple examples of scenarios that can be described by process matrices.

1.1.3 Examples

Consider for instance the case where the output spaces $\mathcal{H}^{A_O}$ and $\mathcal{H}^{B_O}$ of both parties are trivial ($d_{A_O} = d_{B_O} = 1$). As mentioned above, the local operations of the parties then reduce to generalised measurements, described by POVMs $\{E^{A_I}_{a|x}\}_a$ and $\{E^{B_I}_{b|y}\}_b$, whose CJ representations are $\{(E^{A_I}_{a|x})^T\}_a$ and $\{(E^{B_I}_{b|y})^T\}_b$. The validity condition (1.7) in that case simply characterises positive semidefinite matrices $\rho^{A_I B_I}$ with $\mathrm{Tr}(\rho^{A_I B_I}) = 1$ (in particular, the conditions (1.10) are trivially satisfied for one-dimensional output spaces). In other words, the process matrix reduces to a quantum state in $\mathcal{L}(\mathcal{H}^{A_I} \otimes \mathcal{H}^{B_I})$. We thus recover the standard Bell scenario where a bipartite state is shared between Alice and Bob, who perform measurements on their respective part (see Fig. 1.3). Indeed, Eq. (1.5) reduces to the standard Born rule, i.e. the correlations established between Alice and Bob are

$$P(a, b|x, y) = \mathrm{Tr}\left[E^{A_I}_{a|x} \otimes E^{B_I}_{b|y} \cdot \rho\right] \tag{1.11}$$

As is well known, the correlations in that scenario do not allow for signalling between the parties.

Let's consider another example where we do have signalling from Alice to Bob, but not vice versa. For instance, a process matrix can describe a scenario as in Fig. 1.4, where Alice prepares a state $\rho^{A_O}_x$ (depending on her input x), which is sent to Bob via a quantum channel (i.e., a CPTP map) $\mathcal{C} : \mathcal{L}(\mathcal{H}^{A_O}) \to \mathcal{L}(\mathcal{H}^{B_I})$, and where Bob then performs a measurement, which he chooses according to his input y, and which gives rise to a measurement outcome b.

In that scenario we just described, Alice has a trivial, one-dimensional input space, since she just performs a state preparation (with the CJ representation $\rho^{A_O}_x$), and Bob has a trivial output space, since he performs a POVM (with the CJ representation $\{(E^{B_I}_{b|y})^T\}_b$). The process matrix in this case is simply the CJ representation of the channel $C^{A_O B_I}$ connecting the two parties, which clearly satisfies the validity conditions (1.7) (as it is CPTP, i.e., it is positive semidefinite, and satisfies the trace-preserving condition $\mathrm{Tr}_{B_I} C^{A_O B_I} = \mathbb{1}^{A_O}$). The correlations as per Eq. (1.5) are

$$P(b|x, y) = \mathrm{Tr}\left[(\rho^{A_O}_x)^T \otimes E^{B_I}_{b|y} \cdot C^{A_O B_I}\right] \tag{1.12}$$

Using the inverse of the CJ isomorphism (Eq. (1.2)), it can easily be seen that this is indeed the same as $P(b|x, y) = \mathrm{Tr}[\mathcal{C}(\rho^{A_O}_x) \cdot E^{B_I}_{b|y}]$ (which would be the "usual way" of calculating the correlations arising in the scenario we consider here).

These two examples illustrate that the process matrix formalism can be seen as a framework that captures both non-signalling scenarios (such as quantum states) and signalling scenarios (such as quantum channels) in a unified way.

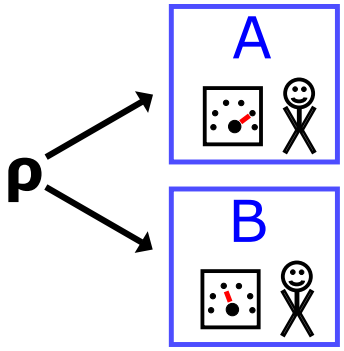

FIGURE 1.3: A bipartite quantum state. In this scenario, the process matrix is the quantum state ρ, and the local operations of Alice and Bob are POVMs. Alice and Bob cannot signal to each other.

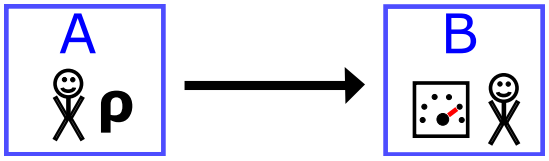

FIGURE 1.4: A quantum channel from Alice to Bob. Here, the process matrix describes the channel, the local operation of Alice is a state preparation and that of Bob is a POVM. In this scenario, Alice can signal to Bob.

1.2 Causal (non)separability

The two examples of process matrices mentioned in the previous subsection are compatible with a definite causal structure, in the sense that they allow for signalling in at most one direction. We will now consider the general question, whether the situation described by a process matrix can be embedded in a well-defined causal structure, with a fixed causal order between the events happening in each party's laboratory, or not.

1.2.1 Process matrices compatible with a definite causal order

A process matrix is said to be "compatible with (the causal order) $A \prec B$" (sometimes abbreviated to just "$A \prec B$", e.g., in superscripts) if all the correlations it generates are compatible with a causal order where A acts before B, which is to be understood operationally: such a process matrix $W^{A\prec B}$ does not allow for any signalling from B to A. More precisely, whatever the CP and CPTP maps $M^{A_{IO}}_{a|x}, M^{B_{IO}}_{y^{(\prime)}}$ of A and B, the resulting correlations respect the no-signalling condition $P(a|x,y) = P(a|x,y')$, or $\mathrm{Tr}[(M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{y})^T \cdot W^{A\prec B}] = \mathrm{Tr}[(M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{y'})^T \cdot W^{A\prec B}]$ according to Eq. (1.5). This constrains $W^{A\prec B}$ to be in a linear subspace $\mathcal{L}^{A\prec B}$ of $\mathcal{L}^{\{A,B\}}$, formed

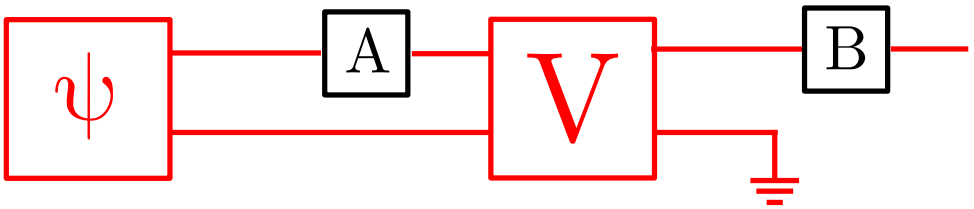

FIGURE 1.5: A channel with memory with A before B. The process matrix describing such a circuit is compatible with $A \prec B$. Vice versa, any process matrix compatible with $A \prec B$ can be realised in terms of a channel with memory, and the initial state $|\psi\rangle$, as well as the isometry V, can be constructed from the process matrix.

by the Hermitian operators in $\mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}})$ that satisfy the linear conditions [4, 11]

$$\begin{aligned} {}_{[1-B_O]}W &= 0 \\ {}_{[1-A_O]B_{IO}}W &= 0 \end{aligned} \tag{1.13}$$

To see that $\mathcal{L}^{A\prec B}$ is indeed a linear subspace of $\mathcal{L}^{\{A,B\}}$, note that the condition ${}_{[1-B_O]}W = 0$ implies the second and third lines in Eq. (1.10).

The above example of a bipartite quantum state, for instance, is compatible with both causal orders $A \prec B$ and $B \prec A$, and the second example of a quantum channel from Alice to Bob is compatible with the causal order $A \prec B$.

More generally, it can be shown [44, 45] that the set of process matrices that do not allow for signalling from B to A corresponds precisely to the "standard quantum circuits" with A acting before B, that is, the circuits which have the form shown in Fig. 1.5. In such a circuit, A receives one part of a bipartite state and acts on it, before its output, as well as the ancillary system, are sent through an isometric channel V to the input of B and an additional ancilla, which is discarded. Such a circuit is also called a *quantum comb* [45] or a *channel with memory* [46]. Its process matrix is compatible with $A \prec B$. Vice versa, every process matrix compatible with $A \prec B$ has a realisation in terms of a channel with memory. The proof is constructive, that is, the initial state $|\Psi\rangle$ and the isometry V can be calculated explicitly from the process matrix [44, 45] (cf. also Chapter 3).

Likewise, process matrices that do not allow signalling from A to B are said to be compatible with the causal order $B \prec A$, and will typically be denoted $W^{B\prec A} \in \mathcal{L}^{B\prec A}$. They satisfy the constraints analogous to Eq. (1.13) (with A and B interchanged), and correspond to the channels with memory from B to A.

1.2.2 Causally (non)separable process matrices

One can also conceive of situations where the causal order is not fixed to be the same for all experimental runs, but where there is instead a probabilistic mixture of the two possibilities. Such a scenario is described by a convex combination of process matrices compatible with $A \prec B$

and $B \prec A$, respectively. Process matrices of this form remain compatible with an underlying causal framework and are the subject of the following definition, first introduced by Oreshkov, Costa and Brukner [4]:

Definition 1.2 (Bipartite causal (non)separability [4])**.** A bipartite process matrix W is said to be *causally separable* if and only if it can be written as a convex combination

$$W \;=\; q\, W^{A\prec B} \;+\; (1{-}q)\, W^{B\prec A}\,, \tag{1.14}$$

with $q \in [0,1]$ and where $W^{A\prec B}$ and $W^{B\prec A}$ are two process matrices compatible with the causal orders $A \prec B$ and $B \prec A$, respectively.

A process matrix that cannot be decomposed as above is said to be *causally nonseparable.*

Interestingly, causally nonseparable process matrices indeed exist. The first example of such a process matrix was introduced in Ref. [4]:

$$W^{\mathrm{ocb}} = \frac{1}{4}\left[\mathbb{1}^{A_I A_O B_I B_O} + \frac{1}{\sqrt{2}}(\mathbb{1}^{A_I}\hat{\mathrm{Z}}^{A_O}\hat{\mathrm{Z}}^{B_I}\mathbb{1}^{B_O} + \hat{\mathrm{Z}}^{A_I}\mathbb{1}^{A_O}\hat{\mathrm{X}}^{B_I}\hat{\mathrm{Z}}^{B_O})\right] \tag{1.15}$$

Here, $\hat{\mathrm{X}}, \hat{\mathrm{Y}}, \hat{\mathrm{Z}}$ denote the Pauli matrices, $\mathbb{1}$ denotes the 2×2 identity matrix and tensor products between all matrices are implicit. It is straightforward to check that W^{ocb} indeed satisfies the constraints (1.7), and is thus a valid process matrix.

The existence of such causally nonseparable process matrices means that, among the possible scenarios that can occur under the assumptions that we made (i.e., parties in local laboratories that generate valid correlations by performing operations in accordance with quantum theory, but that are not embedded into any causal order *a priori*), some are indeed incompatible with any definite causal order between the parties. What do these scenarios correspond to in practice? This is, in fact, an open question. Contrary to the case of (bipartite) causally separable process matrices (which most generally describe probabilistic mixtures of fixed-order quantum circuits, as we discussed just above), it is not clear what physical interpretation—if any—can be given to bipartite causally nonseparable process matrices, such as W^{ocb}. An example of a causally nonseparable process matrix with a physical realisation exists, however, in the tripartite case, namely the quantum switch, which will be discussed further below.

1.2.3 Witnesses of causal nonseparability

How can we prove that a given process matrix (such as, for instance, W^{ocb} above) is indeed causally nonseparable?

For W^{ocb} in particular, it was shown in Ref. [4] that this process matrix allows for the generation of *noncausal correlations*, which is impossible for causally separable process matrices and which therefore proves its causal nonseparability. We will explain the concept of noncausal correlations in detail in the next subsection.

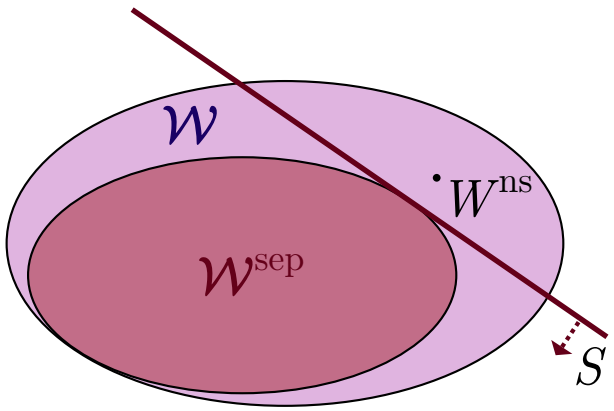

FIGURE 1.6: The separating hyperplane theorem guarantees that for any causally nonseparable process matrix, there exists a witness certifying its causal nonseparability.

Another approach, first introduced in Ref. [11] and further studied in Ref. [13], is to construct a *witness of causal nonseparability*—or "*causal witness*" for short, a Hermitian operator that certifies the causal nonseparability of the process matrix under consideration, and that can be accessed experimentally through suitable operations performed by the parties.

When the normalisation constraint $\operatorname{Tr} W = d_{A_O} d_{B_O}$ in Eq. (1.7) is ignored, the various sets of (non-normalised) process matrices we considered above can be characterised as *convex cones*. The set of (non-normalised) valid process matrices is the convex cone

$$\mathcal{W} := \mathcal{P} \cap \mathcal{L}^{\{A,B\}} \tag{1.16}$$

where we denote by $\mathcal{P}$ the convex cone of positive semidefinite matrices. The convex cones of (non-normalised) process matrices compatible with the fixed orders $A \prec B$ and $B \prec A$ are $\mathcal{W}^{A\prec B} := \mathcal{P} \cap \mathcal{L}^{A\prec B}$ and $\mathcal{W}^{B\prec A} := \mathcal{P} \cap \mathcal{L}^{B\prec A}$, respectively, and the convex cone of (non-normalised) causally separable process matrices is given by their Minkowski sum

$$\mathcal{W}^{\text{sep}} := \mathcal{W}^{A\prec B} + \mathcal{W}^{B\prec A} = \{W^{A\prec B} + W^{B\prec A} | W^{A\prec B} \in \mathcal{W}^{A\prec B} \text{ and } W^{B\prec A} \in \mathcal{W}^{B\prec A}\}. \tag{1.17}$$

A causal witness is then defined as a Hermitian operator S such that

$$\operatorname{Tr}[S \cdot W^{\text{sep}}] \geq 0 \tag{1.18}$$

for all $W^{\text{sep}} \in \mathcal{W}^{\text{sep}}$. The set of causal witnesses is precisely the dual cone of $\mathcal{W}^{\text{sep}}$, $\mathcal{S} = (\mathcal{W}^{\text{sep}})^*$ [11, 13].

For any causally nonseparable W^{ns}, it is known (as a consequence from the *separating hyperplane theorem*, cf. Fig. 1.6) that there exists a causal witness S such that $\operatorname{Tr}[S \cdot W^{\text{ns}}] < 0$ [11, 13].

The causal nonseparability of W^{ocb}, for instance, can be certified by the witness[3]

$$S^{\text{ocb}} = \frac{1}{4}\left[\mathbb{1}^{A_I A_O B_I B_O} - (\mathbb{1}^{A_I}\hat{\mathrm{Z}}^{A_O}\hat{\mathrm{Z}}^{B_I}\mathbb{1}^{B_O} + \hat{\mathrm{Z}}^{A_I}\mathbb{1}^{A_O}\hat{\mathrm{X}}^{B_I}\hat{\mathrm{Z}}^{B_O})\right], \tag{1.19}$$

for which one finds that $\mathrm{Tr}[S^{\text{ocb}} \cdot W^{\text{ocb}}] = 1 - \sqrt{2} < 0$, proving that W^{ocb} is causally nonseparable.

In general, the construction of causal witnesses (or of explicit decompositions of causally separable process matrices) can be efficiently performed with semidefinite programming, as first described in Ref. [11]; we will follow here the slightly different approach of [13]. The question of whether a given W is causally separable can be reformulated as the optimisation problem of how much white noise can be added to a process matrix before it becomes causally separable. Let $\mathbb{1}^\circ = \mathbb{1}^{A_{IO}B_{IO}}/d_{A_I}d_{B_I}$ be the "white noise" process matrix (which corresponds to a scenario where the parties receive fully mixed states $\mathbb{1}^{A_I}/d_{A_I}$ and $\mathbb{1}^{B_I}/d_{B_I}$ at their inputs, and is causally separable), and consider the noisy process matrix

$$W(r) = \frac{1}{1+r}(W + r\mathbb{1}^\circ). \tag{1.20}$$

Since the normalisation is irrelevant to membership of $\mathcal{W}^{\text{sep}}$, determining whether W is causally separable can be thus phrased as the SDP optimisation problem

$$\begin{aligned} &\min\ r \\ \text{s.t.}\quad &W + r\mathbb{1}^\circ \in \mathcal{W}^{\text{sep}}, \end{aligned} \tag{1.21}$$

which can be efficiently solved using standard software by writing $\mathcal{W}^{\text{sep}}$ in terms of SDP constraints (see [13] and Chapter 2 for further details). The solution to this problem, r^*, gives the *random robustness* $\max(r^*, 0)$ of W, and a value $r^* > 0$ implies that W is causally nonseparable [11, 13].

Eq. (1.21) is known as the primal problem, and is related to the dual problem

$$\begin{aligned} &\min\ \mathrm{Tr}[S \cdot W] \\ \text{s.t.}\quad &S \in \mathcal{S} \quad \text{and} \quad \mathrm{Tr}[S \cdot \mathbb{1}^\circ] = 1, \end{aligned} \tag{1.22}$$

defined over the dual cone $\mathcal{S}$ of $\mathcal{W}^{\text{sep}}$ [11, 13]. The optimal solution S^* is a witness of the causal nonseparability of W whenever $\mathrm{Tr}[S^* \cdot W] < 0$. The Strong Duality Theorem for SDP problems moreover relates these two problems, stating that their solutions satisfy

$$r^* = -\mathrm{Tr}[S^* \cdot W]. \tag{1.23}$$

This implies in particular that the witness S^* thus obtained is optimal when W is subject to white noise, in the sense that it witnesses the causal nonseparability of all noisy process matrices $W(r)$ with r sufficiently small ($r < r^*$) so as for $W(r)$ to remain causally nonseparable.

[3] The fact that S^{ocb} is indeed a causal witness follows from ${}_{A_O}S^{\text{ocb}} \geq 0$ and ${}_{B_O}S^{\text{ocb}} \geq 0$, which imply that $\mathrm{Tr}[S^{\text{ocb}} \cdot W^{\text{sep}}] \geq 0$ for all $W^{\text{sep}} \in \mathcal{W}^{\text{sep}}$ [11].

In a physical realisation of a given process matrix W, one can "measure" a causal witness S [11, 13], by decomposing it as a linear combination of the form

$$S = \sum_{x,y,a,b} \gamma_{x,y,a,b} \cdot (M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{b|y})^T \tag{1.24}$$

with suitable CP maps $M^{A_{IO}}_{a|x}$, and $M^{B_{IO}}_{b|y}$, and coefficients $\gamma_{x,y,a,b}$. Having each party implement the respective instruments then allows to experimentally determine the quantity

$$\mathrm{Tr}[S \cdot W] = \sum_{x,y,a,b} \gamma_{x,y,a,b} \cdot \mathrm{Tr}[(M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{b|y})^T \cdot W] = \sum_{x,y,a,b} \gamma_{x,y,a,b} \cdot P(a,b|x,y). \tag{1.25}$$

by appropriately combining the measurement statistics.

1.3 Noncausal correlations

The above-described measurement of a causal witness is a device-dependent test, in the sense that in order to realise it, the parties need to perform specific operations that are described by quantum theory. The incompatibility of certain scenarios with an underlying causal order can also be certified in a device-independent way, through the violation of so-called causal inequalities. These are constraints that the correlations established between parties need to satisfy whenever there is a definite background causal structure.

Before outlining how (non)causal correlations can be characterised in general, let us consider an illustrative example. Assume, as above, that one has two parties A(lice) and B(ob), receiving binary inputs x and y, respectively, and generating binary outputs a and b, respectively. The two parties are supposed to maximise the success probability in a *guess your neighbour's input (GYNI) game* [12], whose rules are as follows: the task of A is to guess the input of B, i.e. she needs to output y, while B needs to guess the input of A, i.e. he needs to output x. For sake of simplicity, we assume that the inputs x and y are uniform, i.e. $p(x) = p(y) = \frac{1}{2}$. A definite causal structure between the parties implies an upper bound on the success probability. If, for instance, A acts in the past of B, she can signal her input x to B and B can successfully output x. However, A cannot receive any signal from B and therefore cannot know the input of B in any way. Thus, she has to make a random guess, which will be successful with probability $\frac{1}{2}$. The overall success probability for the parties to win the game is thus upper-bounded by $\frac{1}{2}$. The same is true if B is in the past of A or for a probabilistic mixture of the two cases. Therefore, if A and B manage to establish correlations between them that violate this bound of $\frac{1}{2}$, it follows that they cannot be embedded into a definite causal structure.
Formally, (non)causal correlations can be defined in the following way [12, 15]:
If B is in the future of A, the choice of his input y cannot influence what output A obtains, i.e.

the no-signalling condition

$$P(a|x,y) := \sum_b P(a,b|x,y) = P(a|x,y') \quad \forall x,y,y',a \tag{1.26}$$

is satisfied. A probability distribution satisfying that constraint is denoted by $P^{A\prec B}(a,b|x,y)$. Likewise, the analogous no-signalling constraint from A to B reads

$$P(b|x,y) := \sum_a P(a,b|x,y) = P(b|x',y) \quad \forall x,x',y,b \tag{1.27}$$

A probability distribution satisfying that constraint is denoted by $P^{B\prec A}(a,b|x,y)$.
In general, one can also consider a convex mixture

$$P(a,b|x,y) = q \cdot P^{A\prec B}(a,b|x,y) + (1-q) \cdot P^{B\prec A}(a,b|x,y) \qquad q \in [0,1], \tag{1.28}$$

which can arise in a situation where the causal order is still definite, but determined probabilistically in each run of the experiment. Correlations of that form are called *causal*, while correlations that cannot be decomposed in this way are called *noncausal*.

When characterising (non)causal correlations, one can draw upon concepts that are known from the study of Bell (non)locality. Any correlation $P(a,b|x,y)$ can be seen as a point in a high-dimensional vector space, with each coordinate given by the probability for a particular value of the inputs and outputs. In this geometrical picture, the set of causal correlations forms a *convex polytope*, i.e. a convex geometrical object with a finite number of vertices. The facets of this polytope define *causal inequalities*, that is, linear constraints which are respected by all causal correlations, and the violation of which thus indicates that a correlation is noncausal.

The analysis of (non)causal correlations is thus mathematically analogous to the analysis of Bell (non)local correlations, where the local correlations form the *local polytope*, the facets of which define *Bell inequalities*. Conceptually the two notions are of course quite different. In the case of Bell (non)locality, the correlations under consideration (both inside and outside of the local polytope) are non-signalling, while in the case of (non)causal correlations, the correlations of interest are signalling (with (non)causal correlations being those for which the signalling is (in)compatible with an underlying, definite causal structure).

The polytope of bipartite causal correlations for binary inputs and outputs has been fully characterised in Ref. [12]. In this scenario, there are two nontrivial equivalence classes of facets, one of which is given by[4]

$$\frac{1}{4} \sum_{x,y,a,b} \delta_{a,y}\delta_{b,x} P(a,b|x,y) \leq \frac{1}{2}. \tag{1.29}$$

[4]That is, Eq. (1.29) is one representative of the equivalence class, and the other representatives can be obtained by relabeling inputs and outputs.

where $\delta_{i,j}$ is the Kronecker delta. For uniform inputs and outputs, this inequality can be written as $P(a = y, b = x) \leq \frac{1}{2}$, and its interpretation is precisely the GYNI game we discussed above. The other equivalence class of facets is given by

$$\frac{1}{4} \sum_{x,y,a,b} \delta_{x(a\oplus y),0} \delta_{y(b\oplus x),0} P(a, b|x, y) \leq \frac{3}{4}. \tag{1.30}$$

where $\oplus$ denotes addition modulo 2. This causal inequality corresponds to a *lazy guess your neighbour's input (LGYNI) game*, where A and B need to guess each other's input only when their respective input is 1, while their output does not matter if their input is 0. This is in fact an interesting difference with Bell nonlocality, where one cannot have nontrivial scenarios in that "lazy" case with binary inputs and a trivial output for one of the inputs.

Interestingly, some causally nonseparable quantum processes can generate correlations which violate causal inequalities and are thus incompatible with a causal order even in that stronger, device-independent sense. This is the case for W^{ocb}, as shown in Ref. [4], where a task similar to the GYNI game above was defined, and it was demonstrated that there exists a choice of instruments such that the correlations obtained through the generalised Born rule (1.5) violate the bound that is imposed by a definite causal structure. The violation of causal inequalities with process matrix correlations in the simplest bipartite case was then further studied in Ref. [12].

1.4 Towards multipartite scenarios

In this section, we outline how some of the concepts introduced above have previously been generalised to scenarios with more than two parties, in order to lay the basis for the multipartite considerations in the following chapters, and we point out some of the complexities that arise in multipartite situations.

1.4.1 Multipartite process matrices

The process matrix framework itself generalises rather easily to the multipartite case [11, 15, 47].

Let us first introduce some generalised notations. We shall consider N parties denoted by A_k for $k \in \{1, \ldots, N\} \coloneqq \mathcal{N}$, with corresponding inputs and outputs denoted by x_k, and a_k, respectively. We define the input and output vectors $\vec{x} \coloneqq (x_1, \ldots, x_N)$ and $\vec{a} \coloneqq (a_1, \ldots, a_N)$. The "incoming" and "outgoing" Hilbert spaces for each party are denoted by $\mathcal{H}^{A_I^k}, \mathcal{H}^{A_O^k}$ (of dimensions $d_{A_I^k}, d_{A_O^k}$, respectively). We also define $\mathcal{H}^{A_{IO}^k} \coloneqq \mathcal{H}^{A_I^k} \otimes \mathcal{H}^{A_O^k}$, and $d_{A_{IO}^k} \coloneqq d_{A_I^k} d_{A_O^k}$.

For a subset $\mathcal{K} \subseteq \mathcal{N}$ of parties, we will denote by $\vec{x}_\mathcal{K}$ and $\vec{a}_\mathcal{K}$ the vectors of inputs and outputs restricted to the parties in $\mathcal{K}$, and use shorthand notations like $\mathcal{H}^{A_{IO}^\mathcal{K}} \coloneqq \bigotimes_{k\in\mathcal{K}} \mathcal{H}^{A_{IO}^k}$ ($= \mathbb{R}$ if $\mathcal{K} = \emptyset$), $\mathbb{1}^\mathcal{K} \coloneqq \bigotimes_{k\in\mathcal{K}} \mathbb{1}^{A_{IO}^k} = \mathbb{1}^{A_{IO}^\mathcal{K}}$, and $\text{Tr}_\mathcal{K}$ for the trace over all (incoming and outgoing) systems of the parties in $\mathcal{K}$—i.e., $\text{Tr}_{A_{IO}^\mathcal{K}}$ or $\text{Tr}_{A_{II'O}^\mathcal{K}}$, as appropriate (see below). For notational

simplicity, we shall sometimes identify the parties' names with their labels, and singletons of parties (e.g., $\{A_k\}$) with the parties themselves (e.g., A_k) or the corresponding label, so that $\mathcal{N} = \{1, \ldots, N\} \equiv \{A_1, \ldots, A_N\}$, $\mathcal{N}\backslash\{A_k\} \equiv \mathcal{N}\backslash k$, $\mathrm{Tr}_{\{A_k\}} \equiv \mathrm{Tr}_k$, etc.

The CP maps corresponding to the parties' operations are then denoted by $M^{A_{IO}^k}_{a_k|x_k}$, the corresponding CPTP maps $M^{A_{IO}^k}_{x_k} := \sum_{a_k} M^{A_{IO}^k}_{a_k|x_k}$, and the overall process is represented by a Hermitian operator $W \in \mathcal{L}(\mathcal{H}^{A_{IO}^{\mathcal{N}}})$. The resulting correlations are obtained through a generalised Born rule as before:

$$P(\vec{a}|\vec{x}) = \mathrm{Tr}\left[(M^{A_{IO}^1}_{a_1|x_1})^T \otimes \cdots \otimes (M^{A_{IO}^N}_{a_N|x_N})^T \cdot W\right]. \tag{1.31}$$

As in the bipartite case, the parties may also share some ancillary state ρ in some extra incoming spaces $\mathcal{L}(\mathcal{H}^{A_{I'}^1} \otimes \cdots \otimes \mathcal{H}^{A_{I'}^N}) = \mathcal{L}(\mathcal{H}^{A_{I'}^{\mathcal{N}}})$, and extend their local operations to act on these spaces as well. Requiring again the nonnegativity and normalisation of all obtainable probabilities, including for arbitrary extensions $W \otimes \rho$ of W, imposes validity constraints on W. In the general multipartite case, they read

$$W \geq 0, \quad W \in \mathcal{L}^{\mathcal{N}}, \quad \text{and} \quad \mathrm{Tr}\, W = \prod_{k \in \mathcal{N}} d_{A_O^k} \tag{1.32}$$

for some particular linear subspace $\mathcal{L}^{\mathcal{N}}$ of the space of Hermitian operators $\in \mathcal{L}(\mathcal{H}^{A_{IO}^{\mathcal{N}}})$; see Sec. 2.2.1 and Refs. [4, 11].

The no-signalling constraints can readily be generalised to the N-partite case, allowing the notion of compatibility with a fixed causal order to be extended accordingly. For instance, a process matrix is said to be compatible with the fixed causal order $A_1 \prec A_2 \prec \cdots \prec A_N$ if no party or group of parties can signal to other parties in their causal "past" (as defined by the specified causal order)—which translates into the constraint that $P(a_1, \ldots, a_k|\vec{x}) = P(a_1, \ldots, a_k|x_1, \ldots, x_k)$ for all $k = 1, \ldots, N-1$. As before, this constrains such a process matrix $W^{A_1 \prec \cdots \prec A_N}$ to be in a linear subspace $\mathcal{L}^{A_1 \prec \cdots \prec A_N}$ of $\mathcal{L}^{\mathcal{N}}$; see Sec. 2.2.1 for an explicit characterisation of $\mathcal{L}^{A_1 \prec \cdots \prec A_N}$.

1.4.2 Multipartite causal (non)separability and (non)causal correlations

The generalisation of the concepts of causal (non)separability and of (non)causal correlations to multiple parties is less straightforward than the extension of the framework itself. In particular, one of the complexities that arise in the multipartite case is the possiblity of *dynamical* causal orders—that is, the causal order of parties in the future can depend on operations of parties in the past [14, 15, 47] (see Fig. 1.7). Here, we show how (non)causal correlations can be defined and characterised in the multipartite case. The multipartite generalisation of the concept of causal (non)separability will be detailed in Chapter 2.

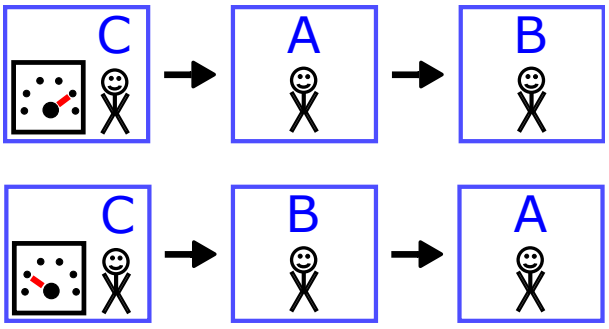

FIGURE 1.7: In multipartite scenarios, dynamical orders can arise—that is, the causal order of parties in the future can depend on events in the past. For instance, the causal order of Alice and Bob could depend on Charlie's outcome.

In order to generalise Eq. (1.28) to multipartite scenarios, one could, in the first place, simply consider a convex combination of correlations compatible with different fixed causal orders. However, this does not include scenarios with dynamical causal orders, and is therefore too restrictive to capture all correlations that should be considered compatible with a well-defined causal order. The problem of how to define multipartite causal correlations while allowing for dynamical causal orders was considered in Refs. [14, 15], which came to the conclusion that an appropriate definition is via an iterative "unraveling argument": A multipartite correlation is causal if it is possible to identify, up to some probability, a party that acts first, and such that, for any behaviour of this first party, the conditional correlations shared by the remaining parties are again causal. The formal definition of N-partite causal correlations given in Ref. [14] (which is equivalent to that first introduced in Ref. [15]) is as follows:

Definition 1.3 (N-partite causal correlations)**.** For $N = 1$, any correlation $P(a_1|x_1)$ is causal. For $N \geq 2$, an N-partite correlation $P(\vec{a}|\vec{x})$ is said to be *causal* if and only if it can be decomposed as

$$P(\vec{a}|\vec{x}) = \sum_{k\in\mathcal{N}} q_k \, P_k(a_k|x_k) \, P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k}), \tag{1.33}$$

with $q_k \geq 0$, $\sum_k q_k = 1$, where (for each k) $P_k(a_k|x_k)$ is a single-partite (and hence causal) correlation and (for each k, x_k, a_k) $P_{k,x_k,a_k}(\vec{a}_{\mathcal{N}\setminus k}|\vec{x}_{\mathcal{N}\setminus k})$ is a causal $(N-1)$-partite correlation.

In Ref. [14], it was shown that N-partite causal correlations also form a polytope, with its facets defining N-partite causal inequalities.

The simplest nontrivial multipartite situation is the "lazy tripartite" scenario in which three parties A, B and C receive binary inputs x, y and z, respectively, and generate outputs a, b and c, respectively, which are 0 when the respective input is 0, while they can be either 0 or 1 when the respective input is 1. The polytope for this case was characterised in Ref. [14], and it was demonstrated that the corresponding causal inequalities can be violated with process matrix correlations. A simple example of such a tripartite causal inequality is

$$\frac{1}{4} \sum_{x,y,a,b} \delta_{x(a\oplus y),0}\delta_{y(b\oplus x),0} P(a, b, c = 0|x, y, z = 0) \leq \frac{3}{4}. \tag{1.34}$$

which is a conditional version of the LGYNI inequality (1.30), where A and B are supposed to guess each other's input when their respective input is 1, and conditioned on the input 0 of C. Writing out Eq. (1.34) more explicitly, we obtain

$$\frac{1}{4}\left[P(0,0,0|0,0,0)+P(0,0,0|0,1,0)+P(0,0,0|1,0,0)+P(1,1,0|1,1,0)\right]\leq\frac{3}{4} \tag{1.35}$$

With $P(0,0,0|0,0,0)=1$, $P(0,0,0|0,1,0)=1-P(0,1,0|0,1,0)$ and $P(0,0,0|1,0,0)=1-P(1,0,0|1,0,0)$, this can then be simplified to

$$P(1,0,0|1,0,0)+P(0,1,0|0,1,0)-P(1,1,0|1,1,0)\geq 0. \tag{1.36}$$

1.5 A tripartite process: The quantum switch

A lot of studies of the multipartite case have focused on a restricted tripartite situation in which one party has no (or, equivalently, a trivial) outgoing system. This particular scenario has been studied extensively because of its relevance for a practical protocol where the causal order between two parties A and B, which perform some unitary operations U_A and U_B on a *target qubit* initialised in a state $|\psi\rangle^t$, is controlled by another (two-dimensional) quantum system. If this control qubit is initialised in the state $|0\rangle^c$, the operation U_A is applied before U_B, while for a control qubit in the state $|1\rangle^c$, U_B is applied before U_A. If the control qubit is initialised in a superposition state $|+\rangle^c=\frac{1}{\sqrt{2}}(|0\rangle^c+|1\rangle^c)$, the overall transformation on the joint state of the target and control systems is thus

$$|\psi\rangle^t\otimes|+\rangle^c\rightarrow\frac{1}{\sqrt{2}}(U_BU_A\,|\psi\rangle^t\otimes|0\rangle^c+U_AU_B\,|\psi\rangle^t\otimes|1\rangle^c)\,, \tag{1.37}$$

i.e., the unitaries are applied in a "superposition of orders". The output state is then sent to a third party C (Charlie) who can measure the control qubit, and possibly also the target qubit. The protocol just described can be generalised to the case where A and B's operations are general quantum instruments instead of unitaries. This so-called *quantum switch* constitutes a new resource for quantum computation that goes beyond causally ordered quantum circuits [3]. It has attracted particular interest as a consequence of being readily implementable, and can for example be implemented in an interferometric setup using two polarizing beam splitters, as shown schematically in Fig. 1.8. The polarization state of the input photon serves as the control qubit, while some internal degree of freedom of the photon, such as for instance its orbital angular momentum, serves as the target qubit. The two polarizing beam splitters transmit or reflect the photon depending on its polarization state such that a horizontally polarized photon goes first to the party A, then to the party B, while a vertically polarized photon, goes first to B, then to A. If the input photon is in a superposition of vertical and horizontal polarization, the orders of the operations are superposed. Indeed, several laboratory implementations along those lines have been experimentally realised [16–22].

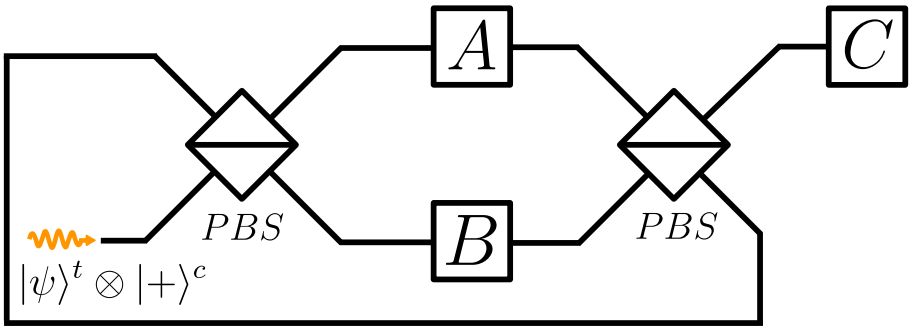

FIGURE 1.8: Optical implementation of the quantum switch with polarising beam splitters (PBS).

The quantum switch can naturally be described in the process matrix formalism [11, 15] where it corresponds to the following tripartite process matrix for parties A (Alice), B (Bob) and C (Charlie):

$$W^{\text{switch}} = |w\rangle\langle w| \tag{1.38}$$

where $|w\rangle$ is the "process vector"

$$|w\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{C_I^c} |\psi\rangle^{A_I} |\mathbb{1}\rangle\!\rangle^{A_O B_I} |\mathbb{1}\rangle\!\rangle^{B_O C_I^t} + |1\rangle^{C_I^c} |\psi\rangle^{B_I} |\mathbb{1}\rangle\!\rangle^{B_O A_I} |\mathbb{1}\rangle\!\rangle^{A_O C_I^t} \right) \tag{1.39}$$

Here, $|\psi\rangle$ is the initial state of the target qubit, $|\mathbb{1}\rangle\!\rangle^{A_O B_I} = \sum_{j=0,1} |j\rangle^{A_O} |j\rangle^{B_I}$ is the pure CJ representation of an identity qubit channel from A to B, $|\mathbb{1}\rangle\!\rangle^{B_O A_I} = \sum_{j=0,1} |j\rangle^{B_O} |j\rangle^{A_I}$ is the pure CJ representation of an identity qubit channel from B to A, and tensor products are implicit. The party C has no outgoing space, and its incoming space is $\mathcal{H}^{C_I} = \mathcal{H}^{C_I^t} \otimes \mathcal{H}^{C_I^c}$, as it can measure both the control and the target qubits.

The expression (1.39) is indeed quite intuitive. The process vector $|\psi\rangle^{A_I} |\mathbb{1}\rangle\!\rangle^{A_O B_I} |\mathbb{1}\rangle\!\rangle^{B_O C_I^t}$ corresponds to a quantum circuit with identity channels, where Alice acts on the state $|\psi\rangle$ before Bob. Similarly for the process vector $|\psi\rangle^{B_I} |\mathbb{1}\rangle\!\rangle^{B_O A_I} |\mathbb{1}\rangle\!\rangle^{A_O C_I^t}$, which describes a circuit in which Bob acts on $|\psi\rangle$ before Alice. The overall process vector $|w\rangle$ is a controlled superposition of the two terms, with the control qubit going directly to Charlie.

In this tripartite process matrix, Charlie indeed has no outgoing system and therefore cannot signal to the other parties. The situation is thus relatively similar to the bipartite case, since the only relevant causal orders are those where Charlie acts last, i.e., $A \prec B \prec C$ and $B \prec A \prec C$. This observation led Araújo *et al.* to propose the following definition of causal separability (as an initial, "1-step" generalisation of Definition 1.2) for this particular scenario:

Definition 1.4 (Araújo *et al.*'s causal separability [11])**.** In a tripartite scenario where party C has no outgoing system, a process matrix W is said to be *causally separable* if and only if it can be written as a convex combination

$$W \;=\; q\, W^{A\prec B\prec C} \;+\; (1-q)\, W^{B\prec A\prec C}\,, \tag{1.40}$$

with $q \in [0,1]$ and where $W^{A \prec B \prec C}$ and $W^{B \prec A \prec C}$ are two process matrices compatible with the causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively.

The process matrix W^{switch} describing the quantum switch is indeed causally nonseparable as per Definition 1.4. This follows straightforwardly from the fact that it is a rank-one projector [11, 15]. One can also consider the case where Charlie measures only the control qubit, and discards the target system. The process matrix corresponding to this situation is $\text{Tr}_{C_I^t} W^{\text{switch}}$, and it is also causally nonseparable, as shown in [11, 13], where explicit witnesses of causal nonseparability have been constructed. However, when Charlie discards the control qubit, or both the control and target qubits, the resulting process matrix is causally separable. In the latter case, for instance, we obtain the bipartite process matrix

$$\text{Tr}_{C_I^c C_I^t} W^{\text{switch}} = \frac{1}{2}\big(|\psi\rangle\langle\psi|^{A_I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_O B_I} \otimes \mathbb{1}^{B_O} + |\psi\rangle\langle\psi|^{B_I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{B_O A_I} \otimes \mathbb{1}^{A_O}\big) \tag{1.41}$$

which just describes a probabilistic mixture of two fixed-order circuits. The quantum switch thus indeed requires the coordinated action of three parties to exhibit causal nonseparability. As shown in Refs. [11, 15], a consequence of this fact is also that all correlations that the process matrix of the quantum switch can generate are causal, i.e. the quantum switch cannot violate causal inequalities. It is, in fact, an open question whether a violation of a causal inequality through some physical, experimentally realisable process is possible in a meaningful way.

1.6 Information processing advantages using indefinite causal structures

As mentioned in the introduction, one of the motivations for the investigation of indefinite quantum causal structures is the fact that scenarios without definite causal order are more general than causally ordered, standard quantum circuits, and that new resources for quantum information processing become available when the assumption of a definite causal structure is relaxed. In this subsection we present a brief overview over some quantum information processing tasks that have been studied in this context, and some of the advantages that have been identified.

As the quantum switch is so far the only causally nonseparable process with a physical interpretation, the majority of studies have concentrated on this particular example. In Ref. [23] it was shown that by superposing the orders of operations, one can indeed achieve a computational advantage over fixed-order quantum circuits in a particular task. Assume that Alice and Bob apply unitaries U_A and U_B that either commute or anticommute, and the goal of the task is to determine which of the two statements is true. For that purpose, Charlie measures the control qubit in the $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$-basis. Writing the output state (1.37) of the quantum switch as $\frac{1}{2}(\{U_B, U_A\} |\psi\rangle^t \otimes |+\rangle^c + [U_B, U_A] |\psi\rangle^t \otimes |-\rangle^c)$, it can be seen that the control qubit is found in $|-\rangle$ with certainty if the operations U_A and U_B anticommute, and in $|+\rangle$ if they commute. Thus,

using the quantum switch, just one application of each operation is necessary to determine with certainty which of the two statements is true. This is not possible when A and B are connected in a standard quantum circuit [11, 23]. Therefore, the quantum switch indeed represents a new computational resource beyond causally ordered quantum circuits. This task was further studied in Ref. [11] in the context of causal nonseparability and causal witnesses.

Based on this discrimination problem between commuting and anticommuting unitaries, it was shown in subsequent works that the quantum switch allows for advantages in certain communication complexity tasks where two or more separate parties cooperate in order to compute a function of their local input data, while using a minimal amount of communication [27, 28]. In Ref. [28], a task was found for which there is an exponential separation in communication complexity (i.e., in the number of qubits that need to be communicated to accomplish the task) between the protocol using the quantum switch and any causally ordered protocol.

In Ref. [25], the above bipartite unitary discrimination task was generalised, and it was found that using N gates in a superposition of orders allows for an advantage in computation complexity, compared to fixed-order quantum circuits. An algorithm was presented that requires $O(N^2)$ queries to the operations in the fixed-order case, but only $O(N)$ queries in the protocol using a superposition of orders.

Another curious feature of the quantum switch has been pointed out in a recent series of papers (Refs. [29, 48, 49]). It was shown that combining noisy channels in a superposition of orders can enhance the classical and quantum channel capacity. For instance, when the two operations applied by A and B are completely depolarizing channels, information about the initial state of the target system they act upon can be acquired after the passage through the switch, despite the fact that a completely depolarizing channel by itself (and, by extension, any causally ordered composition of two such channels) always outputs a completely mixed state [29]. We will come back to this counterintuitive effect in Chapter 4.

Chapter 2

On the definition and characterisation of multipartite causal (non)separability

This chapter is based on the publication:
Julian Wechs, Alastair A. Abbott, and Cyril Branciard. On the definition and characterisation of multipartite causal (non)separability, *New J. Phys. 21 013027 (2019)* (Ref. [32]).

The text (from Sec. 2.1.2) corresponds to the main text of the published paper (from Sec. 3.2), with minor modifications to avoid repetitions, to make notations consistent and to integrate the content into the broader context of the thesis. The technical details from the Appendix of Ref. [32] are not reproduced here. We refer to the published paper at the relevant passages in the text. The study of multipartite causal (non)separability has been one of the central parts of my thesis, and I have contributed to all parts of the research presented in this chapter. In particular, I established the equivalence between our multipartite definition and that of "extensible causal (non)separability", developed the multipartite characterisations and a large part of the technical points in Ref. [32].

While the definition of causal (non)separability is unambiguous in the bipartite case, its generalisation to the multipartite case is not so straightforward. Two seemingly different generalisations have previously been proposed in the literature, one for the restricted tripartite scenario of the quantum switch (Sec. 1.5), and one for the general multipartite case. In this chapter, we present the results from Ref. [32], where we compared the two, and showed that they are in fact inequivalent. We then proposed our own definition of causal (non)separability for the general case, which—although a priori subtly different—turned out to be equivalent to the concept of "extensible causal (non)separability" introduced in Ref. [15], and which we argued is a more natural definition for general multipartite scenarios. We then turned to the question of how multipartite causal (non)separable process matrices can be characterised in terms of simple

conditions, and how multipartite causal (non)separability can be certified in practice. After recalling the characterisations of bipartite and tripartite causally (non)separable process matrices (cf. Refs. [11, 15]), we derived necessary, as well as sufficient conditions to characterise causally (non)separable process matrices in the general N-partite case. These allow one to devise practical tests, by generalising the tool of witnesses of causal nonseparability.

2.1 Defining multipartite causal (non)separability

2.1.1 Araújo *et al.*'s definition

The multipartite case was first considered in the restricted tripartite situation of the quantum switch, where one party has no (or, equivalently, a trivial) outgoing system (see Section 1.5). For this special case, Araújo *et al.* introduced the definition 1.4, which they used to characterise and study the causal nonseparability of the process matrix describing the quantum switch (Eq. (1.39)). This definition has subsequently been used e.g. in Refs. [13, 19, 40].

2.1.2 Oreshkov and Giarmatzi's definitions

While Araújo *et al.*'s definition recalled above applied only to a particular tripartite situation, Oreshkov and Giarmatzi (OG) considered in Ref. [15] the general multipartite case—taking into account, in particular, the possibility of dynamical causal orders. They defined in fact two possible generalisations of bipartite causal (non)separability, namely what they called the notions of "causal (non)separability" and "extensible causal (non)separability".

The definition they proposed for causal separability is recursive, in analogy with the definition of multipartite "causal correlations" [14, 15] (see Section 1.4.2). Oreshkov and Giarmatzi invoked an analogous "unraveling argument" for causally separable processes.

More specifically, their definition is based on the concept of a "conditional (process) matrix", defined for a given matrix W and a given CP map $M_k \coloneqq M^{A_{IO}^k}_{a_k|x_k}$ applied by a party A_k as

$$W_{|M_k} \coloneqq \mathrm{Tr}_k \left[M_k^T \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W \right]. \tag{2.1}$$

In general, even if W is a valid process matrix, $W_{|M_k}$ thus defined may not be a valid process matrix (in which case we shall just talk about a "conditional matrix"). In fact, as we will see in Sec. 2.2.1, a process matrix W is compatible with party A_k acting first (i.e., it does not allow signalling from the other parties to A_k) if and only if for any CP map M_k the conditional matrix

$W_{|M_k}$, as defined in Eq. (2.1), is (up to normalisation[1]) a valid $(N-1)$-partite process matrix for the parties in $\mathcal{N} \setminus k$. In that case, the conditional process matrix $W_{|M_k}$ then represents the process shared by these $N-1$ parties, conditioned on party A_k performing the CP map $M_k = M^{A^k_{IO}}_{a_k|x_k}$ (i.e., conditioned on both receiving the input x_k and obtaining the outcome a_k).

Oreshkov and Giarmatzi then proposed the following (recursive) definition:[2]

Definition 2.1 (Oreshkov and Giarmatzi's causal separability [15])**.** For $N = 1$, any process matrix is causally separable. For $N \geq 2$, an N-partite process matrix W is said to be *causally separable* if and only if it can be decomposed as

$$W = \sum_{k \in \mathcal{N}} q_k \, W_{(k)}, \tag{2.2}$$

with $q_k \geq 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}$ is a process matrix compatible with party A_k acting first, and is such that for any possible CP map $M_k \in \mathcal{L}(\mathcal{H}^{A^k_{IO}})$ applied by party A_k, the conditional $(N-1)$-partite process matrix $(W_{(k)})_{|M_k} := \mathrm{Tr}_k[M_k^T \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W_{(k)}]$ is itself causally separable.

As outlined in the previous section, the process matrix framework allows for process matrices to be extended by providing additional ancillary states to the parties. Taking this into account, OG introduced a second definition of causal separability for process matrices that are causally separable even under arbitrary such extensions:

Definition 2.2 (Oreshkov and Giarmatzi's extensible causal separability [15])**.** An N-partite process matrix W is said to be *extensibly causally separable* if and only if it is causally separable (as per Definition 2.1 above), and it remains so under any extension with incoming systems in an arbitrary joint quantum state—i.e., if and only if for any extension $\mathcal{H}^{A^{\mathcal{N}}_{I'}}$ of the parties' incoming spaces and any ancillary quantum state $\rho \in \mathcal{L}(\mathcal{H}^{A^{\mathcal{N}}_{I'}})$, $W \otimes \rho$ is causally separable.

It is easy to see that OG's causal separability (CS) and extensible causal separability (ECS) are equivalent in the bipartite case, and, indeed, equivalent to Definition 1.2 given in Sec. 1.2.2: the process matrix $W \otimes \rho$ obtained by attaching an ancillary state ρ to a causally separable process matrix W of the form of Eq. (1.14) remains of the same form, with $W^{A \prec B} \otimes \rho$ ($W^{B \prec A} \otimes \rho$) compatible with A acting before B (B before A), and for both terms $W^{A \prec B} \otimes \rho$ and $W^{B \prec A} \otimes \rho$, whatever operation the first party applies, the resulting conditional process matrix for the other party is single-partite, hence trivially causally separable.

[1] That is, it does not satisfy the third constraint in Eq. (1.32). In this chapter, we will for convenience often suppress this third constraint, and we will refer to any matrix satisfying the first two constraints in Eq. (1.32) (i.e., $W \geq 0$ and $W \in \mathcal{L}^{\mathcal{N}}$) as a (*valid*) process matrix. The normalisation can always be restored afterwards. Here, for a properly normalised process matrix W compatible with A_k first (i.e., which always gives $P(a_k|\vec{x}) = P(a_k|x_k)$) and a trace-non-increasing CP map $M_k = M_{a_k|x_k}$, one has $\mathrm{Tr}\, W_{|M_k} = P(a_k|x_k) \prod_{j \in \mathcal{N}\setminus k} d_{A^j_O}$, so that $W_{|M_k}$ must be divided by the factor $P(a_k|x_k)$ to also be properly normalised according to Eq. (1.32).

[2] More precisely, what we presented in Ref. [32] as their definition is actually presented in Ref. [15] (in a slightly different, but equivalent way) as a characterisation following from a more fundamental recursive definition of causally separable processes (not necessarily quantum mechanical).

However, OG's CS and ECS are not equivalent in the general multipartite case and thus indeed represent two different possible multipartite generalisations of the same bipartite concept. Of course ECS implies CS, but the converse is not true in general—the result of a phenomenon called "activation of causal nonseparability" in Ref. [15]. An explicit example of a CS process matrix that is not ECS was indeed given in [15], in a tripartite scenario where one party has no incoming system; we will see another example in the following subsection.

2.1.3 Comparison

We thus now have three potential generalisations of the concept of causal separability to the particular tripartite situation where one party has no outgoing system—namely, the two different definitions of causal separability (Definitions 1.4 and 2.1), and that of extensible causal separability (Definition 2.2). How do they relate to one another? Are the two definitions of causal separability indeed equivalent? These questions are answered by the following result:

Proposition 2.3. *In a tripartite scenario where party C has no outgoing system, Araújo* et al.*'s definition of causal separability (Definition 1.4) is equivalent to Oreshkov and Giarmatzi's definition of extensible causal separability (Definition 2.2), but nonequivalent to their definition of causal separability (Definition 2.1).*

The equivalence between Definitions 1.4 and 2.2 for this particular tripartite scenario is proved explicitly in Appendix B.1.1 (page 25) of the published paper [32], which we refer to for more details; we simply summarise the argument here as follows. Clearly, any process matrix W of the form of Eq. (1.40) is ECS, as any $W \otimes \rho$ is also of that form (and of the form also of Eq. (2.2)), and for any $W^{A\prec B\prec C}$ and any M_A, the conditional process $(W^{A\prec B\prec C})_{|M_A}$ is compatible with the order $B \prec C$ (hence it is causally separable; similarly for any $W^{B\prec A\prec C}$ and any M_B). The proof that an ECS process matrix W necessarily has the form of Eq. (1.40) is based on a "teleportation technique" (Lemma B.1 in Appendix B of [32]), already used in Ref. [15], that consists in introducing an ancillary system in a maximally entangled state ρ shared by two parties, e.g. A and C. By definition, the global process matrix $W \otimes \rho^{A_{I'}C_{I'}}$ has a decomposition of the form (2.2). It is then easy to see that the terms W_A and W_B compatible with parties A or B acting first are in fact compatible, since C has no outgoing system, with the causal orders $A \prec B \prec C$ and $B \prec A \prec C$, respectively, and thus contribute to the terms $W^{A\prec B\prec C}$ and $W^{B\prec A\prec C}$ in Eq. (1.40). For the term W_C compatible with C acting first, letting C project his systems in $\mathcal{L}(\mathcal{H}^{C_{II'}}) := \mathcal{L}(\mathcal{H}^{C_I} \otimes \mathcal{H}^{C_{I'}})$ onto the maximally entangled state effectively "teleports" his system to A. By definition, the conditional bipartite process matrix then shared by A and B must be causally separable, and must therefore have a decomposition of the form (1.14), which also leads to a decomposition of the form (1.40) for W_C.

In order to prove the nonequivalence between Araújo *et al.* and OG's definitions of causal separability, we showed in Ref. [32] that OG's CS and ECS are nonequivalent—i.e., that there can be "activation of causal nonseparability" (according to OG's terminology)—in the scenario

where party C has no outgoing system. Note that this scenario differs from that in which OG already gave an example of activation of causal nonseparability: they indeed considered a tripartite case where C has no *incoming* system, rather than no *outgoing* system.

Consider for that the following process matrix:

$$W^{\text{act.}} := \frac{1}{8}\Big[\mathbb{1}(\mathbb{1}\mathbb{1}-\hat{\mathrm{Z}}\hat{\mathrm{Z}})\mathbb{1}\mathbb{1} + \frac{\sqrt{3}}{4}\,\mathbb{1}(\hat{\mathrm{X}}\hat{\mathrm{X}}+\hat{\mathrm{Y}}\hat{\mathrm{Y}})(\hat{\mathrm{Z}}\mathbb{1}+\mathbb{1}\hat{\mathrm{Z}}) + \frac{1}{2}\hat{\mathrm{Z}}(\hat{\mathrm{Z}}\mathbb{1}-\mathbb{1}\hat{\mathrm{Z}})\mathbb{1}\mathbb{1} + \frac{1}{4}\hat{\mathrm{X}}(\hat{\mathrm{X}}\hat{\mathrm{Y}}-\hat{\mathrm{Y}}\hat{\mathrm{X}})(\hat{\mathrm{Z}}\mathbb{1}-\mathbb{1}\hat{\mathrm{Z}})\Big], \tag{2.3}$$

where the subsystems are written, for convenience, in the order $C_I A_I B_I A_O B_O$ (i.e., $W^{\text{act.}} \in \mathcal{L}(\mathcal{H}^{C_I} \otimes \mathcal{H}^{A_I} \otimes \mathcal{H}^{B_I} \otimes \mathcal{H}^{A_O} \otimes \mathcal{H}^{B_O})$).

We note first that $W^{\text{act.}}$ is compatible with Charlie acting first—i.e., with the order $C \prec \{A, B\}$.[3] (Indeed, it satisfies Eq. (2.10) given later, for $A_k = C$.) Any CP map applied by Charlie—i.e., since C has no outgoing system, any element of a POVM in his qubit incoming space $\mathcal{L}(\mathcal{H}^{C_I})$—can be written as $M_{\vec{c}} = \mathbb{1} + \vec{c}\cdot\vec{\sigma}$, where $\vec{\sigma} := (\hat{\mathrm{X}}, \hat{\mathrm{Y}}, \hat{\mathrm{Z}})$ and $\vec{c} := (c_{\mathrm{X}}, c_{\mathrm{Y}}, c_{\mathrm{Z}})$ is a 3-dimensional real vector with $|\vec{c}| \leq 1$, so that $M_{\vec{c}} \geq 0$ (and where we ignore the trace-nonincreasing constraint, and indeed the overall normalisation of $M_{\vec{c}}$, since it is irrelevant for our argument). The resulting conditional matrix for parties A and B (as defined in Eq. (2.1)) is then

$$(W^{\text{act.}})_{|M_{\vec{c}}} := \mathrm{Tr}_{C_I}[(M_{\vec{c}}^{C_I})^T \otimes \mathbb{1}^{A_I B_I A_O B_O} \cdot W^{\text{act.}}] = \frac{1}{2}W^{A\prec B}_{|M_{\vec{c}}} + \frac{1}{2}W^{B\prec A}_{|M_{\vec{c}}} \tag{2.4}$$

with (written in the order $A_I B_I A_O B_O$)

$$W^{A\prec B}_{|M_{\vec{c}}} := \frac{1}{4}\Big[(\mathbb{1}\mathbb{1}-\hat{\mathrm{Z}}\hat{\mathrm{Z}})\mathbb{1}\mathbb{1} + \frac{\sqrt{3}}{2}(\hat{\mathrm{X}}\hat{\mathrm{X}}+\hat{\mathrm{Y}}\hat{\mathrm{Y}})\hat{\mathrm{Z}}\mathbb{1} + \frac{c_{\mathrm{Z}}}{2}(\hat{\mathrm{Z}}\mathbb{1} - \mathbb{1}\hat{\mathrm{Z}})\mathbb{1}\mathbb{1} + \frac{c_{\mathrm{X}}}{2}(\hat{\mathrm{X}}\hat{\mathrm{Y}}-\hat{\mathrm{Y}}\hat{\mathrm{X}})\hat{\mathrm{Z}}\mathbb{1}\Big] \tag{2.5}$$

and with $W^{B\prec A}_{|M_{\vec{c}}}$ of a similar form, obtained from $W^{A\prec B}_{|M_{\vec{c}}}$ by changing $\hat{\mathrm{Z}}^{A_O}\mathbb{1}^{B_O}$ to $\mathbb{1}^{A_O}\hat{\mathrm{Z}}^{B_O}$ and c_{X} to $-c_{\mathrm{X}}$.

Note that $W^{A\prec B}_{|M_{\vec{c}}}$ and $W^{B\prec A}_{|M_{\vec{c}}}$ are valid, causally ordered process matrices, compatible with $A \prec B$ and $B \prec A$, respectively (their eigenvalues are found to be $0, \frac{1}{2}(1 \pm \sqrt{\frac{3+c_{\mathrm{X}}^2+c_{\mathrm{Z}}^2}{4}}) \geq 0$ for $|\vec{c}| \leq 1$, and they satisfy the appropriate form of Eq. (2.12) given later). From Eq. (2.4) and the definition of causal separability in the bipartite case (Definition 1.2), we conclude that for any CP map (i.e. here, any POVM element) $M_{\vec{c}}$ applied by Charlie, $(W^{\text{act.}})_{|M_{\vec{c}}}$ is a (bipartite) causally separable process matrix. Therefore, according to OG's Definition 2.1, $W^{\text{act.}}$ is a tripartite CS process matrix (with a single term in the decomposition (2.2), corresponding to C first).

A crucial feature of the decomposition (2.4) is that $(W^{\text{act.}})_{|M_{\vec{c}}}$, $W^{A\prec B}_{|M_{\vec{c}}}$ and $W^{B\prec A}_{|M_{\vec{c}}}$ all depend on Charlie's operation $M_{\vec{c}}$. Even though any valid process matrix in $\mathcal{L}(\mathcal{H}^{C_I} \otimes \mathcal{H}^{A_I} \otimes \mathcal{H}^{B_I} \otimes \mathcal{H}^{A_O} \otimes \mathcal{H}^{B_O})$ (including $W^{A\prec B}_{|M_{\vec{c}}}$ and $W^{B\prec A}_{|M_{\vec{c}}}$) is compatible with C acting last (since C has no outgoing system), the decomposition (2.4) still does not allow us to obtain a decomposition of

[3] As C has no outgoing system, $W^{\text{act.}}$ is also compatible with C acting last. But to prove that $W^{\text{act.}}$ is CS (according to OG's definition) as we do below we need to consider C acting first.

the form $W^{\text{act.}} = \frac{1}{2}W^{A\prec B\prec C} + \frac{1}{2}W^{B\prec A\prec C}$ for $W^{\text{act.}}$ (or even with different weights q, $1-q$), as in Eq. (1.40). Indeed, such a decomposition for $W^{\text{act.}}$, with $W^{A\prec B\prec C}$ and $W^{B\prec A\prec C}$ valid process matrices compatible with the indicated causal order, does not exist. This can be shown using Araújo *et al.*'s technique of "witnesses of causal nonseparability" [11, 13]: one can construct a witness for $W^{\text{act.}}$, and we give one explicitly in Appendix C of [32].

Since, as stated above, the existence of such a decomposition (as in Definition 1.4) would be equivalent in the scenario considered here to OG's ECS (Definition 2.2), this implies that although $W^{\text{act.}}$ is CS according to OG's Definition (see above), it is not ECS. This provides an explicit example of "activation of causal nonseparability" in that scenario.

Hence, OG's CS does not reduce (contrary to OG's ECS) to Araújo *et al.*'s definition of causal separability in this particular scenario. Definitions 1.4 and 2.1 of causal separability are therefore inconsistent. Our aim now is to rectify this inconsistency.

2.1.4 Our choice of definition

To fix this, we proposed in Ref. [32] our own definition of multipartite causal separability, which indeed resolves the inconsistency pointed out above, and which we argue is a more natural definition for general multipartite scenarios. Similarly to OG, we choose a recursive definition, based on the concept of a conditional process matrix and very much in the spirit of the recursive definitions that have been given for multipartite causal correlations [14, 15]. For a process matrix to be compatible with a definite causal order, there should, in any run of the experiment, be a designated party that acts first (which party this is can be determined probabilistically, just like in the bipartite case) and the conditional process matrix for the remaining parties, which depends on the action of the first party, should again be causally separable for any CP map that the first party applies.

For several reasons, we consider it important to allow extensions with extra incoming systems, similar to OG's *extensible* causal separability. Firstly, the whole process matrix framework is constructed so as to allow for shared ancillary systems between the parties. For consistency, we should thus take into account such extensions with shared incoming quantum states when defining causal (non)separability. Indeed, entanglement is a very different resource from causal nonseparability: entangled systems do not by themselves allow signalling between parties, and should be able to be distributed between parties prior to an experiment without "activating" causal nonseparability. (Note, however, that entanglement can still play a crucial role in causal nonseparability, as e.g., in the quantum switch, where the control and target systems can end up being entangled after the parties' operations.) While a "resource theory" for causal nonseparability has not yet been developed, it is reasonable to expect that providing additional shared (entangled) incoming states should be a free operation in such an approach. These considerations led us to propose the following definition.

Definition 2.4 (N-partite causal separability)**.** For N=1, any process matrix is causally separable. For $N \geq 2$, an N-partite process matrix W is said to be *causally separable* if and only if, for any extension $\mathcal{H}^{A_{I'}^{\mathcal{N}}}$ of the parties' incoming spaces and any ancillary quantum state $\rho \in \mathcal{L}(\mathcal{H}^{A_{I'}^{\mathcal{N}}})$, $W \otimes \rho$ can be decomposed as

$$W \otimes \rho = \sum_{k \in \mathcal{N}} q_k \, W_{(k)}^{\rho}, \tag{2.6}$$

with $q_k \geq 0$, $\sum_k q_k = 1$, and where for each k, $W_{(k)}^{\rho} \in \mathcal{L}(\mathcal{H}^{A_{II'O}^{\mathcal{N}}})$ is a process matrix compatible with party A_k acting first, and is such that for any CP map $M_k \in \mathcal{L}(\mathcal{H}^{A_{II'O}^{k}})$ applied by party A_k, the conditional $(N-1)$-partite process matrix[4] $(W_{(k)}^{\rho})_{|M_k} := \mathrm{Tr}_k[M_k^T \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W_{(k)}^{\rho}]$ is itself causally separable.

Note that there is a subtle difference between our definition here and that of OG's ECS (Definition 2.2). We indeed require all conditional process matrices appearing at all levels of the recursive decomposition to remain causally separable under extension with arbitrary ancillary states, while OG impose this *a priori* only for the original process matrix. In fact, although *prima facie* different, these definitions turn out to be equivalent; the proof of this is given in Appendix D of the published paper [32].

From Definition 2.4 we recover the natural, intuitive definition of Araújo *et al.* [11] in the particular tripartite case where one party has a trivial outgoing system—a case of practical relevance, as the quantum switch is the first example of a causally nonseparable process that has been demonstrated and studied in laboratory experiments [16–22]. One can also readily verify that process matrices that are causally separable by Definition 2.4 cannot generate noncausal correlations (as defined in Refs. [14, 15], Sec. 1.4.2); an explicit proof is given in Appendix E of [32].

From now on, whenever we talk about causal (non)separability we will refer to our Definition 2.4.

2.2 Characterising multipartite causal (non)separability

With the definition of causal (non)separability given above, we now turn to the question of how to characterise causally separable process matrices in terms of simple conditions and how to demonstrate multipartite causal nonseparability in practice.

For that we will start by reviewing the characterisations of valid process matrices and of process matrices compatible with fixed causal orders, before recalling the characterisations of causally separable process matrices in the bipartite and tripartite cases, where we will give conditions for causal separability that are both necessary and sufficient. We then present a generalisation

[4] Note that compared to Eq. (2.1), we take here $\mathcal{H}^{A_{II'O}^{k}} := \mathcal{H}^{A_{IO}^{k}} \otimes \mathcal{H}^{A_{I'}^{k}}$, $M_k := M_{a_k|x_k}^{A_{II'O}^{k}}$, $\mathrm{Tr}_k := \mathrm{Tr}_{A_{II'O}^{k}}$ and $\mathbb{1}^{\mathcal{N}\setminus k} := \bigotimes_{j \in \mathcal{N}\setminus k} \mathbb{1}^{A_{II'O}^{j}}$ in the definition of the conditional matrix.

to the N-partite case which, for $N \geq 4$, gives two conditions, one necessary and one sufficient, whose coincidence remains an open question.

As in Sec. 1.2.3 we will not concern ourselves with the normalisation of process matrices (which can always be imposed later). Our characterisations will then be given in terms of linear subspaces of Hermitian matrices (e.g., the spaces $\mathcal{L}^{\mathcal{N}}$ and $\mathcal{L}^{A_1 \prec \cdots \prec A_N}$ introduced already in Sec. 1.4.1); when adding the requirement of positive semidefiniteness, the corresponding sets of (nonnormalised) process matrices will thus be closed convex cones of positive semidefinite matrices. This will allow the conditions we give to be checked efficiently with semidefinite programming (SDP) techniques. In particular, by generalising the techniques used for the bipartite and restricted tripartite cases in Refs. [11, 13], we will extend the idea of witnesses of causal nonseparability to the multipartite case and show how multipartite witnesses can be constructed efficiently, allowing this causal nonseparability to be verified experimentally by having each party perform appropriately chosen measurements [17, 19] (see also Sec. 1.2.3).

2.2.1 Valid process matrices and compatibility with a fixed causal order

Recall from Sec. 1.4.1 that the conditions for a process matrix W to be valid arise from requiring that the generalised Born rule (1.31) should give valid probability distributions, even when the parties share arbitrary ancillary systems. The fact that these probabilities should be nonnegative imposes that W must be positive semidefinite, while the requirement that these probabilities must sum to 1 implies that any valid (but, once again, not necessarily normalised) W must be in a linear space $\mathcal{L}^{\mathcal{N}}$ [4, 11]. As shown in Ref. [11], and Appendix A.1 (page 17) of the relevant paper [32], this subspace is formed by the Hermitian matrices in $\mathcal{L}(\mathcal{H}^{A_{IO}^{\mathcal{N}}})$ that satisfy the following conditions:

$$W \in \mathcal{L}^{\mathcal{N}} \Leftrightarrow \forall\, \mathcal{X} \subsetneq \mathcal{N}, \mathcal{X} \neq \emptyset,\ \mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W \in \mathcal{L}^{\mathcal{X}} \quad \text{and} \quad \textstyle\prod_{i\in\mathcal{N}}[1-A_O^i]W = 0 \tag{2.7}$$

$$\Leftrightarrow \forall\, \mathcal{X} \subseteq \mathcal{N}, \mathcal{X} \neq \emptyset,\ \textstyle\prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\mathcal{X}}W = 0\,. \tag{2.8}$$

Written in the form of Eq. (2.7), the validity constraint for W says that all reduced matrices $\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W$ shared by the parties of any strict subset $\mathcal{X}$ of $\mathcal{N}$ (obtained after tracing out the parties that are not in $\mathcal{X}$) must be valid, and that W must further satisfy the additional constraint that $\prod_{i\in\mathcal{N}}[1-A_O^i]W = 0$. The form of Eq. (2.8) expresses explicitly all the (linearly independent) constraints that these recursive validity conditions imply on W.[5] Denoting again by $\mathcal{P}$ the convex cone of positive semidefinite matrices, the set of valid process matrices is then the convex cone

$$\mathcal{W} = \mathcal{P} \cap \mathcal{L}^{\mathcal{N}}\,. \tag{2.9}$$

[5] Note that the constraint in Eq. (2.8) can also be written as $\prod_{i\in\mathcal{X}}[1-A_O^i](\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W) = 0$. In this chapter we generically use the form of Eq. (2.8) for ease of notation; it may be useful, however, to keep in mind that this type of constraint is in fact a constraint on the reduced matrix $\mathrm{Tr}_{\mathcal{N}\setminus\mathcal{X}}\, W$ shared by the parties in $\mathcal{X}$, as written more explicitly in Eq. (2.7).

In order to discuss the causal separability of process matrices, it is necessary to also characterise the subspaces of such matrices that are compatible with certain fixed causal relations between (subsets of) parties. Such causal relations, as for the particular cases of fixed causal orders discussed in the previous sections, are understood via the notion of signalling: if a (group of) parties is in the causal future of some others, then there is no way for them to signal to those earlier parties.

We first consider the case of process matrices that are compatible with a given party A_k acting first:[6] regardless of the operation performed by the other parties $A_{k'}$ (for all $k' \neq k$), the marginal probability distribution for A_k obtained from (1.31) must not depend on the CPTP maps $M_{x_{k'}}^{A_{IO}^{k'}}$ chosen by those other parties. As already mentioned in the previous section (and shown in Appendix A.2 (page 18) of our publication [32]), a given process matrix W satisfies this condition if and only if, whatever CP map M_k is applied by A_k, the conditional process matrix $W_{|M_k}$, as defined in Eq. (2.1), is a valid $(N-1)$-partite process matrix for the remaining parties in $\mathcal{N}\backslash A_k$.

We can in fact ignore here the assumption that $M_k \geq 0$, and the above constraint is equivalent to imposing that $W_{|M_k} \in \mathcal{L}^{\mathcal{N}\backslash A_k}$ for any Hermitian matrix $M_k \in \mathcal{L}(\mathcal{H}^{A_{IO}^k})$. Such a constraint defines a linear subspace of the space of Hermitian matrices in $\mathcal{L}(\mathcal{H}^{A_{IO}^{\mathcal{N}}})$. Taking its intersection with the subspace $\mathcal{L}^{\mathcal{N}}$, we denote the linear subspace of valid process matrices compatible with party A_k first by $\mathcal{L}^{A_k \prec (\mathcal{N}\backslash A_k)}$. We find, using Eq. (2.8) above (and after removing redundant constraints; see Appendix A.2 of Ref. [32]):

$$\begin{aligned} W \in \mathcal{L}^{A_k \prec (\mathcal{N}\backslash A_k)} &\Leftrightarrow \; W \in \mathcal{L}^{\mathcal{N}} \quad \text{and} \quad \forall \text{ Hermitian } M_k \in \mathcal{L}(\mathcal{H}^{A_{IO}^k}),\; W_{|M_k} \in \mathcal{L}^{\mathcal{N}\backslash A_k} \\ &\Leftrightarrow \; {}_{[1-A_O^k]} A_{IO}^{\mathcal{N}\backslash k} W = 0 \quad \text{and} \quad \forall\, \mathcal{X} \subseteq \mathcal{N}\backslash k, \mathcal{X} \neq \emptyset, \prod_{i\in\mathcal{X}} {}_{[1-A_O^i]} A_{IO}^{\mathcal{N}\backslash k \backslash \mathcal{X}} W = 0. \end{aligned} \tag{2.10}$$

In Appendix A.2 of the published paper [32] we also derive constraints for more general causal orders of the form $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \cdots \prec \mathcal{K}_K$, for various disjoint subsets $\mathcal{K}_i$ of $\mathcal{N}$.

In the case where the subsets $\mathcal{K}_i$ define a full partition of $\mathcal{N}$ (i.e., where $\bigcup_{i=1}^K \mathcal{K}_i = \mathcal{N}$), we found that the corresponding subspace $\mathcal{L}^{\mathcal{K}_1 \prec \cdots \prec \mathcal{K}_K}$ is characterised by

$$\begin{aligned} &W \in \mathcal{L}^{\mathcal{K}_1 \prec \cdots \prec \mathcal{K}_K} \\ &\quad \Leftrightarrow \; \forall\, k = 1, \ldots, K,\; \forall \text{ Hermitian } M_{\mathcal{K}_{(<k)}} \in \mathcal{L}(\mathcal{H}^{A_{IO}^{\mathcal{K}_{(<k)}}}),\; \mathrm{Tr}_{\mathcal{K}_{(>k)}} W_{|M_{\mathcal{K}_{(<k)}}} \in \mathcal{L}^{\mathcal{K}_k} \\ &\quad \Leftrightarrow \; \forall\, k = 1, \ldots, K,\; \forall\, \mathcal{X}_k \subseteq \mathcal{K}_k, \mathcal{X}_k \neq \emptyset,\; \prod_{i\in\mathcal{X}_k} {}_{[1-A_O^i]} A_{IO}^{\mathcal{K}_k \backslash \mathcal{X}_k} A_{IO}^{\mathcal{K}_{(>k)}} W = 0 \end{aligned} \tag{2.11}$$

with $\mathcal{K}_{(<k)} := \bigcup_{i=1}^{k-1} \mathcal{K}_i$ for $k = 2, \ldots, K$, $\mathcal{K}_{(<1)} = \mathcal{K}_{(>K)} = \emptyset$, $W_{|M_{\mathcal{K}_{(<1)}}} = W$.

Of particular interest is the specific case in which each $\mathcal{K}_i$ is a singleton, which gives constraints on a process matrix W being compatible with a fixed causal order such as $A_1 \prec A_2 \prec \cdots \prec A_N$.

[6] Note that a process matrix can be compatible with several different causal relations between parties. For example, if a matrix W does not allow any party to signal to another, then it is compatible with any party or group of parties acting first.

Such a W must be compatible with A_1 acting first (and must therefore satisfy Eq. (2.10) for $k = 1$—in particular, the first constraint on its second line); then, whatever CP map M_1 party A_1 applies, the resulting conditional process matrix $W_{|M_1}$ must then be a valid $(N-1)$-partite process matrix, compatible with party A_2 acting first (and must therefore satisfy Eq. (2.10) for $k = 2$—in particular, the first constraint on its second line—with $\mathcal{N}$ replaced by $\mathcal{N}\backslash\{1\}$); etc. By iterating this argument (up until the party A_N), we find that the linear subspace $\mathcal{L}^{A_1\prec\cdots\prec A_N}$ of process matrices compatible with the causal order $A_1 \prec \cdots \prec A_N$ is characterised by (cf. Refs. [11, 44, 45])

$$W \in \mathcal{L}^{A_1\prec\cdots\prec A_N} \Leftrightarrow \forall\, k = 1, \ldots, N,\ {}_{[1-A_O^k]}A_{IO}^{(>k)} W = 0\,, \tag{2.12}$$

with $A_{IO}^{(>k)} = A_{IO}^{\{k+1,\ldots,N\}}$ (with $A_{IO}^{(>N)} = A_{IO}^{\emptyset} = 1$).

2.2.2 Bipartite and tripartite causally (non)separable process matrices

In the bipartite scenario, the above characterisation of the subspaces $\mathcal{L}^{A\prec B}$ and $\mathcal{L}^{B\prec A}$ allows us, from Definition 1.2, to give the following explicit characterisation of causally separable process matrices.

Proposition 2.5 (Characterisation of bipartite causally separable process matrices)**.** *A matrix $W \in \mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}})$ is a valid bipartite causally separable process matrix if and only if it can be decomposed as*

$$W = W_{(A,B)} + W_{(B,A)} \tag{2.13}$$

where, for each permutation (X, Y) of the two parties A and B, $W_{(X,Y)}$ is a positive semidefinite matrix satisfying

$${}_{[1-X_O]}Y_{IO} W_{(X,Y)} = 0,\ {}_{[1-Y_O]}W_{(X,Y)} = 0 \tag{2.14}$$

(i.e., $W_{(X,Y)}$ is a valid process matrix compatible with the causal order $X \prec Y$).

Note that, in contrast to Eq. (1.14) in Definition 1.2, we did not write the weights q and $1 - q$ explicitly in Eq. (2.13). Instead, for convenience and consistency with the characterisations of tripartite and N-partite causally separable processes which will follow, we decomposed W in terms of nonnormalised process matrices, writing $W_{(A,B)} = q\, W^{A\prec B}$ and $W_{(B,A)} = (1-q)\, W^{B\prec A}$.

As we discussed in Sec. 2.1, the tripartite case of causal separability was already studied by Oreshkov and Giarmatzi under the name "extensible causal separability" in Ref. [15]. In their Proposition 3.3 they provided a characterisation of tripartite (extensible) causal separability, albeit describing the constraints in a different way. In our approach, this characterisation can be expressed as follows:

Proposition 2.6 (Characterisation of tripartite causally separable process matrices)**.** *A matrix $W \in \mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}} \otimes \mathcal{H}^{C_{IO}})$ is a valid tripartite causally separable process matrix (as per*

Definition 2.4) if and only if it can be decomposed as

$$\begin{aligned} W &= \overbrace{W_{(A)}}^{} + \overbrace{W_{(B)}}^{} + \overbrace{W_{(C)}}^{} \\ &= \overbrace{W_{(A,B,C)} + W_{(A,C,B)}}^{} + \overbrace{W_{(B,A,C)} + W_{(B,C,A)}}^{} + \overbrace{W_{(C,A,B)} + W_{(C,B,A)}}^{} \end{aligned} \tag{2.15}$$

where, for each permutation of the three parties (X, Y, Z)*,* $W_{(X,Y,Z)}$ *and* $W_{(X)} := W_{(X,Y,Z)} + W_{(X,Z,Y)}$ *are positive semidefinite matrices satisfying*

$$_{[1-X_O]Y_{IO}Z_{IO}}W_{(X)} = 0\,, \tag{2.16}$$

$$_{[1-Y_O]Z_{IO}}W_{(X,Y,Z)} = 0\,, \quad _{[1-Z_O]}W_{(X,Y,Z)} = 0\,. \tag{2.17}$$

The proof of this characterisation was sketched in Ref. [15] using a somewhat different terminology to what we employ; in particular, they express causal constraints in terms of restrictions of what terms are "allowed" in a Hilbert-Schmidt basis decomposition of a matrix. We gave a more detailed proof in Appendix B.1.2 (page 26) of Ref. [32], which is again based on a "teleportation technique", similar in spirit to the one briefly sketched in Sec. 2.1.3.

Let us break down and analyse the terms appearing in the decomposition (2.15) to understand better this characterisation.

From the constraints in Eq. (2.17) it follows that, that for each party X, the matrix $W_{(X)}(= W_{(X,Y,Z)} + W_{(X,Z,Y)})$ satisfies $_{[1-Y_O]Z_{IO}}W_{(X)} = {}_{[1-Z_O]Y_{IO}}W_{(X)} = {}_{[1-Y_O][1-Z_O]}W_{(X)} = 0$. Together with Eq. (2.16) and the fact that $W_{(X)}$ is positive semidefinite, this implies that $W_{(X)}$ is a valid tripartite process matrix compatible with party X acting first (since it satisfies Eq. (2.10) for $A_k = X$). W is thus decomposed in Eq. (2.15) as a sum of 3 valid process matrices, which ensures in particular that it is itself a valid process matrix.

On the other hand, the matrices $W_{(X,Y,Z)}$ in the decomposition (2.15) are not necessarily valid process matrices. Nevertheless, the constraints (2.17) imply that whatever the CP map M_X applied by the first party X, the conditional process matrix $(W_{(X,Y,Z)})_{|M_X} := \mathrm{Tr}_X[M_X^T \otimes \mathbb{1}^{YZ} \cdot W_{(X,Y,Z)}]$ is a valid bipartite process matrix, compatible with the causal order $Y \prec Z$ (indeed, it satisfies Eq. (2.12) for this causal order: e.g., $_{[1-Y_O]Z_{IO}}[(W_{(X,Y,Z)})_{|M_X}] = (_{[1-Y_O]Z_{IO}}W_{(X,Y,Z)})_{|M_X} = 0$).

The fact that the matrices $W_{(X,Y,Z)}$ are not necessarily valid process matrices, and thus that Eq. (2.15) does not simply decompose W into a combination of process matrices compatible with fixed causal orders, is a consequence of the possibility of dynamical (but still well-defined, albeit not *fixed*) causal orders. In Sec. 2.2.5 we will consider a fourpartite version of the quantum switch as a concrete example of a process matrix allowing for such dynamical causal orders.

2.2.3 General multipartite causally (non)separable process matrices

As we will see below, it is possible to generalise the decomposition of Proposition 2.6 to the case of N-partite causal separability. While the generalisation clearly provides a sufficient condition

for causal separability, the proof that it is also a necessary condition does not readily generalise. Indeed, the proof for the tripartite case relies on the fact that each term $W_{(X)}$ in Eq. (2.15) is the sum of only two "base" terms, something that is not true in the natural generalisation of this decomposition. (To understand this better, we encourage the interested reader to look at the subtleties of that proof [32].)

For the general multipartite case, we therefore provided in Ref. [32] the following, separate, necessary and sufficient conditions. Since these arise from different considerations, we will present and discuss these individually. Indeed, although these coincide in the bipartite and tripartite cases, it remains an open question whether this is the case in general (or if one is both necessary and sufficient but not the other, or if neither are).

2.2.3.1 Necessary condition

The necessary condition is based on the teleportation technique and is a generalisation of the use of this approach in the proof of the tripartite characterisation. Let us outline how it leads to the necessary condition to help understand the condition itself (for a more formal description of the teleportation technique, we refer to Lemma B.1 in Appendix B of [32]). The idea is to consider, in Eq. (2.6) of Definition 2.4, a specific shared incoming ancillary state, as well as specific operations M_k applied by the parties A_k, for which there is a straightforward relation between the forms of the respective N-partite process matrices in which A_k acts first, and the corresponding $(N-1)$-partite conditional process matrices that we obtain after A_k has operated. As the latter are by definition causally separable (and satisfy thus the necessary conditions for $(N-1)$-partite causal separability), this allows us to infer necessary conditions for the causal separability of the original N-partite process matrix.

More precisely, we provide, as ancillary incoming systems, a maximally entangled state between every pair of parties, defining an overall ancillary state ρ. If W is a causally separable process matrix, then, by definition, $W \otimes \rho$ can be decomposed into a sum of process matrices $W^{\rho}_{(k)}$ compatible with a given party A_k acting first (cf. Eq. (2.6) in Definition 2.4); furthermore, as ρ is pure, one can write $W^{\rho}_{(k)} = W_{(k)} \otimes \rho$ with $W_{(k)}$ itself being compatible with A_k first. For each such process matrix $W_{(k)}$ the party A_k can then "teleport" the part of $W_{(k)}$ on their systems A^k_{IO} to another party $A_{k'}$ by applying an appropriate CP map M_k. The effect is that the resulting $(N-1)$-partite conditional process matrix $(W^{\rho}_{(k)})_{|M_k}$ formally has the same form as $W_{(k)}$ (tensored with what is left over of the, now reduced, ancillary state ρ), except that the systems A^k_{IO} are instead attributed ("teleported") to the ancillary incoming system $A^{k'}_{I'}$ of $A_{k'}$. From the definition of causal separability, $(W^{\rho}_{(k)})_{|M_k}$ must itself be causally separable, so the necessary condition can be recursively applied to this $(N-1)$-partite process matrix until the base case of $N = 3$, given by Proposition 2.6, is reached.

In order to state more formally the condition itself, let us introduce the following notation. For a given matrix $W \in \mathcal{L}(\mathcal{H}^{A^{\mathcal{N}}_{IO}})$, we denote by $W^{A^k_{IO} \to A^{k'}_{I'}} \in \mathcal{L}(\mathcal{H}^{A^{\mathcal{N}\setminus k}_{IO}} \otimes \mathcal{H}^{A^{k'}_{I'}})$ the same matrix,

where the systems A^k_{IO} are attributed to some other system $A^{k'}_{I'}$ (of the same dimension as A^k_{IO}). More formally,

$$W^{A^k_{IO}\to A^{k'}_{I'}} := \sum_{i,j} \mathrm{Tr}_k\left[|i\rangle\langle j|^{A^k_{IO}} \otimes \mathbb{1}^{\mathcal{N}\setminus k} \cdot W\right] \otimes |j\rangle\langle i|^{A^{k'}_{I'}}, \tag{2.18}$$

where $\{|i\rangle\}$ is an orthonormal basis of $\mathcal{H}^{A^k_I} \otimes \mathcal{H}^{A^k_O}$.

We then obtain the following recursive necessary condition:

Proposition 2.7 (Necessary condition for general multipartite causal separability)**.** *An N-partite causally separable process matrix $W \in \mathcal{L}(\mathcal{H}^{A^{\mathcal{N}}_{IO}})$ (as per Definition 2.4) must necessarily have a decomposition of the form*

$$W = \sum_{k\in\mathcal{N}} W_{(k)} \tag{2.19}$$

where each $W_{(k)}$ is a valid process matrix compatible with party A_k acting first, and such that for each $k' \neq k$, $W_{(k)}^{A^k_{IO}\to A^{k'}_{I'}}$ is an $(N-1)$-partite causally separable process matrix.

Hence, any constraints satisfied by $(N-1)$-partite causally separable process matrices must also be satisfied by $W_{(k)}$ after re-attributing the system $A^{k'}_{I'}$ back to A^k_{IO}—i.e., after formally replacing $A^{k'}_I$ by $A^{k'}_I A^{k'}_{I'}$ and then $A^{k'}_{I'}$ by A^k_{IO} in the constraints written using the notation defined in Eqs. (1.8)–(1.9).

The decomposition of Eq. (2.19) follows from that of Eq. (2.6) in our definition of causal separability, for the appropriate choice of ancillary state and CP maps, as described above. The full details of the proof of Proposition 2.7 are given in our paper [32] (Appendix B.2.1, page 29).

To further clarify this condition, let us illustrate, in the fourpartite case (with parties A, B, C, D), how one can use it to obtain explicit constraints on causally separable process matrices. Proposition 2.7 implies that a fourpartite causally separable process matrix W must be decomposable as

$$W = W_{(A)} + W_{(B)} + W_{(C)} + W_{(D)}, \tag{2.20}$$

with each $W_{(X)}$ (for $X = A, B, C, D$) being a valid process matrix compatible with party X acting first—hence satisfying Eq. (2.10) for $A_k = X$.[7] For each X and every other party $Y \neq X$, the recursive constraint that $W_{(X)}^{X_{IO}\to Y_{I'}}$ is a tripartite causally separable process matrix further implies, according to Proposition 2.6 (for the 3 parties $Y, Z, T \neq X$) and after re-attributing the system $Y_{I'}$ to X_{IO} (i.e., replacing Y_{IO} by $Y_{I'}Y_{IO}$ and then $Y_{I'}$ by X_{IO} in the constraints), that there must exist a decomposition of $W_{(X)}$ of the form[8]

$$\begin{aligned} W_{(X)} &= W^{[X\to Y]}_{(X,Y)} + W^{[X\to Y]}_{(X,Z)} + W^{[X\to Y]}_{(X,T)} \\ &= W^{[X\to Y]}_{(X,Y,Z,T)} + W^{[X\to Y]}_{(X,Y,T,Z)} + W^{[X\to Y]}_{(X,Z,Y,T)} + W^{[X\to Y]}_{(X,Z,T,Y)} + W^{[X\to Y]}_{(X,T,Y,Z)} + W^{[X\to Y]}_{(X,T,Z,Y)} \end{aligned} \tag{2.21}$$

[7] Note that the existence, for all Y, of a decomposition of the form of Eq. (2.21) satisfying Eq. (2.22) implies all the constraints of Eq. (2.10), except for the first constraint on the second line (i.e., ${}_{[1-X_O]Y_{IO}Z_{IO}T_{IO}}W = 0$).

[8] Here the superscripts $[X \to Y]$ are simply labels to indicate that, for each matrix $W_{(X)}$, there are potentially different decompositions of the form (2.21) for each $Y \neq X$. (The sufficient condition below will in fact precisely be obtained by assuming that these decompositions do not depend on Y.)

where each term appearing in the decomposition is positive semidefinite, $W^{[X\to Y]}_{(X,Y)} = W^{[X\to Y]}_{(X,Y,Z,T)} + W^{[X\to Y]}_{(X,Y,T,Z)}$, etc., and with (for all $X \neq Y \neq Z \neq T$)

$$\begin{aligned}
&{}_{[1-Y_O]}Z_{IO}T_{IO}W^{[X\to Y]}_{(X,Y)} = {}_{[1-Z_O]}X_{IO}Y_{IO}T_{IO}W^{[X\to Y]}_{(X,Z)} = 0,\\
&{}_{[1-Z_O]}T_{IO}W^{[X\to Y]}_{(X,Y,Z,T)} = {}_{[1-T_O]}W^{[X\to Y]}_{(X,Y,Z,T)} = 0,\\
&{}_{[1-Y_O]}T_{IO}W^{[X\to Y]}_{(X,Z,Y,T)} = {}_{[1-T_O]}W^{[X\to Y]}_{(X,Z,Y,T)} = 0,\\
&{}_{[1-T_O]}X_{IO}Y_{IO}W^{[X\to Y]}_{(X,Z,T,Y)} = {}_{[1-Y_O]}W^{[X\to Y]}_{(X,Z,T,Y)} = 0.
\end{aligned} \tag{2.22}$$

Finally, we remark that the constraints obtained by considering teleporting each party X's system to just a single other party Y (i.e., by just demanding the existence of a decomposition of the above form for some other party Y, rather than for *all* other parties $Y \neq X$) yields conditions that are still necessary for the causal separability of W, but which are generally weaker than those given in Proposition 2.7. Indeed, in Appendix F.1 (page 36) of Ref. [32], we give an example of a fourpartite process matrix which satisfies those weaker conditions but not all of those given above.

2.2.3.2 Sufficient condition

A sufficient condition for causal separability can be obtained by considering a stricter form of the recursive decomposition (2.19) in Proposition 2.7. In particular, we demand that W has a decomposition into $W_{(k)}$ compatible with A_k acting first and such that each $W_{(k)}$ itself recursively satisfies the sufficient constraints for an $(N-1)$-partite process matrix without A^k_{IO} being traced out. One can easily verify that the decomposition (2.15) in the tripartite case is a generalisation of this kind from the bipartite case. In the fourpartite case described explicitly above, this means that for each party X there should be a single decomposition of the form (2.21) (i.e., no longer dependent on Y) such that the constraints (2.22) are satisfied without tracing out X_{IO} on the first and fourth lines. The fact that, unlike in the necessary conditions, we only consider a single (recursive) decomposition of each $W_{(k)}$ means that we can give a more explicit formulation for the sufficient condition.

Before stating the sufficient condition, let us introduce some more notations. Let Π denote the set of permutations (generically denoted by π) of $\mathcal{N}$. For an ordered subset $(k_1, \ldots, k_n)$ of $\mathcal{N}$ with n elements (with $1 \leq n \leq N$, $k_i \neq k_j$ for $i \neq j$), let $\Pi_{(k_1,\ldots,k_n)}$ be the set of permutations of $\mathcal{N}$ for which the element k_1 is first, k_2 is second, ..., and k_n is n^{th}—i.e., $\Pi_{(k_1,\ldots,k_n)} = \{\pi \in \Pi \mid \pi(1) = k_1, \ldots, \pi(n) = k_n\}$. With these notations, we have the following sufficient condition, that directly generalises the decomposition of Proposition 2.6.

Proposition 2.8 (Sufficient condition for general multipartite causal separability)**.** *If a matrix $W \in \mathcal{L}(\mathcal{H}^{A^{\mathcal{N}}_{IO}})$ can be decomposed as a sum of $N!$ positive semidefinite operators $W_\pi \geq 0$ in the*

form

$$W = \sum_{\pi\in\Pi} W_\pi, \tag{2.23}$$

such that for any ordered subset of parties $(k_1, \dots, k_n)$ *of* $\mathcal{N}$ *(with* $1 \le n \le N$*,* $k_i \neq k_j$ *for* $i \neq j$*), the partial sum*

$$W_{(k_1,\dots,k_n)} := \sum_{\pi\in\Pi_{(k_1,\dots,k_n)}} W_\pi \tag{2.24}$$

satisfies

$$[1-A_O^{k_n}]A_{IO}^{\mathcal{N}\setminus\{k_1,\dots,k_n\}} W_{(k_1,\dots,k_n)} = 0, \tag{2.25}$$

then W *is a valid causally separable process matrix (as per Definition 2.4).*

This decomposition was also suggested independently by Oreshkov as a possible generalisation of Proposition 2.6 [50] (although following the approach of Refs. [4, 15], Oreshkov expressed it differently, namely in terms of allowed terms in a Hilbert-Schmidt basis decomposition of the matrices $W_{(k_1,\dots,k_n)}$). The proof that the condition above is indeed sufficient is given in the published paper in Appendix B.2.2 (page 29). In order to understand it better, it is nonetheless worth discussing the form of the decomposition and the terms appearing within in a little more detail.

Firstly, one can show by induction that if Eq. (2.25) is satisfied for all $(k_1, \dots, k_n)$, then one also has, for all $(k_1, \dots, k_n)$ with $1 \le n < N$, that

$$\forall\ \mathcal{X} \subseteq \mathcal{N}\setminus\{k_1,\dots,k_n\}, \mathcal{X} \neq \emptyset,\ \prod_{i\in\mathcal{X}}[1-A_O^i]A_{IO}^{\mathcal{N}\setminus\{k_1,\dots,k_n\}\setminus\mathcal{X}} W_{(k_1,\dots,k_n)} = 0. \tag{2.26}$$

Note also that since all $W_\pi \ge 0$, all $W_{(k_1,\dots,k_n)} \ge 0$ as well.

For $n = 1$, Eqs. (2.25) and (2.26) imply that each matrix $W_{(k_1)}\,(\ge 0)$ is a valid process matrix compatible with party A_{k_1} acting first; indeed, Eq. (2.10) is satisfied for $A_k = A_{k_1}$. As $W = \sum_{k_1} W_{(k_1)}$ according to Eqs. (2.23)–(2.24), this ensures in particular that W is indeed a valid process matrix.

Note, however, that in general the matrices $W_{(k_1,\dots,k_n)}$ for $n > 1$ are *not* valid processes matrices compatible with the causal order $A_{k_1} \prec \cdots \prec A_{k_n}$. Indeed, as we already observed in the tripartite case, $W_{(k_1,\dots,k_n)}$ may not generally be a valid process matrix at all. Nevertheless, comparing with Eq. (2.10), one can see that Eqs. (2.25) and (2.26) imply that whatever the CP maps $M_{k_1}, \dots, M_{k_{n-1}}$ applied by the $n-1$ parties $A_{k_1}, \dots, A_{k_{n-1}}$, the conditional matrix $(W_{(k_1,\dots,k_n)})_{|M_{k_1}\otimes\cdots\otimes M_{k_{n-1}}} := \mathrm{Tr}_{k_1,\dots,k_{n-1}}[M_{k_1}^T \otimes \cdots \otimes M_{k_{n-1}}^T \otimes \mathbb{1}^{\mathcal{N}\setminus\{k_1,\dots,k_{n-1}\}} \cdot W_{(k_1,\dots,k_n)}]$ is a valid $(N-n+1)$-partite process matrix, compatible with party A_{k_n} acting first. As we have noted already, the condition of Proposition 2.8 coincides, in the bipartite and tripartite cases, with those given in Propositions 2.5 and 2.6, respectively. Indeed, for these cases, the necessary and sufficient conditions given here coincided. For four-or-more parties it remains an open question whether this is also the case. We performed several numerical searches for process matrices satisfying the necessary but not sufficient conditions (see Appendix F.2 (page 37) of

the published paper [32]) and failed to find any such examples, although the complexity of the numerical searches means that we caution against interpreting this as evidence that the conditions coincide in general. In Appendix B.3 (page 31) of [32], however, we show that they do coincide in the specific fourpartite case with $d_{D_O} = 1$. This is a rather restricted scenario (where any process matrix is compatible with D acting last), but nonetheless includes cases of interest such as the fourpartite variant of the quantum switch we discuss below.

Finally, we note that the decomposition in Proposition 2.8 has consequences beyond the definition of causal separability meriting additional interest: as we show in Chapter 3, it characterises precisely (i.e., providing a necessary and sufficient condition for) quantum circuits with classical control of causal order.

2.2.4 Multipartite witnesses of causal nonseparability

While the previous characterisations provide mathematical descriptions of causally (non)separable process matrices, an important problem is the ability to detect and certify causal nonseparability in practice. Here, we describe how the concept of causal witnesses outlined in Sec. 1.2.3 generalises to the multipartite case, and how the conditions given in the previous subsections allow us to construct causal witnesses for general multipartite scenarios. This will permit a full analysis of the examples in the following section. While the overall approach of causal witnesses—and their formulation as efficiently solvable semidefinite programming (SDP) problems—in the general case mirrors that of the specific scenarios previously studied [11, 13], the validity of the generalisation rests on certain technical details which we prove in Appendix G of the published paper [32].

Propositions 2.5, 2.6, 2.7 and 2.8 allow for the characterisation of the convex cone $\mathcal{W}^{\text{sep}}$ of causally separable processes—or, for the latter two propositions, outer and inner approximations $\mathcal{W}_+^{\text{sep}}$ and $\mathcal{W}_-^{\text{sep}}$ thereof—in terms of Minkowski sums and intersections of linear subspaces and of the cone of positive semidefinite operators $\mathcal{P}$. The set of causal witnesses is then precisely the dual cone of $\mathcal{W}^{\text{sep}}$, $\mathcal{S} = (\mathcal{W}^{\text{sep}})^*$ [11, 13]. A characterisation of $\mathcal{S}$ can, in general, be obtained from the description of $\mathcal{W}^{\text{sep}}$ by using the following duality relations for any two nonempty closed convex cones $\mathcal{C}_1$ and $\mathcal{C}_2$ [51]:

$$(\mathcal{C}_1 + \mathcal{C}_2)^* = \mathcal{C}_1^* \cap \mathcal{C}_2^*, \quad (\mathcal{C}_1 \cap \mathcal{C}_2)^* = \mathcal{C}_1^* + \mathcal{C}_2^* \tag{2.27}$$

(where $\mathcal{C}_1 + \mathcal{C}_2 = \{c_1 + c_2 \mid c_1 \in \mathcal{C}_1, c_2 \in \mathcal{C}_2\}$ is the Minkowski sum of the two cones $\mathcal{C}_1$ and $\mathcal{C}_2$; note that all the cones we shall consider will be nonempty, closed and convex).

Since these cones are convex, the construction of causal witnesses (or of explicit decompositions of causally separable process matrices) can be efficiently performed with SDP also in the multipartite case. The question of whether a given W is causally separable can be reformulated as the optimisation problem of how much white noise can be added to a process matrix before it becomes causally separable, just as in the bipartite case (cf. Sec. 1.2.3).

For more than 3 parties, the witnesses in the set $\mathcal{S}_+ = (\mathcal{W}_+^{\text{sep}})^*$ obtained from the cone $\mathcal{W}_+^{\text{sep}} \supseteq \mathcal{W}^{\text{sep}}$ arising from the necessary condition of Proposition 2.7 are also valid witnesses of $\mathcal{W}^{\text{sep}}$ since $\mathcal{S}_+ \subseteq \mathcal{S}$. On the other hand, by solving the primal SDP problem over the cone $\mathcal{W}_-^{\text{sep}}$ arising from the sufficient condition in Proposition 2.8, one can show the causal separability of any $W \in \mathcal{W}_-^{\text{sep}} \subseteq W^{\text{sep}}$ (through the construction of an explicit causally separable decomposition for W of the form given in Proposition 2.8). Recalling the claim that such process matrices correspond precisely to quantum circuits with classical control of orders (see Chapter 3), the dual cone $\mathcal{S}_-$ is thus the set of "witnesses for no classical control of causal order" (which can thus be found by solving the dual SDP problem).

2.2.5 Examples

In the bipartite scenario and restricted tripartite scenario in which C has no outgoing system, several examples of causally nonseparable process matrices have previously been formulated and studied in detail [4, 11, 13, 15]. Following the approach just outlined, we can readily verify the causal nonseparability of these examples and recover for instance, the results from Refs. [11, 13] for the process matrix of the quantum switch (Eq. (1.39)). In that same restricted tripartite scenario, we have in fact also already looked at another explicit example: the process matrix $W^{\text{act.}}$ (2.3) introduced in Sec. 2.1.3 to show the "activation of causal nonseparability" under OG's definition of causal separability. An explicit witness certifying the causal nonseparability of this process matrix is given in Appendix C of [32], which could thus have been equally well found with the approach of Refs. [11, 13]. Another example of "activation of causal nonseparability" under OG's terminology was given in Ref. [15] in the different tripartite case in which one party, say now A, has only a nontrivial outgoing system, and can thus always be seen as acting first. A witness for this example can be found by solving the dual SDP problem (1.22) using the cone of witnesses corresponding to this restricted tripartite scenario [32].

Of more novel interest is the fourpartite scenario, in which causal separability has not previously been characterised. A particularly interesting and simple example here, which we will illustrate in more detail, is a fourpartite version of the quantum switch, in which a party A(lice) has no incoming system ($d_{A_I} = 1$) and always acts first, while another party D(orothy) has no outgoing system ($d_{D_O} = 1$) and always acts last.

In this version of the switch, Alice prepares the control qubit in some state of her choosing, in general as a function of her input x (in contrast to the tripartite version described in Sec. 1.5, where the control qubit is in a fixed superposition). The target qubit, initially prepared (externally to the 4 parties) in some state $|\psi\rangle$, is then sent to Bob and Charlie, who act in an order that depends on the state of the control qubit: if it is $|0\rangle$ then Bob acts before Charlie ($B \prec C$), while if it is $|1\rangle$ then Charlie acts before Bob ($C \prec B$). If it is in a superposition, then Bob and Charlie act in a superposition of different orders. Finally, both qubits are sent to Dorothy who can perform a measurement on them (for simplicity, we will consider that D simply ignores

the target qubit and thus will trace it out, as this will not change the discussion that follows).[9] Labelling the relevant incoming and outgoing systems (where the superscripts indicate control and target qubits) A_O^c, B_I^t, B_O^t, C_I^t, C_O^t, D_I^t, D_I^c, the process matrix for the quantum switch can be written [11, 15, 38]

$$\begin{aligned} &W^{\text{switch}} = \mathrm{Tr}_{D_I^t} |w\rangle\langle w| \quad \text{with} \\ &|w\rangle = |0\rangle^{A_O^c} |\psi\rangle^{B_I^t} |\mathbb{1}\rangle\!\rangle^{B_O^t C_I^t} |\mathbb{1}\rangle\!\rangle^{C_O^t D_I^t} |0\rangle^{D_I^c} + |1\rangle^{A_O^c} |\psi\rangle^{C_I^t} |\mathbb{1}\rangle\!\rangle^{C_O^t B_I^t} |\mathbb{1}\rangle\!\rangle^{B_O^t D_I^t} |1\rangle^{D_I^c}\,, \end{aligned} \tag{2.28}$$

Note that, while Alice has control over the causal order of the other parties, this switch differs from a classical dynamical control of causal order in that she has coherent quantum control over the control qubit (and thus the causal orders).

In this particular restricted fourpartite scenario, our necessary and sufficient conditions for the causal separability of a process matrix W coincide and reduce to the following proposition (cf. Appendix B.3 (Proposition B7) in the published paper [32]):

Proposition 2.9 (Characterisation of fourpartite causally separable process matrices with $d_{A_I} = 1$ and $d_{D_O} = 1$)**.** *In a fourpartite scenario where party A has no incoming system and party D has no outgoing system, a matrix $W \in \mathcal{L}(\mathcal{H}^{A_O} \otimes \mathcal{H}^{B_{IO}} \otimes \mathcal{H}^{C_{IO}} \otimes \mathcal{H}^{D_I})$ is a valid fourpartite causally separable process matrix (as per Definition 2.4) if and only if*

$$_{[1-A_O]B_{IO}C_{IO}D_I}W = 0 \tag{2.29}$$

and W can be decomposed as

$$W = W_{(A,B,C,D)} + W_{(A,C,B,D)} \tag{2.30}$$

where, for each permutation (X, Y) of the two parties B and C, $W_{(A,X,Y,D)}$ is a positive semidefinite matrix satisfying

$$_{[1-X_O]Y_{IO}D_I}W_{(A,X,Y,D)} = 0, \;\; _{[1-Y_O]D_I}W_{(A,X,Y,D)} = 0\,. \tag{2.31}$$

[9] We note that the quantum switch was also described as a fourpartite process in Ref. [17], with one party acting first, and one acting last. However, in that reference the first party was controlling the target qubit, rather than the control qubit as we consider here. In that case (with the first party controlling the target qubit), the random robustness is increased to 2.767. One could also have here a first party that controls both the target and control qubits (as in Ref. [38]), which further increases the tolerable white noise to 4.686; for simplicity we do not consider this possibility, as our goal here is just to illustrate the role of the control qubit.
Note also that Rubino *et al.* [17] used yet another definition of causal nonseparability, different from the ones discussed in Sec. 2.1, which did not allow for dynamical causal orders. As argued before and discussed in Refs. [14, 15], such a definition is however too restrictive to really characterise processes that are compatible with a well-defined causal order, as one would like the notion of causal separability to do. Nevertheless, it turns out that the witness constructed and experimentally tested in Ref. [17] is not only a witness for fixed (nondynamical) causal orders, but also witnesses causal nonseparability as per our Definition 2.4.

It follows from Proposition 2.9 that the cone $\mathcal{W}^{\text{sep}}$ in the scenario considered here is given by

$$\mathcal{W}^{\text{sep}} = \mathcal{L}_{[1-A_O]B_{IO}C_{IO}D_I} \cap \left[\mathcal{P} \cap \mathcal{L}_{[1-B_O]C_{IO}D_I} \cap \mathcal{L}_{[1-C_O]D_I} + \mathcal{P} \cap \mathcal{L}_{[1-C_O]B_{IO}D_I} \cap \mathcal{L}_{[1-B_O]D_I}\right] \tag{2.32}$$

where $\mathcal{L}_{[1-A_O]B_{IO}C_{IO}D_I}$ denotes the linear subspace of Hermitian matrices $W \in \mathcal{L}(\mathcal{H}^{A_O} \otimes \mathcal{H}^{B_{IO}} \otimes \mathcal{H}^{C_{IO}} \otimes \mathcal{H}^{D_I})$ satisfying ${}_{[1-A_O]B_{IO}C_{IO}D_I}W = 0$, and similarly for $\mathcal{L}_{[1-B_O]C_{IO}D_I}$ etc.

With the duality relations (2.27), it then follows that the dual cone of causal witnesses $\mathcal{S} = (\mathcal{W}^{\text{sep}})^*$ is

$$\mathcal{S} = \mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_{IO}D_I} + (\mathcal{P} + \mathcal{L}^{\perp}_{[1-B_O]C_{IO}D_I} + \mathcal{L}^{\perp}_{[1-C_O]D_I}) \cap (\mathcal{P} + \mathcal{L}^{\perp}_{[1-C_O]B_{IO}D_I} + \mathcal{L}^{\perp}_{[1-B_O]D_I}) \,. \tag{2.33}$$

where we used the fact that $\mathcal{P}$ is self-dual, and where $\mathcal{L}^{\perp}_{[1-A_O]B_{IO}C_{IO}D_I}$ is the orthogonal complement of $\mathcal{L}_{[1-A_O]B_{IO}C_{IO}D_I}$, that is, the linear space of Hermitian operators $S \in \mathcal{L}(\mathcal{H}^{A_O} \otimes \mathcal{H}^{B_{IO}} \otimes \mathcal{H}^{C_{IO}} \otimes \mathcal{H}^{D_I})$ satisfying ${}_{[1-A_O]B_{IO}C_{IO}D_I}S = S$, and similarly for $\mathcal{L}^{\perp}_{[1-B_O]C_{IO}D_I}$ etc.

The causal nonseparability of W^{switch} can thus be verified by solving the dual SDP problem (1.22) and thereby obtaining a witness of its causal nonseparability. Doing so, we find that (up to numerical precision) the random robustness of W^{switch} of 2.343 (note that this does not depend on the choice of initial state of the target qubit, so in solving the SDP problem numerically we can take, e.g., $|\psi\rangle = |0\rangle$). In experimental efforts to measure a witness and verify the causal nonseparability of a process matrix, one may only have access to a restricted set of operations for the parties. Many natural such constraints can also be imposed as SDP constraints, as described in Ref. [13], allowing one to find implementable causal witnesses. A particularly natural such constraint is to restrict B and C's operations to unitary operations (as in the experimental implementation of the tripartite switch in Refs. [16, 19]); we find that the tolerable white noise on W^{switch} to witness its causal nonseparability is reduced, under such a restriction, to 0.746.

It is important to note that if we trace out the last party from W^{switch} (i.e., D_I^c in addition to D_I^t), we obtain

$$\operatorname{Tr}_D W^{\text{switch}} = |0\rangle\langle 0|^{A_O^c} |\psi\rangle\langle\psi|^{B_I^t} |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|^{B_O^t C_I^t} \mathbb{1}^{C_O^t} + |1\rangle\langle 1|^{A_O^c} |\psi\rangle\langle\psi|^{C_I^t} |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|^{C_O^t B_I^t} \mathbb{1}^{B_O^t}, \tag{2.34}$$

which is causally separable since it is of the form of Eq. (2.15) with just the first two terms being nonzero: $\operatorname{Tr}_D W^{\text{switch}} = W_{(A)} = W_{(A,B,C)} + W_{(A,C,B)}$, with $W_{(A,B,C)}$, $W_{(A,C,B)}$ (defined as the first and second terms in Eq. (2.34) above, respectively) and $W_{(A)}$ satisfying the constraints of Eqs. (2.16)–(2.17). This was also the case with the original tripartite version of the quantum switch considered in Sec. 1.5 (in which the control qubit is in the fixed state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$). There, one is left with a simple probabilistic mixture of channels in two different directions after tracing out the last party (cf. Eq. (1.41)). In contrast here, Eq. (2.34) is not compatible with any probabilistic mixture of fixed causal orders: indeed, $W_{(A,B,C)}$ and $W_{(A,C,B)}$ are not valid process matrices, as ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A,B,C)} = -{}_{[1-A_O]B_{IO}C_{IO}}W_{(A,C,B)} \neq 0$ (these terms cancel in

the sum $W_{(A,B,C)} + W_{(A,C,B)}$, so that ${}_{[1-A_O]B_{IO}C_{IO}}W_{(A)} = 0$ as required for $W_{(A)}$ to be a valid process matrix). Rather, $\mathrm{Tr}_D\, W^{\mathrm{switch}}$ is a "classical switch" in which A can incoherently control the causal order between B and C, which thus allows for dynamical causal orders.

2.3 Discussion

In this chapter we presented the results from Ref. [32], where we studied the question of how to generalise the concept of causal (non)separability to the multipartite case. We reviewed several definitions that had been proposed for multipartite scenarios in previous works, namely the definition of causal separability introduced by Araújo *et al.* [11] for a particular tripartite situation, and Oreshkov and Giarmatzi's definitions of causal separability (CS) and extensible causal separability (ECS) [15] for the general multipartite case. We established the equivalence between Araújo *et al.*'s (restricted) definition of causal separability and Oreshkov and Giarmatzi's definition of ECS in the particular tripartite situation considered by Araújo *et al.*, thus linking two *a priori* different definitions for that case. Moreover, by showing that ECS and CS are different in that scenario, we found that the two definitions of causal separability proposed by Araújo *et al.* [11] and by Oreshkov and Giarmatzi [15] were inconsistent, a problem that thus needed to be addressed.

We proposed a new general definition of N-partite causal nonseparability, similar in spirit to the recursive definitions that have been proposed for multipartite causal correlations [14, 15], and more consistent with the fact that the process matrix framework always allows for parties to share additional ancillary systems. Our definition thus avoids some unwanted features of the definition of CS in Ref. [15], such as the "activation" of causal nonseparability by shared entanglement. Moreover, we showed that our definition, although *a priori* different, in fact reduces to the notion of ECS proposed in [15], which also reduces to the definition of Araújo *et al.* [11] in the particular restricted scenario considered there.

We then focused on characterising causally separable process matrices, giving (in the general multipartite case) two conditions—one necessary and one sufficient (Propositions 2.7 and 2.8, respectively)—for a given process matrix to be causally separable. These conditions allowed us to characterise the corresponding sets of process matrices through SDP constraints, and to generalise the tool of witnesses for causal nonseparability to the multipartite case. In the bipartite and tripartite cases, our necessary and sufficient conditions coincide and reduce to those previously described [4, 11, 15]. The principal open question is whether this also holds in the general N-partite case with $N \geq 4$, or whether one of the two is both necessary and sufficient (or if one could derive yet another distinct condition, that would be both necessary and sufficient).

As we show in Chapter 3 (and will present in an upcoming publication [33]), our sufficient condition characterises precisely the processes that can be realised as a quantum circuit with classical control of causal order. If that condition is in fact also necessary, this would thus

confirm the conjecture of Oreshkov and Giarmatzi, that causally separable process matrices (or "extensibly causally separable processes" using their terminology) are those realisable by such "classically controlled quantum circuits" [15]. This would provide more solid founding for our understanding of the notion of causal separability, which would then indeed correspond to our intuition (quantum circuits with possibly dynamical causal orders that are classically controlled). Furthermore, the proof in Chapter 3 (Ref. [33]) would also provide a general explicit construction to realise any given causally separable process matrix in practice.

However, the forms of our necessary and sufficient condition, and the fact that the proof for the necessity of the conditions in the tripartite case does not generalise straightforwardly to more parties, indeed leave open the possibility that our sufficient condition may turn out to not be necessary. If this is the case, it would mean that there exist causally separable process matrices that are not realisable as classically controlled quantum circuits—and which we would not currently know how to realise experimentally. It would certainly be interesting to understand what kind of situations such process matrices correspond to—and if (and how) they can be realised quantum mechanically. This question is reminiscent of the open problem of whether process matrices that allow for the violation of causal inequalities are realisable with "standard" quantum mechanics. Here the question would concern even less extreme situations: causally separable process matrices.

Another question that arises naturally in the multipartite case is whether a given phenomenon is *genuinely* multipartite, in the sense that its occurrence truly requires the coordinated action of a certain number of parties. It would be important for our understanding of multipartite process matrices to define a notion of "genuinely multipartite causal nonseparability", similar to the concept of "genuinely multipartite noncausality" for correlations [35] (see Chapter 5) and analogous to the notions of genuinely multipartite entanglement [52] and nonlocality [53–55]. It would then also be interesting to study whether the definition can be refined to give a hierarchy of degrees of causal nonseparability, similar to the approach in Ref. [35] (Chapter 5) for correlations, and whether the characterisation of the corresponding process matrices and the construction of "witnesses of genuinely multipartite causal nonseparability" are still possible with SDP techniques. These questions are left for further research.

Chapter 3

Quantum circuits with quantum versus classical control of causal order

A manuscript based on the material from this chapter is in preparation:
Julian Wechs, Hippolyte Dourdent, Alastair A. Abbott, and Cyril Branciard. Quantum circuits with quantum versus classical control of causal order, *in preparation (2020)* (Ref. [33]).

The work presented in this chapter has been another central project of my thesis, and I mainly developed the characterisations, proofs and examples presented here.

It is known that process matrices with a fixed causal order correspond to standard, causally ordered quantum circuits. For the bipartite case, we mentioned in Section 1.2.1 that process matrices that are compatible with the causal order $A \prec B$—i.e., that can only generate correlations such that Alice can signal to Bob, but not vice versa—describe precisely the channels with memory from Alice to Bob, as represented in Fig. 1.5. More generally, it has been shown that the subset of process matrices which are compatible with a given causal order corresponds to the set of *quantum combs*, which are fixed-order quantum circuits with open slots into which variable input operations can be inserted. Quantum combs have been described and characterised in Refs. [44, 45, 56]. In this chapter, we go beyond the fixed-order case, and we study more general classes of circuits that can be characterised in terms of their process matrix description.

To lay the basis for the following considerations, we first review quantum circuits with fixed causal order (which we abbreviate to "QC-FO", Sec. 3.3). We recall the description of such circuits in terms of process matrices (or in this case, "quantum combs"), as well as how, given such a process matrix description, one can construct the corresponding quantum circuit. As discussed in previous chapters, we may also consider scenarios where the causal order is not fixed a priori, but established dynamically during the process. This leads us to introduce "quantum circuits with classical control of causal order" (which we refer to as "QC-CC", Sec. 3.4), in

which, as the circuit is being used, the operations are still realised in a well-defined causal order, which is established on the fly. We show how to describe these circuits as process matrices, and characterise the class of process matrices they define. As for the case of fixed causal order, we show how, given a process matrix from this class, one can construct the corresponding circuit. Our study formalises the description of "classically controlled quantum circuits" proposed in Ref. [15]; we note that because the control is classical, the causal order remains well-defined (if not fixed), so that the corresponding processes are causally separable. In fact, the class of process matrices describing QC-CC turns out to coincide precisely with the class of process matrices that satisfy the sufficient condition for multipartite causal separability that we derived in the previous chapter.

It is then natural to consider quantum circuits in which the causal order is controlled coherently. We thus introduce the new class of "quantum circuits with quantum control of causal order" ("QC-QC", Sec. 3.5), which we again fully characterise in terms of their process matrix description. This class contains the quantum switch as a particular example, but also more general types of causally nonseparable quantum processes.

In (Sec. 3.6), we generalise our analysis to probabilistic (post-selected) quantum circuits, and we characterise the classes of "probabilistic process matrices" that can be realised in terms of probabilistic QC-FO, QC-CC and QC-QC.

On the one hand, our results are of foundational importance, because one of the main open questions regarding process matrices is which of them have a physical interpretation. Here, we characterise a large class of process matrices that are realisable through quantum control of causal order. On the other hand, our analysis is also interesting with regard to possible applications of quantum processes with indefinite causal order. In this context, it is natural to study "higher order" quantum information processing tasks that involve the processing of unknown operations. The description of quantum combs in terms of quantum supermaps has been used to formulate and study various tasks of that kind as semidefinite optimisation problems [57]. Based on the characterisations that we provide in this work, this approach can be extended to the more general classes of QC-CC and QC-QC. Our characterisation of the whole class of realisable QC-QCs, together with the characterisation of the corresponding quantum superinstruments, thus allows one to systematically investigate possible applications of quantum processes which go beyond quantum circuits with a well-defined causal order, and for which a concrete realisation scheme exists. We give an example in Sec. 3.7 where we consider some variants of a recently studied black-box discrimination task [58] and show that the success probability obtained using QC-QC is higher than the one for QC-FO and QC-CC.

3.1 The link product

Throughout this chapter, we will again make extensive use of the CJ isomorphism (see Sec. 1.1.1) in order to represent quantum operations in a convenient way, and we will often connect together

various operations in circuits. We will therefore start by introducing a convenient mathematical tool, the *link product* [45, 56], which allows us to obtain the CJ representation of a composition of operations in terms of the CJ representation of the individual operations.

Consider two operators $M_A \in \mathcal{L}(\mathcal{H}^{\mathrm{A}})$ and $M_B \in \mathcal{L}(\mathcal{H}^{\mathrm{B}})$ acting on some tensor product Hilbert spaces $\mathcal{H}^{\mathrm{A}} := \bigotimes_{j\in\mathrm{A}} \mathcal{H}^j$ and $\mathcal{H}^{\mathrm{B}} := \bigotimes_{j\in\mathrm{B}} \mathcal{H}^j$, where some of the constituent Hilbert spaces $\mathcal{H}^j$ may be common to both $\mathcal{H}^{\mathrm{A}}$ and $\mathcal{H}^{\mathrm{B}}$ (i.e., the label sets A and B overlap). The link product of M_A and M_B is then defined as

$$M_A * M_B := \mathrm{Tr}_{\mathrm{A}\cap\mathrm{B}}\left[(M_A \otimes \mathbb{1}^{\mathrm{B}\setminus\mathrm{A}})(M_B^{T_{\mathrm{A}\cap\mathrm{B}}} \otimes \mathbb{1}^{\mathrm{A}\setminus\mathrm{B}})\right] \quad \in \mathcal{L}(\mathcal{H}^{(\mathrm{A}\setminus\mathrm{B})\cup(\mathrm{B}\setminus\mathrm{A})}), \tag{3.1}$$

where the partial trace and partial transpose are taken over the overlapping Hilbert spaces $\mathcal{H}^{\mathrm{A}\cap\mathrm{B}} := \bigotimes_{j\in\mathrm{A}\cap\mathrm{B}} \mathcal{H}^j$, and $\mathbb{1}^{\mathrm{B}\setminus\mathrm{A}}$ ($\mathbb{1}^{\mathrm{A}\setminus\mathrm{B}}$) is the identity operator on $\mathcal{H}^{\mathrm{B}\setminus\mathrm{A}} := \bigotimes_{j\in\mathrm{B}\setminus\mathrm{A}} \mathcal{H}^j$ (on $\mathcal{H}^{\mathrm{A}\setminus\mathrm{B}} := \bigotimes_{j\in\mathrm{A}\setminus\mathrm{B}} \mathcal{H}^j$).

In the case where $M_A = |a\rangle\langle a|$ and $M_B = |b\rangle\langle b|$ are rank-one projectors onto some vectors $|a\rangle \in \mathcal{H}^{\mathrm{A}}$ and $|b\rangle \in \mathcal{H}^{\mathrm{B}}$, we can work on the level of Hilbert space vectors. We have that

$$|a\rangle\langle a| * |b\rangle\langle b| = (|a\rangle * |b\rangle)(\langle a| * \langle b|) \tag{3.2}$$

with the "vector link product"

$$|a\rangle * |b\rangle := (\mathbb{1}^{(\mathrm{A}\setminus\mathrm{B})\cup(\mathrm{B}\setminus\mathrm{A})} \otimes \langle\langle \mathbb{1}|^{\mathrm{A}\cap\mathrm{B}})(\,|a\rangle \otimes |b\rangle\,) \in \mathcal{H}^{(\mathrm{A}\setminus\mathrm{B})\cup(\mathrm{B}\setminus\mathrm{A})}, \tag{3.3}$$

and $\langle a| * \langle b| = (|a\rangle * |b\rangle)^\dagger$.

The link products we defined here allow us to conveniently describe "quantum networks" in which individual quantum operations are connected with each other through certain shared "wires". The basic case is shown in Fig. 3.1. Here, two quantum operations $\mathcal{M}_\mathcal{A} : \mathcal{L}(\mathcal{H}^{X_1}) \to \mathcal{L}(\mathcal{H}^{Y_1 Z})$ and $\mathcal{M}_\mathcal{B} : \mathcal{L}(\mathcal{H}^{X_2 Z}) \to \mathcal{L}(\mathcal{H}^{Y_2})$ (with CJ representations $M_A \in \mathcal{L}(\mathcal{H}^{X_1 Y_1 Z})$ and $M_B \in \mathcal{L}(\mathcal{H}^{X_2 Z Y_2})$) are composed via the shared system Z. The CJ representation of their composition $\mathcal{M}_\mathcal{C} = \mathcal{M}_\mathcal{B} \circ \mathcal{M}_\mathcal{A}$ is then simply given by [45, 56]

$$M_C = M_A * M_B \quad \in \mathcal{L}(\mathcal{H}^{X_1 X_2 Y_1 Y_2}). \tag{3.4}$$

For "pure" quantum operations of the form $\mathcal{M}_\mathcal{A}(\rho) = V_A \rho V_A^\dagger$ and $\mathcal{M}_\mathcal{B}(\rho) = V_B \rho V_B^\dagger$, we can thus use the vector link product in order to obtain the pure CJ representation $|V_C\rangle\rangle$ (cf. Sec. 1.1.1) of their composition in terms of the pure CJ representations $|V_A\rangle\rangle \in \mathcal{H}^{X_1 Y_1 Z}$ and $|V_B\rangle\rangle \in \mathcal{H}^{X_2 Z Y_2}$. It is given by

$$|V_C\rangle\rangle = |V_A\rangle\rangle * |V_B\rangle\rangle \in \mathcal{H}^{X_1 X_2 Y_1 Y_2}. \tag{3.5}$$

Let us review some properties of the link product that will be useful in the following. The link product is commutative (up to a re-ordering of tensor products), and associative when each

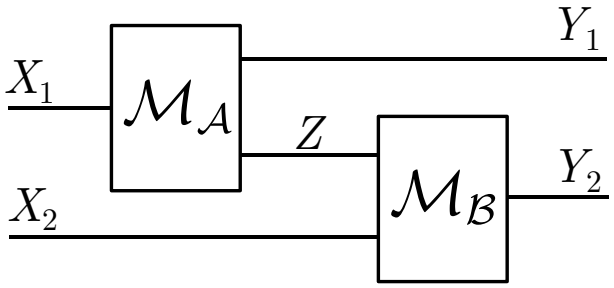

FIGURE 3.1: The link product allows one to obtain the CJ representation of the composition of quantum operations.

Hilbert space $\mathcal{H}^j$ involved in the link product of $M_1 \in \mathcal{L}(\mathcal{H}^{\mathrm{A}_1}), M_2 \in \mathcal{L}(\mathcal{H}^{\mathrm{A}_2}), \ldots, M_n \in \mathcal{L}(\mathcal{H}^{\mathrm{A}_n})$ appears at most twice—i.e., for all $i \neq j \neq k$, $\mathrm{A}_i \cap \mathrm{A}_j \cap \mathrm{A}_k = \emptyset$. In all situations below where we use the link product, this will be the case, which enables us to write the n-fold link product as $M_1 * M_2 * \cdots * M_n$. (or $|m_1\rangle * |m_2\rangle * \cdots * |m_n\rangle$ for the vector link product).

In the particular case where the label sets are disjoint ($\mathrm{A} \cap \mathrm{B} = \emptyset$), we have that $M_A * M_B = M_A \otimes M_B$. If they are identical ($\mathrm{A} = \mathrm{B}$), we have that $M_A * M_B = \mathrm{Tr}[M_A M_B^T]$. The link product of positive semidefinite operators is positive semidefinite.

Other useful properties are $M_A * \mathbb{1}^{\mathrm{B}} = (\mathrm{Tr}_{\mathrm{A}\cap\mathrm{B}}\, M_A) \otimes \mathbb{1}^{\mathrm{B}\backslash\mathrm{A}}$ (if $\mathrm{B} \subseteq \mathrm{A}$, $M_A * \mathbb{1}^{\mathrm{B}} = \mathrm{Tr}_{\mathrm{B}}\, M_A$), and, for $\mathrm{C} \subseteq \mathrm{B}\backslash\mathrm{A}$, $\mathrm{Tr}_{\mathrm{C}}[M_A * M_B] = M_A * (\mathrm{Tr}_{\mathrm{C}}\, M_B)$.

3.2 Process matrices as quantum supermaps

In this chapter, we will use the process matrix formalism in a slightly different way than above. Previously we defined process matrices as objects that take local CP maps to probabilities. Here, we will instead use process matrices to describe "higher order transformations" or *quantum supermaps* [59], which take the local CP maps to a new CP map from some "global past space" to some "global future space" (as in Ref. [60]). The two descriptions are equivalent, but the second point of view is more convenient for our purposes in this chapter.

In the scenario that we consider here, we have N local quantum operations (CP maps) $\mathcal{A}_k$ ($k \in \mathcal{N} := \{1, \ldots, N\}$), with (possibly different) input and output Hilbert spaces $\mathcal{H}^{A_k^I}$ and $\mathcal{H}^{A_k^O}$, respectively, of dimensions d_k^I and d_k^O, respectively (all Hilbert spaces throughout the chapter are again taken to be finite-dimensional), mapping some incoming physical system to some outgoing system. That is, the $\mathcal{A}_k$'s are (any) CP maps $\mathcal{A}_k : \mathcal{L}(\mathcal{H}^{A_k^I}) \to \mathcal{L}(\mathcal{H}^{A_k^O})$. Similarly to above, we will use the notations $\mathcal{H}^{A_k^{IO}} := \mathcal{H}^{A_k^I A_k^O}$, and $\mathcal{H}^{A_\mathcal{K}^{IO}} := \bigotimes_{k\in\mathcal{K}} \mathcal{H}^{A_k^{IO}}$ for any subset $\mathcal{K} \subseteq \mathcal{N}$.

A quantum supermap [59] is a multilinear transformation $(\mathcal{A}_1, \ldots, \mathcal{A}_N) \to \mathcal{M}$ that takes the N quantum operations $\mathcal{A}_k$ to a new CP map $\mathcal{M} : \mathcal{L}(\mathcal{H}^P) \to \mathcal{L}(\mathcal{H}^F)$ from some Hilbert space $\mathcal{H}^P$ in the "global past" of all operations $\mathcal{A}_k$ to some Hilbert space $\mathcal{H}^F$ in their "global future", with CJ

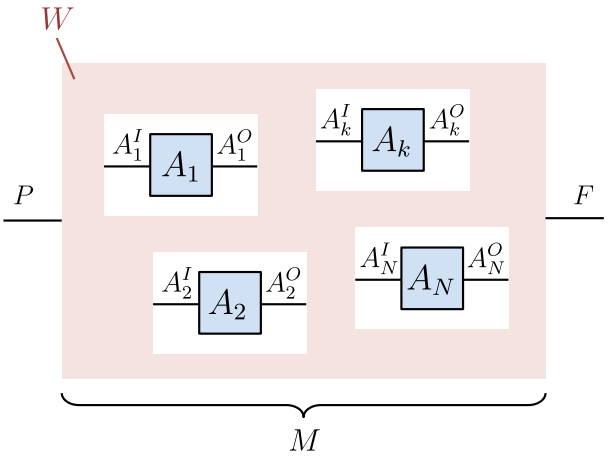

FIGURE 3.2: A process matrix W represents a quantum supermap that takes N quantum operations—i.e., CP maps—$\mathcal{A}_k$ with input and output Hilbert spaces $\mathcal{H}^{A_k^I}$ and $\mathcal{H}^{A_k^O}$, respectively, to a new CP map $\mathcal{M}$ with an input Hilbert space $\mathcal{H}^P$ in the "global past" of all operations $\mathcal{A}_k$ and an output Hilbert space $\mathcal{H}^F$ in their "global future" [4, 60]. The CJ representation M of the global map $\mathcal{M}$ is obtained from the CJ representations A_k of the maps $\mathcal{A}_k$ according to Eq. (3.6), as a function of the process matrix W.

representation[1] $M \in \mathcal{L}(\mathcal{H}^{PF})$; see Fig. 3.2. For now[2], we consider deterministic supermaps that, although possibly taking probabilistic operations as inputs, do not themselves produce random outcomes; that is, we require that the global transformation $\mathcal{M}$ should be a CPTP map whenever the local operations are CPTP maps. Similarly to above, we allow for additional ancillary spaces that the local operations can act upon. That is, we impose that even for extended CP maps $\mathcal{A}_k^e : \mathcal{L}(\mathcal{H}^{A_k^I A_k^{I'}}) \to \mathcal{L}(\mathcal{H}^{A_k^O A_k^{O'}})$, involving some additional input (output) space $\mathcal{H}^{A_k^{I'}}$ ($\mathcal{H}^{A_k^{O'}}$), the supermap (tensored with the identity map on the ancillas) should again output a valid CP map (from $\mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{I'}})$ to $\mathcal{L}(\mathcal{H}^{FA_{\mathcal{N}}^{O'}})$).

Quantum supermaps, as described here, are in one-to-one correspondence with process matrices $W \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$. That is, they can be represented as [60]:

$$M = \mathrm{Tr}_{A_{\mathcal{N}}^{IO}} \left[W (A_1^T \otimes \cdots \otimes A_N^T \otimes \mathbb{1}^{PF}) \right] = W * (A_1 \otimes \cdots \otimes A_N) \quad \in \mathcal{L}(\mathcal{H}^{PF}), \tag{3.6}$$

where $W \geq 0$, $\mathrm{Tr}\, W = d_P\, d_{\mathcal{N}}^O$ and $W \in \mathcal{L}^{P \prec \mathcal{N} \prec F} \subset \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$, the subspace of valid process matrices in the $(N+2)$-partite scenario involving N parties with incoming (outgoing) Hilbert spaces $\mathcal{H}^{A_k^I}$ ($\mathcal{H}^{A_k^O}$), as well as an additional "global past" party with an outgoing space $\mathcal{H}^P$ (and a trivial, one-dimensional incoming space) and a "global future" party with incoming space $\mathcal{H}^F$ (and a trivial, one-dimensional outgoing space).[3] In the second step of Eq. (3.6) we used the link product notation defined above.

In the following three sections, we will study some particular classes of quantum supermaps, and their descriptions in terms of process matrices. We will start with quantum circuits with fixed

[1] Again, we will often refer to a map or its CJ representation indifferently.
[2] That is, in Secs. 3.3—3.5. We will then extend our anaysis to probabilistic supermaps in Sec. 3.6.
[3] An explicit characterisation of this subspace can be obtained as a special case of Eq. (2.11).

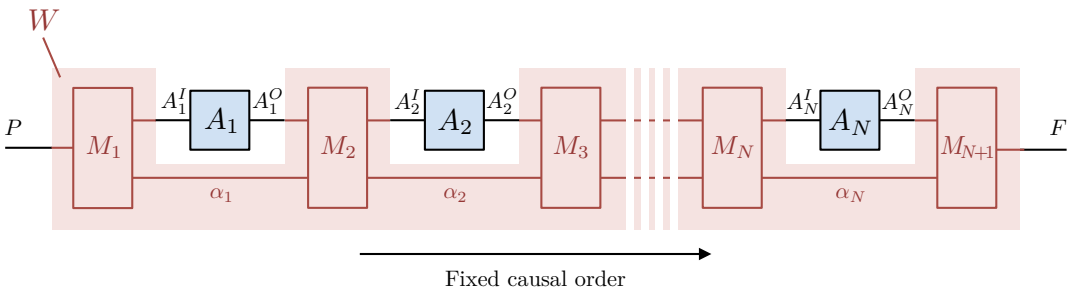

FIGURE 3.3: Quantum circuit with the fixed order of operations $(A_1, A_2, \ldots, A_N)$—or equivalently here, a "quantum comb" [45, 56]. Its process matrix representation is given by $W = M_1 * M_2 * \cdots * M_{N+1}$, as in Proposition 3.2.

causal order.

3.3 Quantum circuits with fixed causal order

Quantum circuits with fixed causal order (QC-FOs) have been studied in details before [45, 56]. Here we simply recall their description and characterisation in terms of process matrices—also called in that case "quantum combs" [45, 56]—so as to make the chapter self-contained and to set up the stage for the following study of quantum circuits without a fixed causal order.

3.3.1 Description

We consider a quantum circuit in which the "external" CP maps $A_1, \ldots, A_N$ are combined (so as to define the global map M, as described above) in a fixed order, which we denote $(A_1, A_2, \ldots, A_N)$ (meaning that A_1 is applied before A_2 which is applied before A_3, ...(etc) ..., applied before A_N). That is, the CP maps are inserted into a causally ordered circuit composed of "internal" circuit operations that take the output of each CP map to the input of the subsequent one. As we want the circuit to give rise to a deterministic supermap, the internal circuit operations are taken to be deterministic quantum channels (CPTP maps). Moreover, they may involve additional ancillary systems or "memories" $\alpha_1, \ldots, \alpha_N$ that are entangled with the system the CP maps act upon.

Such a circuit is depicted on Fig. 3.3.

The circuit starts with a CPTP map (a "quantum channel") $\mathcal{M}_1 : \mathcal{L}(\mathcal{H}^P) \to \mathcal{L}(\mathcal{H}^{A_1^I \alpha_1})$, with CJ representation $M_1 \in \mathcal{L}(\mathcal{H}^{P A_1^I \alpha_1})$, which takes as input some state in the "global past" Hilbert space $\mathcal{H}^P$ and which outputs a state of the system in the input Hilbert space $\mathcal{H}^{A_1^I}$ of the first external CP map A_1, possibly entangled to some ancillary system in some Hilbert space $\mathcal{H}^{\alpha_1}$.

Then for $1 \leq n \leq N-1$, the output state of the CP map A_n in the Hilbert space $\mathcal{H}^{A_n^O}$ (possibly entangled to some ancillary system in some Hilbert space $\mathcal{H}^{\alpha_n}$) is mapped to the input state

of the CP map A_{n+1} in the Hilbert space $\mathcal{H}^{A^I_{n+1}}$ (possibly entangled to some ancillary system in some Hilbert space $\mathcal{H}^{\alpha_{n+1}}$) via a CPTP map $\mathcal{M}_{n+1} : \mathcal{L}(\mathcal{H}^{A^O_n \alpha_n}) \to \mathcal{L}(\mathcal{H}^{A^I_{n+1}\alpha_{n+1}})$, with CJ representation $M_{n+1} \in \mathcal{L}(\mathcal{H}^{A^O_n \alpha_n A^I_{n+1}\alpha_{n+1}})$.

After the last operation A_N, a CPTP map $\mathcal{M}_{N+1} : \mathcal{L}(\mathcal{H}^{A^O_N \alpha_N}) \to \mathcal{L}(\mathcal{H}^F)$, with CJ representation $M_{N+1} \in \mathcal{L}(\mathcal{H}^{A^O_N \alpha_N F})$, takes the output state of A_N, together with the ancillary state in $\mathcal{H}^{\alpha_N}$, to the global output state of the full circuit, in the "global future" Hilbert space $\mathcal{H}^F$.

Note that this description includes as a particular case the situation where the CP maps A_n (or just some of them) are used in parallel. Composing some CP maps in parallel is equivalent to composing them sequentially in an arbitrary order, with internal circuit operations that send the input systems to the respective CP maps one at a time, while passing on the outputs of the preceding CP maps, as well as the inputs of the subsequent ones, via the ancillary systems (for an example, see subsection 3.3.3).

Let us consider that we input some state $\rho \in \mathcal{L}(\mathcal{H}^P)$ into the circuit. Using the link product notation defined above, it is easy to see, recursively, that for $1 \leq n \leq N+1$, the (non-normalised) joint quantum state of the system (in $\mathcal{H}^{A^I_n}$) and the ancilla (in $\mathcal{H}^{\alpha_n}$) right after the internal CPTP map M_n is

$$M_n * A_{n-1} * \cdots * M_2 * A_1 * M_1 * \rho = (M_1 * M_2 * \cdots * M_n) * (\rho \otimes A_1 \otimes \cdots \otimes A_{n-1}), \tag{3.7}$$

where we used the properties of the link product to reorder the terms (and to change some of the link products into tensor products).

All internal circuit operations $M_1, \ldots, M_{n+1}, \ldots M_{N+1}$ need to be trace-preserving. Using Eq. (3.7), we can write this condition as

$$\begin{aligned} \mathrm{Tr}[M_1 * \rho] = \mathrm{Tr}[\rho] \quad \text{and} \\ \mathrm{Tr}[(M_1 * M_2 * \cdots * M_{n+1}) * (\rho \otimes A_1 \otimes \cdots \otimes A_n)] \\ = \mathrm{Tr}[(M_1 * M_2 * \cdots * M_n) * (\rho \otimes A_1 \otimes \cdots \otimes A_n)]. \end{aligned} \tag{3.8}$$

for $1 \leq n \leq N$. That is, the trace of the state right after the operation M_1 (M_{n+1}) needs to be equal to the trace of the state right before the operation.[4]

[4]Note that the trace preserving condition, as written in Eq. (3.8) is more general than imposing directly that $\mathrm{Tr}_{A^I_1\alpha_1} M_1 = \mathbb{1}^P$, $\mathrm{Tr}_{A^I_{n+1}\alpha_{n+1}} M_{n+1} = \mathbb{1}^{A^O_n \alpha_n}$ for $1 \leq n \leq N-1$, and $\mathrm{Tr}_F M_{N+1} = \mathbb{1}^{A^O_N \alpha_N}$ (which are the "standard" trace-preserving conditions, when we impose that the operations under consideration act trace-preservingly on all possible states in their input space; cf. Sec. 1.1.1). In Eq. (3.8), we require that each circuit operation M_n acts trace-preservingly on any input state it can receive in a given QC-FO, and this may not necessarily be *any* state in $\mathcal{L}(\mathcal{H}^{A^O_n \alpha_n})$. In other words, there may be situations where, at some given time step t_n, not the entire space $\mathcal{H}^{A^O_n \alpha_n}$ is "populated". In the case of QC-FO, this remark is rather trivial (the only possibility for such a situation to occur is if only some subspace of the full ancillary space $\mathcal{H}^{\alpha_n}$ is "used" at a given time-step t_n, in which case we could just restrict our considerations to this subspace). This issue appears however also in the case of QC-QC, where the situation is more complex (cf. footnote 18 and the remark after the proof of Proposition 3.9).

Definition 3.1. Any circuit that abides by the previous description, as represented in Fig. 3.3, with the internal circuit operations M_n (for $1 \leq n \leq N+1$) satisfying the trace-preserving conditions in Eq. (3.8) for any initial state $\rho \in \mathcal{L}(\mathcal{H}^P)$ and CP maps $A_k \in \mathcal{L}(\mathcal{H}^{A_k^{IO}})$, is called a *quantum circuit with fixed causal order* (QC-FO).

Considering the output state (3.7) for $n = N+1$, it is immediate to see, according to the discussion in Sec. 3.2, that:

Proposition 3.2. *The process matrix corresponding to the quantum circuit of Fig. 3.3, with the fixed causal order* $(A_1, A_2, \ldots, A_N)$*, is*

$$W = M_1 * M_2 * \cdots * M_{N+1} \quad \in \mathcal{L}(\mathcal{H}^{P A_{\mathcal{N}}^{IO} F}). \tag{3.9}$$

In this case, W is precisely a "quantum comb", as defined in Refs. [45, 56].

3.3.2 Characterisation

The following proposition characterises the process matrices of QC-FO (cf also Refs. [44, 45]):

Proposition 3.3. *For a given matrix* $W \in \mathcal{L}(\mathcal{H}^{P A_{\mathcal{N}}^{IO} F})$*, let us define the reduced matrices (for* $1 \leq n \leq N$*, and relative to the fixed order* $(A_1, A_2, \ldots, A_N)$*)*

$$W_{(n)} := \frac{1}{d_n^O d_{n+1}^O \cdots d_N^O} \operatorname{Tr}_{A_n^O A_{\{n+1,\ldots,N\}}^{IO} F} W \in \mathcal{L}(\mathcal{H}^{P A_{\{1,\ldots,n-1\}}^{IO} A_n^I}).$$

The process matrix $W \in \mathcal{L}(\mathcal{H}^{P A_{\mathcal{N}}^{IO} F})$ *of a quantum circuit with the fixed causal order* $(A_1, A_2, \ldots, A_N)$ *is a positive semidefinite matrix such that its reduced matrices* $W_{(n)}$ *just defined satisfy*

$$\begin{aligned} &\operatorname{Tr}_F W = W_{(N)} \otimes \mathbb{1}^{A_N^O}, \\ &\forall\, n = 1, \ldots, N-1, \quad \operatorname{Tr}_{A_{n+1}^I} W_{(n+1)} = W_{(n)} \otimes \mathbb{1}^{A_n^O}, \\ &\text{and} \quad \operatorname{Tr}_{A_1^I} W_{(1)} = \mathbb{1}^P. \end{aligned} \tag{3.10}$$

Conversely, any positive semidefinite matrix $W \in \mathcal{L}(\mathcal{H}^{P A_{\mathcal{N}}^{IO} F})$ *whose reduced matrices* $W_{(n)}$ *satisfy the constraints of Eq. (3.10) is the process matrix of a quantum circuit with the fixed causal order* $(A_1, A_2, \ldots, A_N)$*.*

Equivalent results were already proven in Refs. [44, 45]. We give a self-contained proof here, which we will later generalise to the other classes of circuits.

Proof. Writing out the link product in Eq. (3.8), we obtain

$$\begin{aligned}\operatorname{Tr}[\operatorname{Tr}_{A_1^I\alpha_1} M_1 \cdot \rho^T] &= \operatorname{Tr}[\rho] \quad \text{and} \\ \operatorname{Tr}[\operatorname{Tr}_{A_{n+1}^I\alpha_{n+1}}(M_1 * M_2 * \cdots * M_{n+1}) \cdot (\rho \otimes A_1 \otimes \cdots \otimes A_n)^T] \\ &= \operatorname{Tr}[(\operatorname{Tr}_{\alpha_n}(M_1 * M_2 * \cdots * M_n) \otimes \mathbb{1}^{A_n^O}) \cdot (\rho \otimes A_1 \otimes \cdots \otimes A_n)^T] \end{aligned} \tag{3.11}$$

(with $A_{N+1}^I \coloneqq F$ and trivial α_{N+1}). Since we require Eq. (3.11) to hold for any initial states $\rho \in \mathcal{L}(\mathcal{H}^P)$ and CP maps $A_k \in \mathcal{L}(\mathcal{H}^{A_k^{IO}})$, this is equivalent to

$$\begin{aligned}\operatorname{Tr}_{A_1^I\alpha_1} M_1 &= \mathbb{1}^P \quad \text{and} \\ \operatorname{Tr}_{A_{n+1}^I\alpha_{n+1}}(M_1 * M_2 * \cdots * M_{n+1}) &= \operatorname{Tr}_{\alpha_n}(M_1 * M_2 * \cdots * M_n) \otimes \mathbb{1}^{A_n^O} \end{aligned} \tag{3.12}$$

Consider the process matrix $W = M_1 * M_2 * \cdots * M_{N+1}$ of a QC-FO as per Proposition 3.2, with the operations M_n satisfying the conditions Eq. (3.12). W is clearly positive semidefinite, as all $M_n \geq 0$. Defining $W_{(N+1)} \coloneqq W$, the reduced matrices $W_{(n)}$ defined in Proposition 3.3 are obtained recursively (from $n = N$, down to $n = 1$) as $W_{(n)} = \frac{1}{d_n^O} \operatorname{Tr}_{A_n^O A_{n+1}^I} W_{(n+1)}$. Similarly, Eq. (3.12) implies that $\operatorname{Tr}_{\alpha_n}(M_1 * \cdots * M_n) = \frac{1}{d_n^O} \operatorname{Tr}_{A_n^O A_{n+1}^I}[\operatorname{Tr}_{\alpha_{n+1}}(M_1 * \cdots * M_n * M_{n+1})]$. Since $W_{(n)}$ and $\operatorname{Tr}_{\alpha_n}(M_1 * \cdots * M_n)$ are equal for $n = N + 1$ (with a trivial $\mathcal{H}^{\alpha_{N+1}}$) and satisfy the same recursive property, it follows that they are the same for all $n = 1, \ldots, N + 1$:

$$W_{(n)} = \operatorname{Tr}_{\alpha_n}(M_1 * \cdots * M_n). \tag{3.13}$$

The constraints of Eq. (3.10) are then simply equivalent to (and therefore readily implied by) the trace-preserving conditions of Eq. (3.12).

To prove the converse, we are going to show that for any given positive semidefinite matrix $W \in \mathcal{L}(\mathcal{H}^{PA_N^{IO}F})$ that satisfies Eq. (3.10), with reduced matrices $W_{(n)}$ as defined in Proposition 3.3, we can find positive semidefinite matrices $M_1 \in \mathcal{L}(\mathcal{H}^{PA_1^I\alpha_1})$, ..., $M_n \in \mathcal{L}(\mathcal{H}^{A_n^O\alpha_n A_{n+1}^I\alpha_{n+1}})$,..., $M_{N+1} \in \mathcal{L}(\mathcal{H}^{A_N^O\alpha_N F})$, whose link product is the required $W_{(n)}$ for all n, i.e. that satisfy $W_{(n)} = \operatorname{Tr}_{\alpha_n}(M_1 * M_2 * \cdots * M_n)$ for $1 \leq n \leq N + 1$ (and that therefore satisfy the trace-preserving condition (3.8), since the matrices $W_{(n)}$ satisfy Eq. (3.12)).

We will give such circuit operations explicitly.

Since $W \geq 0$, all $W_{(n)} \geq 0$ as well (for $1 \leq n \leq N + 1$), and they admit a spectral decomposition of the form

$$W_{(n)} = \sum_i |w_{(n)}^i\rangle\langle w_{(n)}^i| \tag{3.14}$$

for some eigenbasis consisting of $r_n \coloneqq \operatorname{rank} W_{(n)}$ (nonnormalised and nonzero) orthogonal vectors $|w_{(n)}^i\rangle \in \mathcal{H}^{PA_{\{1,\ldots,n-1\}}^{IO}A_n^I}$. Let us introduce, for each $n = 1, \ldots, N + 1$, some r_n-dimensional

ancillary Hilbert space $\mathcal{H}^{\alpha_n}$ with its computational basis $\{|i\rangle^{\alpha_n}\}_{i=1}^{r_n}$, and define

$$|w_{(n)}\rangle := \sum_i |w^i_{(n)}\rangle \otimes |i\rangle^{\alpha_n} \ \in \mathcal{H}^{PA^{IO}_{\{1,\dots,n-1\}}A^I_n\alpha_n}, \tag{3.15}$$

such that $W_{(n)} = \mathrm{Tr}_{\alpha_n} |w_{(n)}\rangle\langle w_{(n)}|$.

We then define $|V_1\rangle\rangle := |w_{(1)}\rangle$, and, for $1 \leq n \leq N$,

$$|V_{n+1}\rangle\rangle := \sum_i \frac{|i\rangle^{\alpha_n} \langle w^i_{(n)}|}{\langle w^i_{(n)}|w^i_{(n)}\rangle} \otimes \mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}} |w_{(n+1)}\rangle \ \in \mathcal{H}^{A^O_n\alpha_n A^I_{n+1}\alpha_{n+1}}. \tag{3.16}$$

We then find that, for $1 \leq n \leq N$,

$$|w_{(n)}\rangle * |V_{n+1}\rangle\rangle = \sum_i \frac{|w^i_{(n)}\rangle\langle w^i_{(n)}|}{\langle w^i_{(n)}|w^i_{(n)}\rangle} \otimes \mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}} |w_{(n+1)}\rangle = |w_{(n+1)}\rangle\,. \tag{3.17}$$

The last step follows from the fact that the projector $\Pi_{(n)} := \sum_i \frac{|w^i_{(n)}\rangle\langle w^i_{(n)}|}{\langle w^i_{(n)}|w^i_{(n)}\rangle}$, tensored with $\mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}}$, acts as the identity on $|w_{(n+1)}\rangle$[5].

It then follows recursively, for $1 \leq n \leq N+1$ that $|w_{(n)}\rangle = |V_1\rangle\rangle * \cdots * |V_n\rangle\rangle$, and therefore $W_{(n)} = \mathrm{Tr}_{\alpha_n}(M_1 * M_2 * \cdots * M_n)$, where $M_n = |V_n\rangle\rangle\langle\langle V_n|$ (and $W = M_1 * M_2 * \cdots * M_{N+1}$, with $M_{N+1} = \mathrm{Tr}_{\alpha_{N+1}} |V_{N+1}\rangle\rangle\langle\langle V_{N+1}|$). This completes the proof of Proposition 3.3. □

Note that the realisation that we constructed in the above proof is not unique, and different circuits may be described by the same process matrix. In particular, a given process matrix of this class may be compatible with different fixed causal orders. An example is given in the next subsection.

Finally, note that the set of QC-FO process matrices characterised by Proposition 3.3 corresponds precisely to the set of process matrices in the subspace $\mathcal{L}^{P\prec A_1\prec\cdots\prec A_N\prec F} \subset \mathcal{L}^{P\prec\mathcal{N}\prec F}$ that we already came across earlier (cf. Secs. 1.4.1 and 2.2.1). That is, the form of circuits considered above, obtained (constructively) by combining the CP maps A_k one after the other, and linking them by quantum operations possibly also acting on some ancillary systems ("channels with memory"), are the most general quantum supermaps that respect the fixed causal order $(A_1, A_2, \dots, A_N)$, i.e., the most general supermaps that allow only for correlations compatible with the fixed order $(A_1, A_2, \dots, A_N)$ to be established.

[5]This can be seen as follows. Denoting by $\Pi^\perp_{(n)} := \mathbb{1}^{PA^{IO}_{\{1,\dots,n-1\}}A^I_n} - \Pi_{(n)}$ the orthogonal projector to $\Pi_{(n)}$ in $\mathcal{H}^{PA^{IO}_{\{1,\dots,n-1\}}A^I_n}$, we have

$$\begin{aligned}\mathrm{Tr}[(\Pi^\perp_{(n)} \otimes \mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}})\, |w_{(n+1)}\rangle\langle w_{(n+1)}|] &= \mathrm{Tr}[(\Pi^\perp_{(n)} \otimes \mathbb{1}^{A^O_n})\, \mathrm{Tr}_{A^I_{n+1}} W_{(n+1)}] \\ &= \mathrm{Tr}[(\Pi^\perp_{(n)} \otimes \mathbb{1}^{A^O_n})(W_{(n)} \otimes \mathbb{1}^{A^O_n})] = 0,\end{aligned} \tag{3.18}$$

so that $(\Pi^\perp_{(n)} \otimes \mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}})\, |w_{(n+1)}\rangle = 0$, and therefore $(\Pi_{(n)} \otimes \mathbb{1}^{A^O_n A^I_{n+1}\alpha_{n+1}})\, |w_{(n+1)}\rangle = |w_{(n+1)}\rangle$.

3.3.3 Examples

As a simple example of a QC-FO, consider the case where two CP maps A_1 and A_2 are applied successively to the input state from the global past, and then the output is sent to the global future. This scenario corresponds to a QC-FO with the order (A_1, A_2), with circuit operations that are identity channels, and that do not involve additional ancillas. The corresponding process matrix as per Proposition 3.2 is

$$W_C = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{PA_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O A_2^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O F} \tag{3.19}$$

and it is straightforward to verify that it satisfies the characterisation of Proposition 3.3.

Another example is a scenario where a bipartite state is prepared in the global past and sent in parallel to A_1 and A_2, whose output is then sent to the global future. The process matrix corresponding to that situation is

$$W_S = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_1 A_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_2 A_2^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O F_1} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O F_2} \tag{3.20}$$

with the global past $\mathcal{H}^P = \mathcal{H}^{P_1} \otimes \mathcal{H}^{P_2}$ and the global future $\mathcal{H}^F = \mathcal{H}^{F_1} \otimes \mathcal{H}^{F_2}$ composed of two spaces each. W_S is the process matrix of a QC-FO compatible with both orders (A_1, A_2) and (A_2, A_1), as can be verified from Proposition 3.3. Indeed, a realisation of that process matrix as a QC-FO conforming to the description above with the causal order (A_1, A_2) is given through the circuit operations $M_1 = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_1 A_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_2 \alpha_1}$, $M_2 = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O \alpha_2} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{\alpha_1 A_2^I}$ and $M_3 = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{\alpha_2 F_1} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O F_2}$, corresponding to a circuit where the state in P_1 is sent directly to the input of A_1, and its output is transmitted to F_1 via the ancilla α_2, while A_2 receives its input via the ancilla α_1, and its output is directly transmitted to F_2. A realisation of W_S in terms of a circuit with the order (A_2, A_1) is given through the circuit operations $M_1' = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_1 \alpha_1} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_2 A_2^I}$, $M_2' = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{\alpha_1 A_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O \alpha_2}$ and $M_3' = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O F_1} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{\alpha_2 F_2}$. It can easily be checked that $M_1 * M_2 * M_3 = W_S = M_1' * M_2' * M_3'$.

This example illustrates the fact that a given process matrix may have different realisations, and, more particularly, that process matrices in the class of QC-FO may be compatible with different causal orders.

3.4 Quantum circuits with classical control of causal order

3.4.1 Description

We now move on to a more general scenario in which the causal order between the N quantum operations A_k is still well-defined, but not fixed from the outset. Instead, it is established dynamically, with the operations in the past determining the causal order of the operations in the future. In order to make sure that we still end up with a valid quantum supermap, the

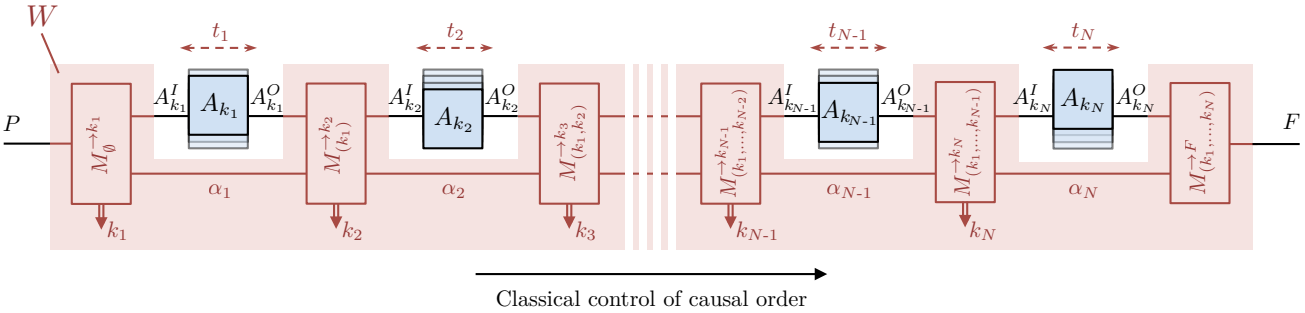

FIGURE 3.4: Quantum circuit with classical control of causal order (QC-CC). The causal order is controlled, and established dynamically, by the outcomes k_n of the operations $M^{\to k_n}_{(k_1,\dots,k_{n-1})}$, represented by the double-stroke arrows. The superimposed boxes A_{k_n} at each time slot t_n indicate that any operation A_k can *a priori* be applied at any time slot; we illustrate here the case where the ("causal") order of operations ends up being $(k_1, k_2, \dots, k_N)$. The process matrix W that represents the circuit above is a (classical) combination of the different contributions corresponding to the different (dynamically established) orders $(k_1, k_2, \dots, k_N)$; see Proposition 3.5.

protocol that we describe here is constructed in such a way that each of the operations A_k is applied once and only once. This ensures in particular that the overall process is linear in the external input operations A_k.

We consider a quantum circuit as represented schematically in Fig. 3.4, with N "open slots" at different time slots t_n $(1 \leq n \leq N)$. At each time slot, one (and only one) operation A_k will be applied (and each operation A_k can *a priori* be applied at any time slot t_n). Here, compared to the previous case of QC-FOs, it is however not pre-defined, which operation is applied at each time slot t_n.

Before the first time slot t_1, and between each pair of consecutive time slots t_n, t_{n+1} (for $1 \leq n \leq N-1$), the circuit applies some quantum operation, which determines, in particular, which operation A_k shall be applied next (while also transforming the input state). In the end, a last operation is applied, that takes the output of the operation applied at the last time slot to the global output of the circuit in $\mathcal{H}^F$ ("global future"). The operations applied by the circuit thus need to admit several possible classical outcomes, according to which the subsequent operation is determined, while also mapping the output state of the preceding party to the input state of the succeeding one. In other words, they need to be quantum instruments.

Before t_1, the circuit thus applies some quantum instrument $\{M_\emptyset^{\to k_1}\}_{k_1 \in \mathcal{N}}$, where each operation $M_\emptyset^{\to k_1} \in \mathcal{L}(\mathcal{H}^{PA^I_{k_1}\alpha_1})$, attached to the classical output value k_1 that tells which operation shall be applied first, takes the "global past" system in $\mathcal{H}^P$ to the incoming space $\mathcal{H}^{A^I_{k_1}}$ of the operation A_{k_1} and (possibly) also to some ancillary system in some Hilbert space $\mathcal{H}^{\alpha_1}$.[6,7]

[6] In principle, this Hilbert space could also be conditioned on k_1. We can however always embed the different $\mathcal{H}^{\alpha(k_1)}$ into a larger ancillary space $\mathcal{H}^{\alpha_1}$, which is common to all k_1, and therefore assume one ancillary space at the first time step without loss of generality (and similarly for the subsequent time steps).

[7] Note that the various operations $M_\emptyset^{\to k_1}$ that form the instrument $\{M_\emptyset^{\to k_1}\}_{k_1 \in \mathcal{N}}$ (and similarly for the operations $M^{\to k_{n+1}}_{(k_1,\dots,k_n)}$ that form the instruments $\{M^{\to k_{n+1}}_{(k_1,\dots,k_n)}\}_{k_{n+1} \in \mathcal{N} \setminus \{k_1,\dots,k_n\}}$ below) do not have the same output spaces. This is however not a problem; if necessary one can always define a common output space into which one embeds all their respective output spaces, as we will do in the next section.

Between the time slots t_n and t_{n+1}, for $1 \leq n \leq N-1$, conditioned on the sequence $(k_1, \ldots, k_n)$ of operations that have already been performed,[8] the circuit then applies some quantum instrument $\{M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}\}_{k_{n+1}\in\mathcal{N}\setminus\{k_1,\ldots,k_n\}}$, where each operation $M^{\to k_{n+1}}_{(k_1,\ldots,k_n)} \in \mathcal{L}(\mathcal{H}^{A^O_{k_n}\alpha_n A^I_{k_{n+1}}\alpha_{n+1}})$, attached to the (classical) output value k_{n+1}, takes the output system of the last performed operation A_{k_n}, together with the ancillary system in $\mathcal{H}^{\alpha_n}$, to the incoming space $\mathcal{H}^{A^I_{k_{n+1}}}$ of some yet unperformed operation $A_{k_{n+1}}$ (hence with $k_{n+1} \in \mathcal{N}\setminus\{k_1, \ldots, k_n\}$) and some ancillary system in some Hilbert space $\mathcal{H}^{\alpha_{n+1}}$.[9]

After the time slot t_N all operations A_k have been performed (once and only once indeed, as required). The circuit then applies some operation $M^{\to F}_{(k_1,\ldots,k_N)} \in \mathcal{L}(\mathcal{H}^{A^O_{k_N}\alpha_N F})$ that takes the output system of the last operation A_{k_N}, together with the ancillary system in $\mathcal{H}^{\alpha_N}$, to the "global future" output space $\mathcal{H}^F$ of the circuit.

Consider inputting some quantum state $\rho \in \mathcal{L}(\mathcal{H}^P)$ into the circuit. In the case where the operations $M^{\to k_1}_{\emptyset}$, $M^{\to k_2}_{(k_1)}$, $M^{\to k_3}_{(k_1,k_2)}$, $\ldots$, $M^{\to k_n}_{(k_1,\ldots,k_{n-1})}$ are applied, the (unnormalised) state at the n-th time step is (similarly to Eq. (3.7))

$$\begin{aligned} &M^{\to k_n}_{(k_1,\ldots,k_{n-1})} * A_{k_{n-1}} * \cdots * A_{k_3} * M^{\to k_3}_{(k_1,k_2)} * A_{k_2} * M^{\to k_2}_{(k_1)} * A_{k_1} * M^{\to k_1}_{\emptyset} * \rho \\ &\qquad = (M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * M^{\to k_3}_{(k_1,k_2)} * \cdots * M^{\to k_n}_{(k_1,\ldots,k_{n-1})}) * (\rho \otimes A_{k_1} \otimes \cdots \otimes A_{k_{n-1}}). \end{aligned} \tag{3.21}$$

Here, we need to impose the trace-preserving condition on the quantum instruments $\{M^{\to k_1}_{\emptyset}\}_{k_1\in\mathcal{N}}$ and $\{M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}\}_{k_{n+1}\in\mathcal{N}\setminus\{k_1,\ldots,k_n\}}$—i.e., the trace needs to be preserved when we sum over their classical outcomes. These conditions thus read

$$\begin{aligned} &\sum_{k_1\in\mathcal{N}} \mathrm{Tr}[\rho * M^{\to k_1}_{\emptyset}] = \mathrm{Tr}[\rho], \\ &\forall\, n = 1, \ldots, N-1,\ \forall\,(k_1, \ldots, k_n), \\ &\sum_{k_{n+1}\in\mathcal{N}\setminus\{k_1,\ldots,k_n\}} \mathrm{Tr}[(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_n}_{(k_1,\ldots,k_{n-1})} * M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}) * (\rho \otimes A_{k_1} \otimes \cdots \otimes A_{k_n})] \\ &\qquad\qquad = \mathrm{Tr}[(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_n}_{(k_1,\ldots,k_{n-1})}) * (\rho \otimes A_{k_1} \otimes \cdots \otimes A_{k_n})] \\ &\text{and } \forall\,(k_1, \ldots, k_N), \\ &\mathrm{Tr}[(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_N}_{(k_1,\ldots,k_{N-1})} * M^{\to F}_{(k_1,\ldots,k_N)}) * (\rho \otimes A_{k_1} \otimes \cdots \otimes A_{k_N})] \\ &\qquad\qquad = \mathrm{Tr}[(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_N}_{(k_1,\ldots,k_{N-1})}) * (\rho \otimes A_{k_1} \otimes \cdots \otimes A_{k_N})] \end{aligned} \tag{3.22}$$

[8] In accordance with the assumption that each operation can only be applied once, all sequences $(k_1, \ldots, k_n)$ we shall write assume that all k_i's $(1 \leq i \leq n)$ are different; for $n = N$, a sequence $(k_1, \ldots, k_N)$ shall thus contain each operation label $k_i \in \mathcal{N}$ once and only once. When we write the k_i's within parentheses as in $(k_1, \ldots, k_n)$, their order matters (as opposed to $\{k_1, \ldots, k_n\}$ which denotes an unordered set).

[9] Note that the quantum instruments $\{M^{\to k_N}_{(k_1,\ldots,k_{N-1})}\}_{k_N\in\mathcal{N}\setminus\{k_1,\ldots,k_{N-1}\}}$ between the last two time slots t_{N-1} and t_N consist of a single operation $M^{\to k_N}_{(k_1,\ldots,k_{N-1})}$ (which is thus CPTP), since they admit only one classical output value k_N, corresponding to the only remaining external circuit operation A_{k_N} that has not been applied yet.

Definition 3.4. Any circuit that abides by the above description, as represented in Fig. 3.4, with the internal circuit operations $M_{\emptyset}^{\to k_1}$, $M_{(k_1,\dots,k_n)}^{\to k_{n+1}}$ and $M_{(k_1,\dots,k_N)}^{\to F}$ satisfying the trace-preserving conditions in Eqs. (3.22) for any initial state ρ and CP maps A_k, is called a *quantum circuit with classical control of causal order* (QC-CC).

After the last operation, in the case where the order of the operations A_k, established dynamically, ends up being $(k_1, k_2, \dots, k_N)$, the state (3.21) reads

$$W_{(k_1,\dots,k_N,F)} * (\rho \otimes A_1 \otimes \cdots \otimes A_N), \tag{3.23}$$

with

$$W_{(k_1,\dots,k_N,F)} := M_{\emptyset}^{\to k_1} * M_{(k_1)}^{\to k_2} * M_{(k_1,k_2)}^{\to k_3} * \cdots * M_{(k_1,\dots,k_{N-1})}^{\to k_N} * M_{(k_1,\dots,k_N)}^{\to F}. \tag{3.24}$$

In order to obtain the process matrix describing the overall, deterministic supermap, we should however not post-select the state on the realisation of one particular causal order, and we need to sum over all realisations that can occur. Combining all causal orders $(k_1, \dots, k_N)$, and recalling the discussion in Sec. 3.2, we get:

Proposition 3.5. *The process matrix corresponding to the quantum circuit with classical control of causal order depicted on Fig. 3.4 is*

$$W = \sum_{(k_1,\dots,k_N)} W_{(k_1,\dots,k_N,F)} \tag{3.25}$$

with $W_{(k_1,\dots,k_N,F)} \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$ *defined in Eq.* (3.24).

3.4.2 Characterisation

Process matrices corresponding to QC-CC can be characterised as follows:

Proposition 3.6. *The process matrix* $W \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$ *of a quantum circuit with classical control of causal order can be decomposed in terms of positive semidefinite matrices* $W_{(k_1,\dots,k_N,F)} \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$ *and* $W_{(k_1,\dots,k_n)} \in \mathcal{L}(\mathcal{H}^{PA_{\{k_1,\dots,k_{n-1}\}}^{IO}A_{k_n}^{I}})$*, for all nonempty ordered subsets* $(k_1, \dots, k_n)$ *of* $\mathcal{N}$ *(with* $1 \leq n \leq N$*,* $k_i \neq k_j$ *for* $i \neq j$*), in such a way that*

$$W = \sum_{(k_1,\dots,k_N)} W_{(k_1,\dots,k_N,F)} \tag{3.26}$$

and

$$\forall\,(k_1,\dots,k_N),\qquad \mathrm{Tr}_F\, W_{(k_1,\dots,k_N,F)} = W_{(k_1,\dots,k_N)} \otimes \mathbb{1}^{A^O_{k_N}},$$

$$\forall\, n=1,\dots,N-1,\ \forall\,(k_1,\dots,k_n),\qquad \sum_{k_{n+1}\in\mathcal{N}\setminus\{k_1,\dots,k_n\}} \mathrm{Tr}_{A^I_{k_{n+1}}}\, W_{(k_1,\dots,k_n,k_{n+1})} = W_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}},$$

$$\text{and}\quad \sum_{k_1\in\mathcal{N}} \mathrm{Tr}_{A^I_{k_1}}\, W_{(k_1)} = \mathbb{1}^P. \tag{3.27}$$

Conversely, any Hermitian matrix $W \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ that admits a decomposition in terms of positive semidefinite matrices $W_{(k_1,\dots,k_N,F)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ and $W_{(k_1,\dots,k_n)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}})$ satisfying Eqs. (3.26)–(3.27) above is the process matrix of a quantum circuit with classical control of causal order.

Proof. From Eq. (3.22), it follows, in the very same way as in the QC-FO case above, that

$$\sum_{k_1} \mathrm{Tr}_{A^I_{k_1}\alpha_1}\, M^{\to k_1}_{\emptyset} = \mathbb{1}^P, \tag{3.28}$$

$$\begin{aligned}&\forall\, n=1,\dots,N-1,\ \forall\,(k_1,\dots,k_n),\\ &\sum_{k_{n+1}} \mathrm{Tr}_{A^I_{k_{n+1}}\alpha_{n+1}}\big(M^{\to k_1}_{\emptyset} * \cdots * M^{\to k_n}_{(k_1,\dots,k_{n-1})} * M^{\to k_{n+1}}_{(k_1,\dots,k_n)}\big)\\ &\qquad\qquad = \mathrm{Tr}_{\alpha_n}\big(M^{\to k_1}_{\emptyset} * \cdots * M^{\to k_n}_{(k_1,\dots,k_{n-1})}\big) \otimes \mathbb{1}^{A^O_{k_n}},\end{aligned} \tag{3.29}$$

$$\begin{aligned}&\text{and } \forall\,(k_1,\dots,k_N),\\ &\mathrm{Tr}_F\,\big(M^{\to k_1}_{\emptyset} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})} * M^{\to F}_{(k_1,\dots,k_N)}\big) = \mathrm{Tr}_{\alpha_N}\big(M^{\to k_1}_{\emptyset} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})}\big) \otimes \mathbb{1}^{A^O_{k_N}}.\end{aligned} \tag{3.30}$$

Consider the process matrix $W = \sum_{(k_1,\dots,k_N)} W_{(k_1,\dots,k_N,F)}$ of a QC-CC, as per Proposition 3.5, with the $W_{(k_1,\dots,k_N,F)}$'s of the form of Eq. (3.24), and with the internal circuit operations $M^{\to k_{n+1}}_{(k_1,\dots,k_n)} (\geq 0)$ satisfying Eqs. (3.28)–(3.30).

Let us then define, for all $1 \leq n \leq N$ and all $(k_1,\dots,k_n)$, the matrices

$$W_{(k_1,\dots,k_n)} := \mathrm{Tr}_{\alpha_n}\,\big(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_n}_{(k_1,\dots,k_{n-1})}\big) \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}}). \tag{3.31}$$

As all $M^{\to k_{n+1}}_{(k_1,\dots,k_n)} \geq 0$, it directly follows that all $W_{(k_1,\dots,k_n)}$'s (including the $W_{(k_1,\dots,k_N,k_{N+1})} = W_{(k_1,\dots,k_N,F)}$'s for $n = N+1$) are also positive semidefinite. Furthermore, the constraints of Eq. (3.27) are simply equivalent to Eqs. (3.28)–(3.30), and are thus readily satisfied by assumption.

To prove the converse, we will again construct suitable circuit operations explicitly. Consider any positive semidefinite matrix $W \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ that can be decomposed as in Eqs. (3.26)–(3.27). The positive semidefinite matrices $W_{(k_1,\dots,k_n)}$ (for $1 \leq n \leq N+1$) admit a spectral decomposition

of the form

$$W_{(k_1,\dots,k_n)} = \sum_i |w^i_{(k_1,\dots,k_n)}\rangle\langle w^i_{(k_1,\dots,k_n)}|\,, \tag{3.32}$$

for some $r_{(k_1,\dots,k_n)} := \operatorname{rank} W_{(k_1,\dots,k_n)}$ (nonnormalised and nonzero) orthogonal vectors $|w^i_{(k_1,\dots,k_n)}\rangle \in \mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}}$. We then define

$$|w_{(k_1,\dots,k_n)}\rangle := \sum_{i=1}^{r_{(k_1,\dots,k_n)}} |w^i_{(k_1,\dots,k_n)}\rangle \otimes |i\rangle^{\alpha_n} \in \mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}\alpha_n}, \tag{3.33}$$

where we introduced, for each $n = 1,\dots,N+1$, a (sufficiently high-dimensional) ancillary Hilbert space $\mathcal{H}^{\alpha_n}$ with computational basis $\{|i\rangle^{\alpha_n}\}$.[10]

We then define $|V^{\to k_1}_{\emptyset}\rangle\!\rangle := |w_{(k_1)}\rangle$, and, for $1 \le n \le N$,

$$|V^{\to k_{n+1}}_{(k_1,\dots,k_n)}\rangle\!\rangle := \sum_{i=1}^{r_{(k_1,\dots,k_n)}} \frac{|i\rangle^{\alpha_n}\,\langle w^i_{(k_1,\dots,k_n)}|}{\langle w^i_{(k_1,\dots,k_n)}|w^i_{(k_1,\dots,k_n)}\rangle} \otimes \mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}}\,|w_{(k_1,\dots,k_{n+1})}\rangle \;\in \mathcal{H}^{A^O_{k_n}\alpha_n A^I_{k_{n+1}}\alpha_{n+1}} \tag{3.34}$$

We proceed by showing that, for $1 \le n \le N$,

$$|w_{(k_1,\dots,k_n)}\rangle * |V^{\to k_{n+1}}_{(k_1,\dots,k_n)}\rangle\!\rangle = \Pi_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}}\,|w_{(k_1,\dots,k_{n+1})}\rangle = |w_{(k_1,\dots,k_{n+1})}\rangle\,, \tag{3.35}$$

where we used that the projector $\Pi_{(k_1,\dots,k_n)} := \sum_{i=1}^{r_{(k_1,\dots,k_n)}} \frac{|w^i_{(k_1,\dots,k_n)}\rangle\langle w^i_{(k_1,\dots,k_n)}|}{\langle w^i_{(k_1,\dots,k_n)}|w^i_{(k_1,\dots,k_n)}\rangle}$, tensored with $\mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}}$, acts as the identity on $|w_{(k_1,\dots,k_{n+1})}\rangle$.[11]

It follows recursively that, for all $1 \le n \le N+1$, $|w_{(k_1,\dots,k_n)}\rangle = |V^{\to k_1}_{\emptyset}\rangle\!\rangle * \cdots * |V^{\to k_n}_{(k_1,\dots,k_{n-1})}\rangle\!\rangle$, and therefore $W_{(k_1,\dots,k_n)} = \operatorname{Tr}_{\alpha_n}(M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * \cdots * M^{\to k_n}_{(k_1,\dots,k_{n-1})})$, with $M^{\to k_n}_{(k_1,\dots,k_{n-1})} = |V^{\to k_n}_{(k_1,\dots,k_{n-1})}\rangle\!\rangle\langle\!\langle V^{\to k_n}_{(k_1,\dots,k_{n-1})}|$ (and $W_{(k_1,\dots,k_N,F)} = M^{\to k_1}_{\emptyset} * \cdots * M^{\to F}_{(k_1,\dots,k_N)}$, with $M^{\to F}_{(k_1,\dots,k_N)} = \operatorname{Tr}_{\alpha_{N+1}} |V^{\to F}_{(k_1,\dots,k_N)}\rangle\!\rangle\langle\!\langle V^{\to F}_{(k_1,\dots,k_N)}|$).

To sum up, we have indeed explicitly found some positive semidefinite matrices $M^{\to k_1}_{\emptyset}$, $M^{\to k_{n+1}}_{(k_1,\dots,k_n)}$, and $M^{\to F}_{(k_1,\dots,k_N)}$ that compose the matrix W as per Proposition 3.5, and that satisfy Eq. (3.22)

[10] Concretely, this ancillary space needs to be at least of dimension $\max_{(k_1,\dots,k_n)}\{r_{(k_1,\dots,k_n)}\}$.

[11] Similarly to the QC-FO case, we denote by $\Pi^{\perp}_{(k_1,\dots,k_n)} := \mathbb{1}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}} - \Pi_{(k_1,\dots,k_n)}$ the orthogonal projector to $\Pi_{(k_1,\dots,k_n)}$ in $\mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^I_{k_n}}$, and we have

$$\begin{aligned}\sum_{k_{n+1}} &\operatorname{Tr}[(\Pi^{\perp}_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}})\,|w_{(k_1,\dots,k_{n+1})}\rangle\langle w_{(k_1,\dots,k_{n+1})}|]\\ &= \operatorname{Tr}[(\Pi^{\perp}_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}})(W_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}})] = 0.\end{aligned} \tag{3.36}$$

Since the individual summands for each k_{n+1} on the left-hand side of Eq. (3.36) cannot be negative, we conclude that $(\Pi^{\perp}_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}})\,|w_{(k_1,\dots,k_{n+1})}\rangle = 0$ and therefore $(\Pi_{(k_1,\dots,k_n)} \otimes \mathbb{1}^{A^O_{k_n}A^I_{k_{n+1}}\alpha_{n+1}})\,|w_{(k_1,\dots,k_{n+1})}\rangle = |w_{(k_1,\dots,k_{n+1})}\rangle$, as claimed.

(since the $W_{(k_1,\dots,k_n)} := \mathrm{Tr}_{\alpha_n}(M_\emptyset^{\to k_1} * M_{(k_1)}^{\to k_2} * \cdots * M_{(k_1,\dots,k_{n-1})}^{\to k_n})$ satisfy by assumption the constraints Eq. (3.27), and therefore also Eqs. (3.28)–(3.30)). This completes the proof of Proposition 3.6. □

We note that the matrices $W_{(k_1,\dots,k_N,F)}$ in Proposition 3.6 may or may not be valid (deterministic) process matrices. If the $W_{(k_1,\dots,k_N,F)}$'s are valid process matrices (up to normalisation), each compatible with the fixed causal order $(k_1,\dots,k_N)$, then we just end up with a probabilistic mixture of quantum circuits with different fixed causal orders. We recover the case considered in the previous section when there is only one term in the sum of Eq. (3.26), corresponding to the order $(k_1,\dots,k_N) = (1,\dots,N)$; in that case the constraints of Eq. (3.27) simply reduce to those of Eq. (3.10) (with $W_{(1,\dots,n)} \equiv W_{(n)}$ and $W_{(1,\dots,N,F)} \equiv W$).

If, on the other hand, we have truly "dynamical" causal order, with probabilistic operations establishing the causal order on the fly by generating probabilistic classical outcomes, then post-selecting on these outcomes does not result in a valid (deterministic) supermap, and thus the $W_{(k_1,\dots,k_N,F)}$'s are not valid process matrices. (In that case, the $W_{(k_1,\dots,k_N,F)}$ define "probabilistic process matrices", see Sec. 3.6 below.)

Note that the characterisation of Proposition 3.6 coincides precisely with the sufficient condition (Proposition 2.8) obtained in Chapter 2 for the causal separability of general N-partite process matrices. Hence, unsurprisingly, QC-CCs define causally separable processes. This also implies that they can only generate causal correlations, and can therefore not violate any causal inequality.

In particular, in the case of trivial $\mathcal{H}^P$ and $\mathcal{H}^F$, for $N = 2$, the characterisation of Proposition 3.6 just reduces to the characterisation of bipartite causally separable process matrices 2.5, and thus describes a probabilistic mixture of the two possible fixed causal orders $(1,2)$ and $(2,1)$. In order to have truly dynamical order between the N operations A_k, we need $N \geq 3$ (or $N = 2$ and some nontrivial $\mathcal{H}^P$).

3.4.3 Example

An example of a QC-CC is given by a "classical switch", in which two CP maps A_1 and A_2 are applied to a target system in an order that is controlled through another (control) qubit in a classical, incoherent way. We consider a state that is prepared in the global past space $\mathcal{H}^P = \mathcal{H}^{P_t} \otimes \mathcal{H}^{P_c}$, where P_t is the target system, and P_c is the control system. The circuit then performs a measurement on the control qubit, and depending on the (classical) measurement outcome, the target system is sent first to A_1 and then to A_2, or vice versa. The order is thus established dynamically through the preparation of the control state in the global past.

The first circuit operation is a quantum instrument $\{M_\emptyset^{\to 1} := |0\rangle\langle 0|^{P_c} \otimes |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}|^{P_t A_1^I}, M_\emptyset^{\to 2} := |1\rangle\langle 1|^{P_c} \otimes |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}|^{P_t A_2^I}\}$, and the other circuit operations are $M_{(1)}^{\to 2} = |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}|^{A_1^O A_2^I}$ and $M_{(1,2)}^{\to F} =$

$|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O F}$, as well as $M^{\rightarrow 1}_{(2)} = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O A_1^I}$ and $M^{\rightarrow F}_{(2,1)} = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O F}$. The process matrix describing this "classical switch" is thus

$$\begin{aligned} W_{\text{cs}} =& M^{\rightarrow 1}_{\emptyset} * M^{\rightarrow 2}_{(1)} * M^{\rightarrow F}_{(1,2)} + M^{\rightarrow 2}_{\emptyset} * M^{\rightarrow 1}_{(2)} * M^{\rightarrow F}_{(2,1)} \\ =& |0\rangle\langle 0|^{P_c} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_t A_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O A_2^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O F} \\ &+ |1\rangle\langle 1|^{P_c} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{P_t A_2^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_2^O A_1^I} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|^{A_1^O F} \end{aligned} \tag{3.37}$$

Note that in this case, we have truly dynamical causal order, going beyond a probabilistic mixture of two quantum circuits. The two individual summands in Eq. (3.37) do not satisfy the validity constraints for process matrices, and only their sum does. This reflects the fact that the first circuit operation is a probabilistic one, and if we post-select on one of the two outcomes, we do not end up with a valid (deterministic) supermap. (Indeed, the individual terms are probabilistic process matrices, as we define them in Sec. 3.6.) To obtain a valid process, we need to sum over the terms corresponding to the different outcomes.

It was proven in Ref. [3] that such a classical switch is impossible to realise in a standard QC-FO.

3.5 Quantum circuits with quantum control of causal order

We now move on to quantum circuits where the orders are controlled coherently and can be superposed. To do so, we will first give an equivalent description of QC-CCs, which will more naturally lead to our generalisation to QC-QCs.

3.5.1 Revisiting the description of quantum circuits with classical control of causal order

Remember that, in the circuits above, we did not assume that the input and output spaces of the operations A_k need to have the same dimension. Here, we will assume that the input and output spaces are all of the same dimension d, and thus isomorphic. This is an assumption that we can, in fact, make without loss of generality.[12]

In the previous section we imposed that each quantum operation $M^{\rightarrow k_{n+1}}_{(k_1,\ldots,k_n)}$ applied by the circuit between the time slot t_n and t_{n+1} was conditioned on which operations A_k had already been performed, and their order $(k_1,\ldots,k_n)$. This conditioning can be included in the description of the operation applied between t_n and t_{n+1} by introducing some control system C_n in some

[12] Indeed, if the input and output spaces of the operations A_k are not all the same, we can introduce additional ancillary input spaces $\mathcal{H}^{A_k^{I'}}$ (of dimension $d_k^{I'}$) and output spaces $\mathcal{H}^{A_k^{O'}}$ (of dimension $d_k^{O'}$), in such a way that $d_k^I d_k^{I'} = d_k^O d_k^{O'} = d$ for all k. (such $d_k^{I'}$ and $d_k^{O'}$ can always be found; one can simply choose for d the least common multiple of all d_k^I and d_k^O, and then take $d_k^{I'} = d/d_k^I$, and $d_k^{O'} = d/d_k^O$.) The original scenario is then recovered when the additional input and output spaces are traced out.

Hilbert space $\mathcal{H}^{C_n}$ that stores on the fly the dynamically established causal order in some (normalised) state

$$[\![(k_1,\ldots,k_n)]\!]^{C_n} := |(k_1,\ldots,k_n)\rangle\langle(k_1,\ldots,k_n)|^{C_n} \tag{3.38}$$

(with all different states $|(k_1,\ldots,k_n)\rangle \in \mathcal{H}^{C_n}$ being orthogonal), and by defining (for $1 \leq n \leq N-1$) the global operation (in its CJ representation)

$$\tilde{M}_{n+1} := \sum_{(k_1,\ldots,k_n,k_{n+1})} \tilde{M}^{\to k_{n+1}}_{(k_1,\ldots,k_n)} \otimes [\![(k_1,\ldots,k_n)]\!]^{C'_n} \otimes [\![(k_1,\ldots,k_n,k_{n+1})]\!]^{C_{n+1}}, \tag{3.39}$$

applied between t_n and t_{n+1}. The last two terms in the summands

$$[\![(k_1,\ldots,k_n)]\!]^{C'_n} \otimes [\![(k_1,\ldots,k_n,k_{n+1})]\!]^{C_{n+1}},$$

indicate that as the outcome of the operation $M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}$ is k_{n+1}, the causal order (stored in the control system) is updated from $(k_1,\ldots,k_n)$ to $(k_1,\ldots,k_n,k_{n+1})$ [13]. Note also that the various operations $M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}$ have different input spaces $\mathcal{H}^{A^O_{k_n}\alpha_n}$ and output spaces $\mathcal{H}^{A^I_{k_{n+1}}\alpha_{n+1}}$; in order to make the definition of $\tilde{M}_{n+1}$ in Eq. (3.39) consistent, we introduce, at each time slot, a common, d-dimensional "target system" input (output) Hilbert space $\mathcal{H}^{T^I_n}$ ($\mathcal{H}^{T^O_n}$). We then identify each $\mathcal{H}^{A^I_{k_{n+1}}}$ with $\mathcal{H}^{T^I_{n+1}}$ and each $\mathcal{H}^{A^O_{k_n}}$ with $\mathcal{H}^{T^O_n}$. The tilde we added on $\tilde{M}^{\to k_{n+1}}_{(k_1,\ldots,k_n)} \equiv M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}$ in Eq. (3.39) indicates this identification, i.e., it is the notation that we use whenever we let some object act on the common target system. With these notations, we have $\tilde{M}^{\to k_{n+1}}_{(k_1,\ldots,k_n)} \in \mathcal{L}(\mathcal{H}^{T^O_n\alpha_n T^I_{n+1}\alpha_{n+1}})$ and $\tilde{M}_{n+1} \in \mathcal{L}(\mathcal{H}^{T^O_n\alpha_n C'_n T^I_{n+1}\alpha_{n+1} C_{n+1}})$.

In a similar way, for the circuit operations before the time slot t_1 and after the time slot t_N, we define

$$\tilde{M}_1 := \sum_{k_1} \tilde{M}^{\to k_1}_{\emptyset} \otimes [\![(k_1)]\!]^{C_1} \in \mathcal{L}(\mathcal{H}^{PT^I_1\alpha_1 C_1}), \tag{3.40}$$

where the control state $[\![(k_1)]\!]^{C_1}$ records the fact that the first operation to be applied shall be A_{k_1}, and

$$\tilde{M}_{N+1} := \sum_{(k_1,\ldots,k_N)} \tilde{M}^{\to F}_{(k_1,\ldots,k_N)} \otimes [\![(k_1,\ldots,k_N)]\!]^{C'_N} \in \mathcal{L}(\mathcal{H}^{T^O_N\alpha_N C_N F}), \tag{3.41}$$

where $[\![(k_1,\ldots,k_N)]\!]^{C'_N}$ conditions the application of $\tilde{M}^{\to F}_{(k_1,\ldots,k_N)}$ to the causal order of all operations being $(k_1,\ldots,k_N)$.

[13] We use a prime to distinguish between the control system C_n at the beginning of the time slot t_n (i.e., before the application of the input operations A_k) and the control system C'_n at the end of the time slot t_n (after the application of the input operations A_k); cf. also Eq. (3.42).

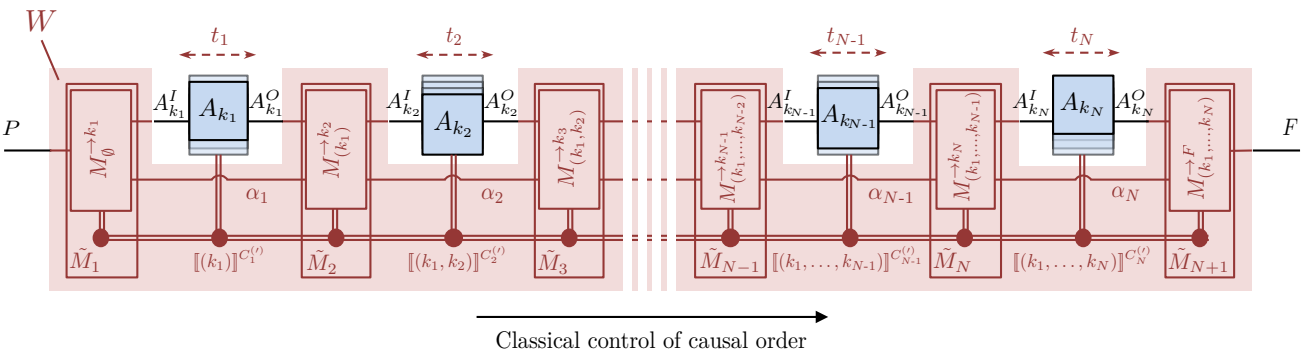

FIGURE 3.5: Another possible representation of a QC-CC, equivalent to Fig. 3.4. Here we show explicitly the transmission of the information about the causal order, established dynamically and stored on the fly in some control system in the states $[\![(k_1,\ldots,k_n)]\!]^{C_n}$. The double-stroke lines indicate that this information is classical. This information is used to control the upcoming operations $\tilde{M}^{\to k_{n+1}}_{(k1,\ldots,k_n)}$, so as to define a joint operation $\tilde{M}_{n+1}$ on $\mathcal{H}^{T_n^O \alpha_n C_n}$, as in Eq. (3.39).

We similarly combine the operations A_k to a global operation

$$\tilde{A}_n := \sum_{(k_1,\ldots,k_n)} \tilde{A}_{k_n} \otimes [\![(k_1,\ldots,k_n)]\!]^{C_n} \otimes [\![(k_1,\ldots,k_n)]\!]^{C'_n} \in \mathcal{L}(\mathcal{H}^{T_n^I T_n^O C_n C'_n}), \tag{3.42}$$

and consider that at each time slot t_n the operation $\tilde{A}_n$ is applied.

This alternative description for a QC-CC, which explicitly involves some physical control system, is illustrated in Fig. 3.5. To see that it is indeed equivalent to the previous description, we note that the operations $\tilde{M}_n$ and $\tilde{A}_n$ described above are applied in a well-defined order; for some input state $\rho \in \mathcal{L}(\mathcal{H}^P)$ sent into the circuit, the output state is then, similarly to Eq. (3.7),

$$\tilde{M}_{N+1} * \tilde{A}_N * \tilde{M}_N * \cdots * \tilde{M}_2 * \tilde{A}_1 * \tilde{M}_1 * \rho. \tag{3.43}$$

Tracing out all control systems C_n in the calculation of the link products above, we find that this output state can be written as

$$\begin{aligned}
&\sum_{(k_1,\ldots,k_N)} \tilde{M}^{\to F}_{(k_1,\ldots,k_N)} * \tilde{A}_N * \tilde{M}^{\to k_N}_{(k_1,\ldots,k_{N-1})} * \cdots * \tilde{M}^{\to k_2}_{(k_1)} * \tilde{A}_1 * \tilde{M}^{\to k_1}_{\emptyset} * \rho \\
&= \sum_{(k_1,\ldots,k_N)} M^{\to F}_{(k_1,\ldots,k_N)} * A_{k_N} * M^{\to k_N}_{(k_1,\ldots,k_{N-1})} * \cdots * M^{\to k_2}_{(k_1)} * A_{k_1} * M^{\to k_1}_{\emptyset} * \rho \\
&= \sum_{(k_1,\ldots,k_N)} W_{(k_1,\ldots,k_N,F)} * (\rho \otimes A_1 \otimes \cdots \otimes A_N),
\end{aligned} \tag{3.44}$$

with $W_{(k_1,\ldots,k_N,F)}$ defined in Eq. (3.24), which implies that the process matrix corresponding to our circuit is indeed $W = \sum_{(k_1,\ldots,k_N)} W_{(k_1,\ldots,k_N,F)}$, as in Proposition 3.5.

It follows straightforwardly from the constraints (3.28)–(3.30) satisfied by the circuit operations $M^{\to k_{n+1}}_{(k_1,\ldots,k_n)}$ that the global circuit operations $\tilde{M}_n$, obtained through combining them in the way described above, act trace-preservingly on any input state of the combined target, control and

ancillary system that they can receive. Note, however, that, conversely, requiring only trace-preservation on the global circuit operations $\tilde{M}_n$ is not sufficient, and does not define a QC-CC in general. In fact, what we need to require in order to recover the constraints (3.28)–(3.30) is

$$\begin{aligned}
&\mathrm{Tr}[\tilde{M}_1 * \rho] = \mathrm{Tr}[\rho],\\
&\forall\, n = 1,\dots,N-1,\ \forall\,(k_1,\dots,k_n),\\
&\sum_{k_{n+1}} \mathrm{Tr}[(\mathbb{1}^{T^I_{n+1}\alpha_{n+1}} \otimes [\![(k_1,\dots,k_n,k_{n+1})]\!]^{C_{n+1}}) \cdot (\tilde{M}_{n+1} * \tilde{A}_n * \tilde{M}_n * \tilde{A}_{n-1} * \cdots * \tilde{M}_1 * \rho)]\\
&\qquad = \mathrm{Tr}[(\mathbb{1}^{T^n_O\alpha_n} \otimes [\![(k_1,\dots,k_n)]\!]^{C'_n}) \cdot (\tilde{A}_n * \tilde{M}_n * \tilde{A}_{n-1} * \cdots * \tilde{M}_1 * \rho)]. \qquad (3.45)\\
&\text{and } \forall\,(k_1,\dots,k_N),\\
&\mathrm{Tr}[(\tilde{M}^{\to F}_{(k_1,\dots,k_N)} \otimes [\![(k_1,\dots,k_N)]\!]^{C'_N}) * (\tilde{A}_N * \tilde{M}_N * \tilde{A}_{N-1} * \cdots * \tilde{M}_1 * \rho)]\\
&\qquad = \mathrm{Tr}[(\mathbb{1}^{T^N_O\alpha_N} \otimes [\![(k_1,\dots,k_N)]\!]^{C'_N}) \cdot (\tilde{A}_N * \tilde{M}_N * \tilde{A}_{N-1} * \cdots * \tilde{M}_1 * \rho)]. \qquad (3.46)
\end{aligned}$$

These constraints mean that the global operations should also preserve the probabilities for each thus-far-established, well-defined causal order $(k_1,\dots,k_n)$ to be realised.

Before moving to the generalisation to QC-QCs, let us also further note that, as we do not make any particular assumption on the ancillary Hilbert spaces $\mathcal{H}^{\alpha_n}$ (e.g., on their dimension), those can be used to "purify" all operations $M^{\to k_n}_{(k_1,\dots,k_{n-1})}$ with $1 \le n \le N$. This implies that without loss of generality, all operations $M^{\to k_n}_{(k_1,\dots,k_{n-1})}$ can be assumed to consist of the application of just one Kraus operator, which we shall denote $V^{\to k_n}_{(k_1,\dots,k_{n-1})}$; in such a case the CJ representations of the operations simply write

$$M^{\to k_n}_{(k_1,\dots,k_{n-1})} = |V^{\to k_n}_{(k_1,\dots,k_{n-1})}\rangle\!\rangle\langle\!\langle V^{\to k_n}_{(k_1,\dots,k_{n-1})}|\,, \qquad (3.47)$$

where $|V^{\to k_n}_{(k_1,\dots,k_{n-1})}\rangle\!\rangle \in \mathcal{H}^{A^O_{k_{n-1}}\alpha_{n-1}A^I_{k_n}\alpha_n}$ (or $|V^{\to k_1}_{\emptyset}\rangle\!\rangle \in \mathcal{H}^{PA^I_{k_1}\alpha_1}$ for $n=1$) is the (pure) CJ representation of $V^{\to k_n}_{(k_1,\dots,k_{n-1})}$; see Sec. 1.1.1. Similarly for the last operations $M^{\to F}_{(k_1,\dots,k_N)}$, one can introduce some ancillary Hilbert space $\mathcal{H}^{\alpha_{N+1}}$ so as to "purify" these operations, before tracing out the ancillary system in $\mathcal{H}^{\alpha_{N+1}}$. Without loss of generality we can thus write

$$M^{\to F}_{(k_1,\dots,k_N)} = \mathrm{Tr}_{\alpha_{N+1}}\, |V^{\to F}_{(k_1,\dots,k_N)}\rangle\!\rangle\langle\!\langle V^{\to F}_{(k_1,\dots,k_N)}|\,, \qquad (3.48)$$

with $|V^{\to F}_{(k_1,\dots,k_N)}\rangle\!\rangle \in \mathcal{H}^{A^O_{k_N}\alpha_N F\alpha_{N+1}}$.

In the following it will indeed be convenient to assume that the circuit operations are pure; we shall thus now describe the circuits in terms of single Kraus operators. For simplicity we shall also assume that the operations A_k correspond to the application of a single Kraus operator—which we also write (with a slight but unambiguous enough conflict of notation) A_k, with a (pure) CJ

representation $|A_k\rangle\rangle$. The general case can then easily be recovered by combining what we shall get for different Kraus operators[14].

3.5.2 Turning the classical control into a coherent control of causal order

As a next step, we now want to turn the classical control described so far into a "quantum control".

Recall that in the case of a classical control, the state $[\![(k_1, \ldots, k_n)]\!]^{C_n}$ of the control system was keeping track of the whole history of which operations had been applied so far. At the end of the day, the complete causal order $(k_1, \ldots, k_N)$ was kept track of in each component $W_{(k_1,\ldots,k_N,F)}$ of W, so that the causal order could be said to be "definite". Here we will now relax this idea of keeping track of the complete causal order. We will thus now just let the control system record which operations have already been applied—recall that we still make the crucial assumption that each operation is applied once and only once, so that one must clearly keep track of this information—and encode which operation should be applied at a given time slot. Importantly, we will not let the control system keep track of the order in which the previous operations were applied; this will allow for different orders $(k_1, \ldots, k_{n-1})$ and $(k'_1, \ldots, k'_{n-1})$ corresponding to the same set $\mathcal{K} = \{k_1, \ldots, k_{n-1}\} = \{k'_1, \ldots, k'_{n-1}\}$ to "interfere". Hence, we shall replace the basis states $|(k_1, \ldots, k_n)\rangle^{C_n}$ of the control system by some basis states of the form $|\mathcal{K}_{n-1}, k_n\rangle^{C_n}$, where $\mathcal{K}_{n-1}$ is the (unordered) set of operations that have already been applied before the time slot t_n, and $k_n \notin \mathcal{K}_{n-1}$ labels the operation to be applied at time slot t_n.

Before the time slot t_1, we let the circuit transform the input state into a state that is sent coherently to all operations A_{k_1} (and some ancillary system in $\mathcal{H}^{\alpha_1}$), while accordingly attaching the control state $|\emptyset, k_1\rangle^{C_1}$ to each component of the superposition. That is, instead of Eq. (3.40), we now let the circuit apply an operation of the form (recall that we are now considering "pure operations")

$$\tilde{V}_1 := \sum_{k_1} \tilde{V}_{\emptyset,\emptyset}^{\to k_1} \otimes |\emptyset, k_1\rangle^{C_1} , \tag{3.49}$$

where $V_{\emptyset,\emptyset}^{\to k_1} : \mathcal{H}^P \to \mathcal{H}^{A^I_{k_1}\alpha_1}$, and where as in the previous subsection we use a tilde on $\tilde{V}_{\emptyset,\emptyset}^{\to k_1} \equiv V_{\emptyset,\emptyset}^{\to k_1}$ to indicate that we embed it into the common target system Hilbert space, that is, we see it as an operation $\tilde{V}_{\emptyset,\emptyset}^{\to k_1} : \mathcal{H}^P \to \mathcal{H}^{T^I_1\alpha_1}$.

Between two time slots t_n and t_{n+1}, for $1 \leq n \leq N-1$, we then let the circuit act coherently on the target, the ancilla and the control systems. We denote by $V_{\mathcal{K}_{n-1},k_n}^{\to k_{n+1}}$ the operation applied to the target and the ancilla systems between the output of the operation A_{k_n} and the input of the

[14]Note that, as we will point out in Chapter 4, the notion of coherent control of a general quantum operation is problematic, and in fact ill-defined without additional information about the quantum operation, beyond its specification in terms of Kraus operators. In the cases we consider here, however, these problems do not arise, since by construction each of the operations A_k is applied once and only once, and we thus deal only with situations that correspond to valid quantum supermaps; see Sec. 4.5.

operation $A_{k_{n+1}}$, when the set of $n-1$ previously performed operations is $\mathcal{K}_{n-1}$ (independently of their order). According to the description above, the corresponding state of the control system is changed in this case from $|\mathcal{K}_{n-1}, k_n\rangle^{C'_n}$ to $|\mathcal{K}_{n-1} \cup \{k_n\}, k_{n+1}\rangle^{C_{n+1}}$. That is, instead of Eq. (3.39), we can describe the global ("pure") operation of the circuit between the time slots t_n and t_{n+1} as

$$\tilde{V}_{n+1} := \sum_{\substack{\mathcal{K}_{n-1} \\ k_n, k_{n+1}}} \tilde{V}^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n} \otimes |\mathcal{K}_n, k_{n+1}\rangle^{C_{n+1}} \langle \mathcal{K}_{n-1}, k_n|^{C'_n}, \tag{3.50}$$

where the sum is over all subsets $\mathcal{K}_{n-1} \subsetneq \mathcal{N}$ such that $|\mathcal{K}_{n-1}| = n-1$ and all $k_n, k_{n+1} \in \mathcal{N} \backslash \mathcal{K}_{n-1}$ (such that $k_n \neq k_{n+1}$), and with $\mathcal{K}_n := \mathcal{K}_{n-1} \cup \{k_n\}$. As before, in Eq.(3.50) we added a tilde on $\tilde{V}^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n} \equiv V^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n}$ to indicate that $V^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n} : \mathcal{H}^{A^O_{k_n} \alpha_n} \to \mathcal{H}^{A^I_{k_{n+1}} \alpha_{n+1}}$ is embedded into the common target system Hilbert spaces and seen as an operation $\tilde{V}^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n} : \mathcal{H}^{T^O_n \alpha_n} \to \mathcal{H}^{T^I_{n+1} \alpha_{n+1}}$.

Similarly for $n = N$, with F replacing $A^I_{k_{N+1}}$, and noting that the control state does not need to be updated (as it would anyway always be in the state $|\mathcal{N}, F\rangle^{C_{N+1}}$) and that the sum reduces to a sum over $k_N \in \mathcal{N}$ (with $\mathcal{K}_{N-1} = \mathcal{N} \backslash \{k_N\}$), we obtain

$$\tilde{V}_{N+1} := \sum_{k_N} \tilde{V}^{\to F}_{\mathcal{N} \backslash \{k_N\}, k_N} \otimes \langle \mathcal{N} \backslash \{k_N\}, k_N|^{C_N}, \tag{3.51}$$

with $\tilde{V}^{\to F}_{\mathcal{N} \backslash \{k_N\}, k_N} : \mathcal{H}^{T^O_N \alpha_N} \to \mathcal{H}^{F \alpha_{N+1}}$. Recall that α_{N+1} will then be traced out.

The CJ representations of the above operations are

$$|\tilde{V}_1\rangle\rangle := \sum_{k_1} |\tilde{V}^{\to k_1}_{\emptyset, \emptyset}\rangle\rangle \otimes |\emptyset, k_1\rangle^{C_1} \in \mathcal{H}^{P T^I_1 \alpha_1 C_1}, \tag{3.52}$$

$$|\tilde{V}_{n+1}\rangle\rangle := \sum_{\substack{\mathcal{K}_{n-1} \\ k_n, k_{n+1}}} |\tilde{V}^{\to k_{n+1}}_{\mathcal{K}_{n-1}, k_n}\rangle\rangle \otimes |\mathcal{K}_{n-1}, k_n\rangle^{C'_n} \otimes |\mathcal{K}_n, k_{n+1}\rangle^{C_{n+1}} \in \mathcal{H}^{T^O_n \alpha_n C'_n T^I_{n+1} \alpha_{n+1} C_{n+1}}, \tag{3.53}$$

$$|\tilde{V}_{N+1}\rangle\rangle := \sum_{k_N} |\tilde{V}^{\to F}_{\mathcal{N} \backslash \{k_N\}, k_N}\rangle\rangle \otimes |\mathcal{N} \backslash \{k_N\}, k_N\rangle^{C'_N} \in \mathcal{H}^{T^O_N \alpha_N C'_N F \alpha_{N+1}}. \tag{3.54}$$

In terms of CP maps (as considered in the previous sections), we then have $\tilde{M}_n = |\tilde{V}_n\rangle\rangle\langle\langle \tilde{V}_n|$ for $n \leq N$, and $\tilde{M}_{N+1} = \mathrm{Tr}_{\alpha_{N+1}} |\tilde{V}_{N+1}\rangle\rangle\langle\langle \tilde{V}_{N+1}|$.

Such a QC-QC is represented on Fig. 3.6.

As we did in the previous subsection, recalling that each operation (here, Kraus operator) $A_k : \mathcal{H}^{A^I_k} \to \mathcal{H}^{A^O_k}$ can *a priori* be applied at each time slot t_n, we embed them in the target system Hilbert spaces $\mathcal{H}^{T^I_n}$ and $\mathcal{H}^{T^O_n}$, and combine them to the global operation

$$|\tilde{A}_n\rangle\rangle := \sum_{\mathcal{K}_{n-1}, k_n} |\tilde{A}_{k_n}\rangle\rangle \otimes |\mathcal{K}_{n-1}, k_n\rangle^{C_n} \otimes |\mathcal{K}_{n-1}, k_n\rangle^{C'_n} \in \mathcal{H}^{T^I_n T^O_n C_n C'_n} \tag{3.55}$$

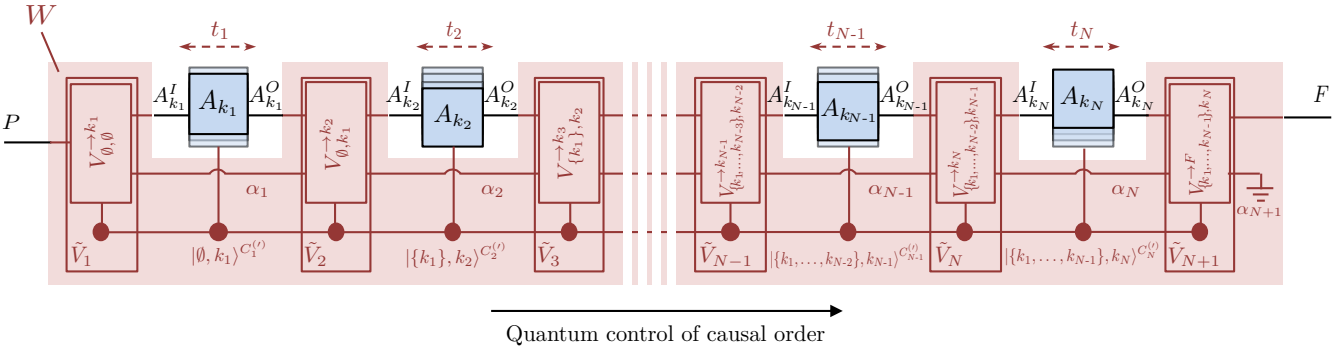

FIGURE 3.6: Quantum circuit with quantum control of causal order (QC-QC). We replaced the classical control system of Fig. 3.5 by a quantum control system with basis states $|\{k_1,\dots,k_{n-1}\},k_n\rangle^{C_n}$, which only stores information about which operations ($\{k_1,\dots,k_{n-1}\}$) have already been applied (but not about their order) and the currently performed operation (k_n). (Note that in contrast to the previous figures, the labels in the "boxes" that make up the circuit refer here to the Kraus operators of each operation, rather than to the (CJ matrix of the) corresponding CP maps). We illustrate here the component $|w_{(k_1,k_2,\dots,k_N)}\rangle\!\rangle$ of the process, corresponding to the order $(k_1,k_2,\dots,k_N)$—which is coherently superposed with other components, corresponding to different orders, so as to define the process matrix W; see Proposition 3.8.

which is applied at the time slot t_n. Note, as in the previous case, that the operations $\tilde{V}_n$ and $\tilde{A}_n$ are applied in a well-defined order. Considering some input state $|\psi\rangle \in \mathcal{H}^P$ sent into the circuit, then, (similarly to Eq. (3.7) and Eq. (3.21)) the state after the n-th circuit operation is

$$|\tilde{V}_n\rangle\!\rangle * |\tilde{A}_{n-1}\rangle\!\rangle * \cdots * |\tilde{V}_2\rangle\!\rangle * |\tilde{A}_1\rangle\!\rangle * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \in \mathcal{H}^{T_n^I \alpha_n C_n}. \tag{3.56}$$

Here, the trace-preserving condition that the circuit operations need to satisfy reads

$$\begin{aligned} \| \, |\psi\rangle \, \|^2 &= \| \, |\psi\rangle * |\tilde{V}_1\rangle\!\rangle \, \|^2 \quad \text{and} \\ \| \, |\tilde{A}_n\rangle\!\rangle * |\tilde{V}_n\rangle\!\rangle * |\tilde{A}_{n-1}\rangle\!\rangle * \cdots * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \, \|^2 &= \\ &\| \, |\tilde{V}_{n+1}\rangle\!\rangle * |\tilde{A}_n\rangle\!\rangle * |\tilde{V}_n\rangle\!\rangle * |\tilde{A}_{n-1}\rangle\!\rangle * \cdots * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \, \|^2, \end{aligned} \tag{3.57}$$

(for $1 \leq n \leq N$, and any choices for the CP maps $(A_1,\dots,A_N)$ and the input state $|\psi\rangle$).

Definition 3.7. Any circuit that abides by the previous description, as represented in Fig. 3.6, with the internal circuit operations $V^{\to k_1}_{\emptyset,\emptyset}$, $V^{\to k_{n+1}}_{\mathcal{K}_{n-1},k_n}$, and $V^{\to F}_{\mathcal{N}\setminus\{k_N\},k_N}$ satisfying the trace-preserving conditions in Eq. (3.57) for any initial state $|\psi\rangle$ and CP maps (Kraus operators) A_k, is called a *quantum circuit with quantum control of causal order* (QC-QC).

After the last circuit operation, the joint output state of the target system in $\mathcal{H}^F$ and the ancillary system in $\mathcal{H}^{\alpha_{N+1}}$ is

$$|\tilde{V}_{N+1}\rangle\!\rangle * |\tilde{A}_N\rangle\!\rangle * |\tilde{V}_N\rangle\!\rangle * \cdots * |\tilde{V}_2\rangle\!\rangle * |\tilde{A}_1\rangle\!\rangle * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \, . \tag{3.58}$$

After some simplification we find that this output state can be written as

$$
\begin{aligned}
&\sum_{(k_1,\dots,k_N)} |\tilde{V}^{\to F}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\!\rangle * |\tilde{A}_N\rangle\!\rangle * |\tilde{V}^{\to k_N}_{\{k_1,\dots,k_{N-2}\},k_{N-1}}\rangle\!\rangle * \cdots * |\tilde{V}^{\to k_2}_{\emptyset,k_1}\rangle\!\rangle * |\tilde{A}_1\rangle\!\rangle * |\tilde{V}^{\to k_1}_{\emptyset,\emptyset}\rangle\!\rangle * |\psi\rangle \\
&= \sum_{(k_1,\dots,k_N)} |V^{\to F}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\!\rangle * |A_{k_N}\rangle\!\rangle * |V^{\to k_N}_{\{k_1,\dots,k_{N-2}\},k_{N-1}}\rangle\!\rangle * \cdots * |V^{\to k_2}_{\emptyset,k_1}\rangle\!\rangle * |A_{k_1}\rangle\!\rangle * |V^{\to k_1}_{\emptyset,\emptyset}\rangle\!\rangle * |\psi\rangle \\
&= \sum_{(k_1,\dots,k_N)} |w_{(k_1,\dots,k_N,F)}\rangle * (|\psi\rangle \otimes |A_1\rangle\!\rangle \otimes \cdots \otimes |A_N\rangle\!\rangle)
\end{aligned} \tag{3.59}
$$

with

$$
|w_{(k_1,\dots,k_N,F)}\rangle := |V^{\to k_1}_{\emptyset,\emptyset}\rangle\!\rangle * |V^{\to k_2}_{\emptyset,k_1}\rangle\!\rangle * |V^{\to k_3}_{\{k_1\},k_2}\rangle\!\rangle * \cdots * |V^{\to k_N}_{\{k_1,\dots,k_{N-2}\},k_{N-1}}\rangle\!\rangle * |V^{\to F}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\!\rangle. \tag{3.60}
$$

By some analogous reasoning to the ones in the previous sections (here applied to "pure" operations), and recalling that the last ancillary system in $\mathcal{H}^{\alpha_{N+1}}$ is to be traced out at the end, we find that:

Proposition 3.8. *The process matrix corresponding to the quantum circuit with quantum control of causal order depicted on Fig. 3.6 is*

$$
\begin{aligned}
W &= \mathrm{Tr}_{\alpha_{N+1}} |w\rangle\langle w| \\
\textit{with} \quad |w\rangle &:= \sum_{(k_1,\dots,k_N)} |w_{(k_1,\dots,k_N,F)}\rangle
\end{aligned} \tag{3.61}
$$

and with $|w_{(k_1,\dots,k_N,F)}\rangle \in \mathcal{H}^{PA^{IO}_{\mathcal{N}}F\alpha_{N+1}}$ *defined in Eq. (3.60).*

3.5.3 Characterisation

The following proposition characterises process matrices of QC-QC:

Proposition 3.9. *The process matrix* $W \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ *of a quantum circuit with quantum control of causal order is such that there exist positive semidefinite matrices* $W_{(\mathcal{K},\ell)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{K}}A^I_\ell})$*, for all strict subsets* $\mathcal{K}$ *of* $\mathcal{N}$ *and all* $\ell \in \mathcal{N}\backslash\mathcal{K}$*, satisfying*

$$
\begin{aligned}
&\mathrm{Tr}_F\, W = \sum_{k\in\mathcal{N}} W_{(\mathcal{N}\backslash\{k\},k)} \otimes \mathbb{1}^{A^O_k}, \\
&\forall\, \emptyset \subsetneq \mathcal{K} \subsetneq \mathcal{N}, \sum_{\ell\in\mathcal{N}\backslash\mathcal{K}} \mathrm{Tr}_{A^I_\ell}\, W_{(\mathcal{K},\ell)} = \sum_{k\in\mathcal{K}} W_{(\mathcal{K}\backslash\{k\},k)} \otimes \mathbb{1}^{A^O_k}, \\
&\text{and} \quad \sum_{\ell\in\mathcal{N}} \mathrm{Tr}_{A^I_\ell}\, W_{(\emptyset,\ell)} = \mathbb{1}^P.
\end{aligned} \tag{3.62}
$$

Conversely, any Hermitian matrix $W \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ such that there exist positive semidefinite matrices $W_{(\mathcal{K},\ell)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{K}}A^{I}_{\ell}})$ for all $\mathcal{K} \subsetneq \mathcal{N}$ and $\ell \in \mathcal{N}\backslash\mathcal{K}$ satisfying Eq. (3.62) is the process matrix of a quantum circuit with quantum control of causal order.

Proof. For some process matrix of a QC-QC as per Proposition 3.8, the matrices $W_{(\mathcal{K},\ell)}$ are given by

$$\begin{aligned} W_{(\mathcal{K},\ell)} &:= \mathrm{Tr}_{\alpha_{n+1}} |w_{(\mathcal{K},\ell)}\rangle\langle w_{(\mathcal{K},\ell)}| \\ \text{with} \quad |w_{(\mathcal{K},\ell)}\rangle &:= \sum_{\substack{(k_1,\ldots,k_n):\\ \{k_1,\ldots,k_n\}=\mathcal{K}}} |V^{\to k_1}_{\emptyset,\emptyset}\rangle\rangle * |V^{\to k_2}_{\emptyset,k_1}\rangle\rangle * |V^{\to k_3}_{\{k_1\},k_2}\rangle\rangle * \cdots \\ &\qquad \cdots |V^{\to k_n}_{\{k_1,\ldots,k_{n-2}\},k_{n-1}}\rangle\rangle * |V^{\to \ell}_{\{k_1,\ldots,k_{n-1}\},k_n}\rangle\rangle \in \mathcal{H}^{PA^{IO}_{\mathcal{K}}A^{I}_{\ell}\alpha_{n+1}}, \end{aligned} \tag{3.63}$$

where the last sum is taken over all ordered sequences $(k_1, \ldots, k_n)$ of $\mathcal{K}$. The fact that they satisfy the constraints given in Proposition 3.9 follows again from the trace-preserving condition (Eq. (3.57)). Tracing out the control systems and simplifying, we obtain

$$\begin{aligned} &\mathrm{Tr}\, |\psi\rangle\langle\psi| = \mathrm{Tr}[\sum_{\ell\in\mathcal{N}} \mathrm{Tr}_{A^I_\ell} W_{(\emptyset,\ell)} \cdot |\psi\rangle\langle\psi|^T] \quad \text{and} \\ &\mathrm{Tr}\,[\sum_{\substack{\mathcal{K}\subset\mathcal{N}\\ |\mathcal{K}|=n}} (\sum_{k\in\mathcal{K}} W_{(\mathcal{K}\backslash\{k\},k)} \otimes \mathbb{1}^{A^O_k}) \cdot (|\psi\rangle\langle\psi| \bigotimes_{r\in\mathcal{K}} |A_r\rangle\rangle\langle\langle A_r|)^T] \\ &\qquad\qquad = \mathrm{Tr}\,[\sum_{\substack{\mathcal{K}\subset\mathcal{N}\\ |\mathcal{K}|=n}} (\sum_{\ell\in\mathcal{N}\backslash\mathcal{K}} \mathrm{Tr}_{A^I_\ell} W_{(\mathcal{K},\ell)}) \cdot (|\psi\rangle\langle\psi| \bigotimes_{r\in\mathcal{K}} |A_r\rangle\rangle\langle\langle A_r|)^T], \end{aligned} \tag{3.64}$$

Once again, since Eq. (3.64) holds for any state and any CP maps, the conditions (3.62) follow[15].

To prove the converse, suppose that we have a Hermitian matrix W satisfying the decomposition in Eq. (3.62). We proceed again very much like in the previous cases, and construct explicit circuit operations that compose W in the desired way. We introduce the spectral decomposition of each of the $W_{(\mathcal{K},\ell)}$ (and of $W_{(\mathcal{N},F)} := W$) of the form

$$W_{(\mathcal{K},\ell)} = \sum_i |w^i_{(\mathcal{K},\ell)}\rangle \langle w^i_{(\mathcal{K},\ell)}|\,, \tag{3.65}$$

for some eigenbasis consisting of $r_{(\mathcal{K},\ell)} := \mathrm{rank}\, W_{(\mathcal{K},\ell)}$ (nonnormalised and nonzero) orthogonal vectors $|w^i_{(\mathcal{K},\ell)}\rangle \in \mathcal{H}^{PA^{IO}_{\mathcal{K}}A^{I}_{\ell}}$. We then define

$$|w_{(\mathcal{K},\ell)}\rangle := \sum_{i=1}^{r_{(\mathcal{K},\ell)}} |w^i_{(\mathcal{K},\ell)}\rangle \otimes |i\rangle^{\alpha_{n+1}} \tag{3.66}$$

[15] To make the sum over subsets in the second and third lines of Eq. (3.64) disappear (i.e., to obtain the condition in Eq. (3.62) corresponding to some individual subset $\emptyset \subsetneq \mathcal{K}' \subsetneq \mathcal{N}$) one can simply choose all A_r with $r \in \mathcal{N}\backslash\mathcal{K}'$ to be zero. Then only the desired term corresponding to one particular subset $\mathcal{K}'$ remains in the sum.

where $\{|i\rangle^{\alpha_{n+1}}\}$ are computational basis states of a sufficiently high-dimensional[16] ancillary Hilbert space $\mathcal{H}^{\alpha_{n+1}}$.

We define now the vectors $|V^{\to\ell}_{\emptyset,\emptyset}\rangle\!\rangle := |w_{(\emptyset,\ell)}\rangle$.

Furthermore, we define, for each $\mathcal{K}$ (with $\emptyset \subsetneq \mathcal{K} \subseteq \mathcal{N}$) the matrix $W_{\mathcal{K}} := \sum_{k\in\mathcal{K}} W_{(\mathcal{K}\setminus\{k\},k)} \otimes \mathbb{1}^{A^O_k}$ $(= \sum_{\ell\in\mathcal{N}\setminus\mathcal{K}} \mathrm{Tr}_{A^I_\ell} W_{(\mathcal{K},\ell)})$, with $r_{\mathcal{K}} := \operatorname{rank} W_{\mathcal{K}}$, and with its spectral decomposition $W_{\mathcal{K}} = \sum_r |w^p_{\mathcal{K}}\rangle\langle w^p_{\mathcal{K}}|$ in terms of non-normalised, orthogonal vectors $\{|w^p_{\mathcal{K}}\rangle\}_{p=1}^{r_{\mathcal{K}}}$. From that, we then define the vectors

$$|V^{\to\ell}_{\mathcal{K}\setminus\{k\},k}\rangle\!\rangle := \left[\sum_{i=1}^{r_{(\mathcal{K}\setminus\{k\},k)}} |i\rangle^{\alpha_n} \langle w^i_{(\mathcal{K}\setminus\{k\},k)}| \otimes \mathbb{1}^{A^O_k} \cdot W^+_{\mathcal{K}}\right] \otimes \mathbb{1}^{A^I_\ell \alpha_{n+1}} |w_{(\mathcal{K},\ell)}\rangle \ \in \mathcal{H}^{A^O_k \alpha_n A^I_\ell \alpha_{n+1}} \tag{3.67}$$

with $W^+_{\mathcal{K}} := \sum_p \frac{|w^p_{\mathcal{K}}\rangle\langle w^p_{\mathcal{K}}|}{\langle w^p_{\mathcal{K}}|w^p_{\mathcal{K}}\rangle^2}$ the Moore-Penrose pseudoinverse of $W_{\mathcal{K}}$. We can then easily check that

$$\sum_{k\in\mathcal{K}} |w_{(\mathcal{K}\setminus\{k\},k)}\rangle * |V^{\to\ell}_{\mathcal{K}\setminus\{k\},k}\rangle\!\rangle = |w_{(\mathcal{K},\ell)}\rangle \ , \tag{3.68}$$

Indeed, inserting Eq. (3.67) into the left handside of Eq. (3.68) yields $W_{\mathcal{K}} W^+_{\mathcal{K}} \otimes \mathbb{1}^{A^I_\ell \alpha_{n+1}} |w_{(\mathcal{K},\ell)}\rangle$, and the projector $\Pi_{\mathcal{K}} = W_{\mathcal{K}} W^+_{\mathcal{K}}$, tensored with $\mathbb{1}^{A^I_\ell \alpha_{n+1}}$, acts as the identity on $|w_{(\mathcal{K},\ell)}\rangle$.[17]

It follows recursively that each $|w_{(\mathcal{K},\ell)}\rangle$ is composed of the vectors that we constructed here as per Eq. (3.63). In particular, this is true for $|w_{(\mathcal{N},F)}\rangle = |w\rangle$. Thus, the vectors that we constructed here indeed compose the desired process vector as per Proposition 3.8. Furthermore, they satisfy the trace preserving condition (3.57) (because the process matrix W satisfies by assumption the constraints (3.62), which implies Eq. (3.64) and thus Eq. (3.57)). This completes the proof of Proposition 3.9. □

Note that, in general, the corresponding global circuit operations

$$|\tilde{V}_{n+1}\rangle\!\rangle := \sum_{\substack{\mathcal{K}\subset\mathcal{N}\\ |\mathcal{K}|=n}} \sum_{\substack{k\in\mathcal{K}\\ \ell\in\mathcal{N}\setminus\mathcal{K}}} |\tilde{V}^{\to\ell}_{\mathcal{K}\setminus\{k\},k}\rangle\!\rangle \otimes |\mathcal{K}\setminus\{k\},k\rangle^{C'_n} \otimes |\mathcal{K},\ell\rangle^{C_{n+1}} \ \in \mathcal{H}^{T^O_n \alpha_n C'_n T^I_{n+1} \alpha_{n+1} C_{n+1}}, \tag{3.69}$$

that we obtain via the above constructive proof, for some given W that can be decomposed as in Eq. (3.62), act only on some subspace of $\mathcal{H}^{T^O_n \alpha_n \mathcal{C}'_n}$.[18] That is, we have here precisely the situation described in footnote 4, where, at the end of each time-slot t_n, only some nontrivial subspace of the target system, the ancilla and the control system is populated.

[16] Here, that space must be at least of dimension $\max_{(\mathcal{K},\ell)}\{r_{(\mathcal{K},\ell)}\}$.

[17] Similarly to the proofs before, using the constraints Eq. (3.62), we find $\sum_\ell \mathrm{Tr}[(\mathbb{1}^{PA^{IO}_{\mathcal{K}} A^I_\ell \alpha_{n+1}} - W_{\mathcal{K}} W^+_{\mathcal{K}} \otimes \mathbb{1}^{A^I_\ell \alpha_{n+1}}) |w_{(\mathcal{K},\ell)}\rangle\langle w_{(\mathcal{K},\ell)}|] = \mathrm{Tr}(W_{\mathcal{K}} - W_{\mathcal{K}} W^+_{\mathcal{K}} W_{\mathcal{K}}) = 0$, and thus also $(\mathbb{1}^{PA^{IO}_{\mathcal{K}} A^I_\ell \alpha_{n+1}} - W_{\mathcal{K}} W^+_{\mathcal{K}} \otimes \mathbb{1}^{A^I_\ell \alpha_{n+1}}) |w_{(\mathcal{K},\ell)}\rangle = 0$, i.e. $W_{\mathcal{K}} W^+_{\mathcal{K}} \otimes \mathbb{1}^{A^I_\ell \alpha_{n+1}} |w_{(\mathcal{K},\ell)}\rangle = |w_{(\mathcal{K},\ell)}\rangle$.

[18] Indeed, $\mathrm{Tr}_{T^I_{n+1}\alpha_{n+1} C_{n+1}} |\tilde{V}_{n+1}\rangle\!\rangle\langle\!\langle \tilde{V}_{n+1}| = \sum_{\mathcal{K}\subset\mathcal{N};|\mathcal{K}|=n} \sum_{k,k'\in\mathcal{K}} \big(\langle w_{(\mathcal{K}\setminus k,k)}|^{T_{\alpha_n}} \otimes \mathbb{1}^{T^O_n} \cdot W^+_{\mathcal{K}} \cdot \mathbb{1}^{T^O_n} \otimes |w_{(\mathcal{K}\setminus k',k')}\rangle^{T_{\alpha_n}} \big) \otimes |\mathcal{K}\setminus k, k\rangle \langle \mathcal{K}\setminus k', k'|^{C'_n} =: \pi_n$ is a projector onto some subspace of $\mathcal{H}^{T^O_n \alpha_n C'_n}$ (since $\pi_n^2 = \pi_n$).

For $N = 2$ already, we find that there is a difference between classical and quantum control; see the quantum switch, which we discuss in the next subsection as an example of a QC-QC.

The previous case of QC-CC is recovered as a particular case of the present one. Indeed, any process matrix that satisfies the QC-CC constraints 3.6 also satisfies the QC-QC constraints 3.9 (with $W_{(\emptyset,\ell)} = W_{(\ell)}$, and $W_{(\mathcal{K},\ell)} = \sum_{\substack{(k_1,\dots,k_n):\\ \{k_1,\dots,k_n\}=\mathcal{K}}} W_{(k_1,\dots,k_n,\ell)}$). But the class of QC-QCs is larger than that of QC-CCs. QC-QCs can in particular be causally nonseparable, as we will see in the examples below.

3.5.4 Examples

Let us show explicitly how the quantum switch fits into the framework of QC-QC that we introduced above.

Here, we consider the quantum switch with a d-dimensional target system and a 2-dimensional control system, where both the target and the control system can be prepared in a variable input state in the global past. We thus have $N = 2$ operations A_1, A_2 (with their input and output Hilbert spaces $\mathcal{H}^{A_I^k}, \mathcal{H}^{A_O^k}$ of dimension d) together with the global past P and global future F, whose Hilbert spaces are of the form $\mathcal{H}^P = \mathcal{H}^{P_\mathrm{t}} \otimes \mathcal{H}^{P_\mathrm{c}}$ and $\mathcal{H}^F = \mathcal{H}^{F_\mathrm{t}} \otimes \mathcal{H}^{F_\mathrm{c}}$, with $\mathcal{H}^{P_\mathrm{t}}$ and $\mathcal{H}^{F_\mathrm{t}}$ of dimension d, and with two-dimensional $\mathcal{H}^{P_\mathrm{c}}$ and $\mathcal{H}^{F_\mathrm{c}}$ (with their computational bases denoted $\{|1\rangle, |2\rangle\}$ here).

The simplest way to see that the quantum switch can be obtained as a QC-QC, as described above, is by taking

$$\begin{aligned} |V_{\emptyset,\emptyset}^{\to k_1}\rangle\!\rangle &= |\mathbb{1}\rangle\!\rangle^{P_\mathrm{t} A_{k_1}^I} \otimes |k_1\rangle^{P_\mathrm{c}}, \\ |V_{\emptyset,k_1}^{\to k_2}\rangle\!\rangle &= |\mathbb{1}\rangle\!\rangle^{A_{k_1}^O A_{k_2}^I}, \\ |V_{\{k_1\},k_2}^{\to F}\rangle\!\rangle &= |\mathbb{1}\rangle\!\rangle^{A_{k_2}^O F_\mathrm{t}} \otimes |k_1\rangle^{F_\mathrm{c}}. \end{aligned} \tag{3.70}$$

With this choice, and according to Proposition 3.8, we obtain the process matrix of the quantum switch

$$\begin{aligned} &W_\mathrm{switch} = |w_\mathrm{s}\rangle\langle w_\mathrm{s}| \quad \text{with} \\ &|w_\mathrm{s}\rangle := |1\rangle^{P_\mathrm{c}} |\mathbb{1}\rangle\!\rangle^{P_\mathrm{t} A_1^I} |\mathbb{1}\rangle\!\rangle^{A_1^O A_2^I} |\mathbb{1}\rangle\!\rangle^{A_2^O F_\mathrm{t}} |1\rangle^{F_\mathrm{c}} + |2\rangle^{P_\mathrm{c}} |\mathbb{1}\rangle\!\rangle^{P_\mathrm{t} A_2^I} |\mathbb{1}\rangle\!\rangle^{A_2^O A_1^I} |\mathbb{1}\rangle\!\rangle^{A_1^O F_\mathrm{t}} |2\rangle^{F_\mathrm{c}} \\ &\qquad\qquad \in \mathcal{H}^{P_\mathrm{c} P_\mathrm{t} A_1^{IO} A_2^{IO} F_\mathrm{t} F_\mathrm{c}}. \end{aligned} \tag{3.71}$$

The version of the quantum switch described in Sec. 1.5 (Eq. (1.39)) is recovered when the control qubit is prepared in the fixed state $|\varphi_\mathrm{c}\rangle^{P_\mathrm{c}} = \frac{1}{\sqrt{2}}(|1\rangle^{P_\mathrm{c}} + |2\rangle^{P_\mathrm{c}})$ (corresponding to applying the operation $|\varphi_\mathrm{c}\rangle\langle\varphi_\mathrm{c}|^T$), and the target system in the fixed state $|\psi\rangle^{P_t}$ (corresponding to applying the operation $|\psi\rangle\langle\psi|^T$).

This description of the quantum switch as a QC-QC extends quite straightforwardly to simple N-partite generalisations of the quantum switch [25], where one has a coherent superposition of the $N!$ fixed causal orders of the operations $A_1,\ldots,A_N$.

However, the class of QC-QC also contains more general process matrices with qualitatively new features. For instance, as one can see from the characterisation (3.62), tracing out the global future F does not necessarily result in a causally separable process matrix, and neither does it necessarily lead to a convex mixture of valid process matrices where some operation comes last.

In the upcoming paper [32], we will present and analyse more in detail an example of such a new type of 3-operation QC-QC, which exhibits both coherent control of causal order and dynamical causal order.

3.5.5 Correlations generated by quantum circuits with quantum control of causal order

A question that arises naturally is whether the class of process matrices describing QC-QC allows for the generation of correlations that can violate causal inequalities. For the quantum switch, it is known (cf. Section 1.5 and Refs. [11, 15]) that this is not the case, and that although being causally nonseparable, it can only generate causal correlations. It turns out that this result generalises to the entire class of QC-QC.

In Appendix A, we prove that:

Proposition 3.10. *Quantum circuits with quantum control of order can only generate causal correlations.*

In the terminology of Refs. [15, 61], QC-QC define extensibly causal processes.

This result implies, *a fortiori*, that QC-CCs, or QC-FOs, cannot violate causal inequalities; this was however already known, as those are causally separable processes.

3.6 Quantum superinstruments

So far, we have studied quantum circuits that, although taking probabilistic operations as inputs, are by themselves deterministic, that is, they arise from the composition of deterministic operations and can be realised without post-selection. In general, one can however also consider circuits consisting of probabilistic operations that admit several classical outcomes. In this section, we are going to characterise the probabilistic supermaps obtained when allowing for probabilistic circuit operations in the classes that we introduced above. To that end, we replace each CPTP map in the above descriptions by a set of (trace non-increasing) CP maps (each corresponding to a given outcome) that sum up to a CPTP map—i.e., by a quantum instrument.

Each combination of such CP maps defines a "probabilistic quantum circuit", that can be represented by a "probabilistic process matrix", and that can be realised by post-selecting on the corresponding classical outcomes (where the probability of post-selection may depend on the input operations of the circuit). The set of all probabilistic quantum circuits thus obtained, for all possible outcomes, defines a "quantum superinstrument".

In what follows, we shall characterise the superinstruments thus obtained—and thereby their elements, the "probabilistic quantum circuits" or "probabilistic process matrices"—for fixed causal order, as well as for classical and quantum control of causal order.

3.6.1 Quantum superinstruments with fixed causal order

A probabilistic quantum circuit with fixed causal order is thus defined as a combination of CP maps $M_1^{[r_1]} \in \mathcal{L}(\mathcal{H}^{PA_1^I\alpha_1})$, $M_{n+1}^{[r_{n+1}]} \in \mathcal{L}(\mathcal{H}^{A_n^O\alpha_n A_{n+1}^I\alpha_{n+1}})$ for $1 \leq n \leq N-1$, and $M_{N+1}^{[r_{N+1}]} \in \mathcal{L}(\mathcal{H}^{A_N^O\alpha_N F})$, which are composed as in Fig. 3.3, with r_1 etc. denoting the classical outcomes, and with each of the CP maps being part of a quantum instrument—that is, with their sum over the classical outcomes yielding a CPTP map.

To simplify the description, we note that the classical outcomes can always be encoded onto suitable orthogonal states of the ancillary systems in the circuit, and the post-selection can be performed at the end as a part of the last map (before F). This allows us to describe any probabilistic QC-FO without loss of generality as a circuit in which all maps are deterministic, except for the last one, which is a CP map $M_{N+1}^{[r]}$ that is part of a quantum instrument, with the corresponding classical outcome denoted by r. The probabilistic process matrix describing such a circuit is then

$$W^{[r]} = M_1 * M_2 * \cdots * M_N * M_{N+1}^{[r]} \quad \in \ \mathcal{L}(\mathcal{H}^{PA_N^{IO}F}). \tag{3.72}$$

The corresponding quantum superinstrument is the set $\{W^{[r]}\}_r$ of such probabilistic process matrices for all classical outcomes.

The characterisation of QC-FO superinstruments follows directly from that of QC-FOs given by Proposition 3.3.

Proposition 3.11. *A quantum superinstrument with a fixed causal order is represented by a set of positive semidefinite matrices $\{W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA_N^{IO}F})\}_r$, whose sum $W := \sum_r W^{[r]}$ is the process matrix of a quantum circuit with the same fixed causal order (as characterised in Proposition 3.3).*

Conversely, any set of positive semidefinite matrices $\{W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA_N^{IO}F})\}_r$ whose sum is the process matrix of a quantum circuit with a fixed order represents a quantum superinstrument with the same fixed causal order.

Proof. Each $W^{[r]}$ in Eq. (3.72) is positive semidefinite, and their sum $\sum_r W^{[r]} = M_1 * M_2 * \cdots * M_N * \sum_r M_{N+1}^{[r]}$ is indeed the process matrix of a QC-FO as per Proposition 3.2, since

$\sum_r M_{N+1}^{[r]}$ is CPTP. Conversely, given some set $\{W^{[r]}\}_r$ of positive semidefinite matrices whose sum is the process matrix of a QC-FO, we introduce an additional Hilbert space $\mathcal{H}^{F'}$ with computational basis $\{|r\rangle\}_r$. We then define the matrix $W_{\text{ext}} = \sum_r W^{[r]} \otimes |r\rangle\langle r|^{F'} \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}FF'})$, which is the process matrix of a (deterministic) QC-FO with the global future space $\mathcal{H}^{FF'}$ as per Proposition 3.3, since $\text{Tr}_{FF'} W_{\text{ext}} = \text{Tr}_F \sum_r W^{[r]}$, and $\sum_r W^{[r]}$ is by assumption the process matrix of a QC-FO, which thus satisfies the constraints of Proposition 3.3. We can therefore decompose $W_{\text{ext}} = M_1 * M_2 * \cdots * M_N * M_{N+1}$, where $M_{N+1} \in \mathcal{L}(\mathcal{H}^{A_N^O \alpha_N FF'})$ is a CPTP map. We thus obtain a realisation of each $W^{[r]} = W_{\text{ext}} * |r\rangle\langle r|^{F'}$ as a probabilistic QC-FO $W^{[r]} = M_1 * M_2 * \cdots * M_N * M_{N+1}^{[r]}$, with all maps CPTP, except for the last one, which is a CP map $M_{N+1}^{[r]} = M_{N+1} * |r\rangle\langle r|^F$, and with $\sum_r M_{N+1}^{[r]} = \text{Tr}_{F'} M_{N+1}$ a CPTP map.

□

An equivalent proof is also given in Ref. [45] (Theorem 4) for the fixed order case.

A given matrix $W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA_{\mathcal{N}}^{IO}F})$ is therefore the process matrix of a probabilistic QC-FO if and only if it is an element of a quantum superinstrument with fixed causal order $\{W^{[r]}\}_r$, characterised as in Proposition 3.11 above.

3.6.2 Quantum superinstruments with classical control of causal order

To move from a deterministic QC-CC to a quantum superinstrument with classical control of causal order, the deterministic objects that need to be replaced by probabilistic ones are the CPTP maps obtained by summing over the classical outcomes of the circuit instruments $\{M_\emptyset^{\to k_1}\}_{k_1 \in \mathcal{N}}$ etc. Equivalently, one replaces these circuit instruments by more "fine-grained" instruments that produce some additional classical outcomes. A probabilistic quantum circuit with classical control thus starts with a CP map $M_\emptyset^{\to k_1[r_1]} \in \mathcal{L}(\mathcal{H}^{PA_{k_1}^I \alpha_1})$, which is part of a quantum instrument $\{M_\emptyset^{\to k_1[r_1]}\}_{k_1 \in \mathcal{N}; r_1}$ with a classical output value k_1 that determines the first operation, as well as an additional classical output r_1. Similarly, the subsequent circuit maps are given by $M_{(k_1,\ldots,k_n)}^{\to k_{n+1}[r_{n+1}]} \in \mathcal{L}(\mathcal{H}^{A_{k_n}^O \alpha_n A_{k_{n+1}}^I \alpha_{n+1}})$, and belong to instruments $\{M_{(k_1,\ldots,k_n)}^{\to k_{n+1}[r_{n+1}]}\}_{k_{n+1} \in \mathcal{N} \setminus \{k_1,\ldots,k_n\}; r_{n+1}}$[19].

Similarly to the QC-FO case above, the fine-grained outcomes can be encoded in the circuit ancillas, and the post-selection can be deferred to the last map. This gives rise to a QC-CC which is deterministic except for the last maps before F—these get replaced by CP maps $M_{(k_1,\ldots,k_N)}^{\to F[r]}$. The probabilistic process matrix of such a probabilistic QC-CC, conditioned on the classical outcome r, is then (cf. Proposition 3.5):

$$W^{[r]} = \sum_{(k_1,\ldots,k_N)} W_{(k_1,\ldots,k_N,F)}^{[r]} \tag{3.73}$$

[19] The set of additional, fine-grained classical outcomes could in principle be conditioned on $(k_1, \ldots, k_n)$. However, we can always take these sets to be the same by taking their union and appending null elements to the instruments where necessary.

with

$$W^{[r]}_{(k_1,\dots,k_N,F)} \coloneqq M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * M^{\to k_3}_{(k_1,k_2)} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})} * M^{\to F\,[r]}_{(k_1,\dots,k_N)}. \tag{3.74}$$

The corresponding quantum superinstrument is the set $\{W^{[r]}\}_r$ of all such probabilistic process matrices $W^{[r]}$, for all values of r.

Quantum superinstruments with classical control of causal order can be characterised as follows:

Proposition 3.12. *A quantum superinstrument with classical control of causal order is represented by a set of positive semidefinite matrices $\{W^{[r]}\}_r$, where the matrices $W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ can be decomposed in terms of positive semidefinite matrices $W^{[r]}_{(k_1,\dots,k_N,F)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ and $W_{(k_1,\dots,k_n)} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\{k_1,\dots,k_{n-1}\}}A^{I}_{k_n}})$, in such a way that*

$$\forall\, r, \quad W^{[r]} = \sum_{(k_1,\dots,k_N)} W^{[r]}_{(k_1,\dots,k_N,F)}, \tag{3.75}$$

and such that the matrices $W_{(k_1,\dots,k_N,F)} \coloneqq \sum_r W^{[r]}_{(k_1,\dots,k_N,F)}$ and $W_{(k_1,\dots,k_n)}$ satisfy Eq. (3.27) of Proposition 3.6.

Conversely, any set $\{W^{[r]}\}_r$ of positive semidefinite matrices with the properties above represents a quantum superinstrument with classical control of causal order.

The proof extends directly from that of Proposition 3.6:

Proof. For a quantum superinstrument composed of circuit operations as per Eqs. (3.73)–(3.74), one has $\sum_r W^{[r]} = \sum_{(k_1,\dots,k_N)} \sum_r W^{[r]}_{(k_1,\dots,k_N,F)} = \sum_{(k_1,\dots,k_N)} M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * M^{\to k_3}_{(k_1,k_2)} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})} * \sum_r M^{\to F\,[r]}_{(k_1,\dots,k_N)}$, which is the process matrix of a QC-CC as per Proposition 3.5, since $\sum_r M^{\to F\,[r]}_{(k_1,\dots,k_N)}$ is a CPTP map. Conversely, for some set of positive semidefinite matrices $W^{[r]}$ that can be decomposed in terms of positive semidefinite matrices $W^{[r]}_{(k_1,\dots,k_N,F)}$ as per Proposition 3.12, we define again the "extended" matrix $W_{\text{ext}} = \sum_r W^{[r]} \otimes |r\rangle\langle r|^{F'} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}FF'})$ by introducing an additional Hilbert space $\mathcal{H}^{F'}$ with computational basis $\{|r\rangle\}_r$. We note now that W_{ext} is the process matrix of a (deterministic) QC-CC with global output space $H^{FF'}$[20], since it has a decomposition as in Proposition 3.6 (with $W_{(k_1,\dots,k_N,FF')} \coloneqq \sum_r W^{[r]}_{(k_1,\dots,k_N,F)} \otimes |r\rangle\langle r|^{F'}$ and $\mathrm{Tr}_{FF'}\, W_{(k_1,\dots,k_N,FF')} = \mathrm{Tr}_F\, W_{(k_1,\dots,k_N,F)}$, which satisfies the corresponding constraints by assumption). Therefore, $W_{(k_1,\dots,k_N,FF')} = M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * M^{\to k_3}_{(k_1,k_2)} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})} * M^{\to FF'}_{(k_1,\dots,k_N)}$, where $M^{\to FF'}_{(k_1,\dots,k_N)} \in \mathcal{L}(\mathcal{H}^{A^{O}_{k_N}\alpha_N FF'})$ is CPTP, and $W^{[r]}_{(k_1,\dots,k_N,F)} = M^{\to k_1}_{\emptyset} * M^{\to k_2}_{(k_1)} * M^{\to k_3}_{(k_1,k_2)} * \cdots * M^{\to k_N}_{(k_1,\dots,k_{N-1})} * M^{\to F\,[r]}_{(k_1,\dots,k_N)}$, with $M^{\to F\,[r]}_{(k_1,\dots,k_N)} = M^{\to FF'}_{(k_1,\dots,k_N)} * |r\rangle\langle r|^{F'}$. □

[20] Note that this is not true in general if we do not impose any further constraints on the matrices $W^{[r]}$. That is, for some set $\{W^{[r]}\}_r$ whose elements sum up to the process matrix of a QC-CC, the extended matrix W_{ext}, as defined here, is not necessarily the process matrix of a deterministic QC-CC (see also the example at the end of Sec. 3.6.3). This is in contrast with QC-FO and QC-QC superinstruments, as well as those for general process matrices, where no additional constraints on the matrices $W^{[r]}$ arise.

We then have that a given matrix $W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ is the process matrix of a probabilistic QC-CC if and only if it is an element of a quantum superinstrument with classical control of causal order $\{W^{[r]}\}_r$, characterised as in Proposition 3.12 above.

In fact, we already encountered an example of a QC-CC superinstrument—the matrices $W_{(k_1,\ldots,k_N,F)}$ in Sec. 3.4 are probabilistic process matrices, and the collection $\{W_{(k_1,\ldots,k_N,F)}\}_{(k_1,\ldots,k_N)}$ defines the corresponding quantum superinstrument, which describes the case where the additional fine-grained outcomes, on which we post-select to obtain the probabilistic QC-CC, are precisely taken to be the outcomes $k_1,\ldots,k_N$ that determine the order of the operations. (That is, the circuit instruments are $\{M_\emptyset^{\to k_1[r_1]}\}_{k_1\in\mathcal{N};r_1\in\mathcal{N}}$, with $M_\emptyset^{\to k_1[r_1]} = \delta_{r_1k_1}\, M_\emptyset^{\to k_1}$ and similarly for the subsequent ones).

3.6.3 Quantum superinstruments with quantum control of causal order

To obtain a probabilistic quantum circuit with quantum control of causal order, we again replace all deterministic circuit operations in our description of QC-QC by quantum instruments. Similarly to Sec. 3.5, we can again assume without loss of generality that the CP maps constituting these instruments consist of the application of one single Kraus operator, since the circuit ancillas can be used to purify all maps accordingly. Every operation in Eqs. (3.49)–(3.51) therefore gets replaced by the Kraus operator of some CP map that is part of an instrument.

As above, we can encode the classical outcomes in the circuit ancillas, and perform the post-selection at the end as a part of the last operation. Therefore, any probabilistic QC-QC can be described as a deterministic one, where only the last operation $|\tilde{V}_{N+1}\rangle\!\rangle$ (Eq. (3.54)) gets replaced by

$$|\tilde{V}^{[r]}_{N+1}\rangle\!\rangle := \sum_{k_N} |\tilde{V}^{\to F\,[r]}_{\mathcal{N}\setminus\{k_N\},k_N}\rangle\!\rangle \otimes |\mathcal{N}\setminus\{k_N\},k_N\rangle^{C'_N}\,, \tag{3.76}$$

such that $\sum_r |\tilde{V}^{[r]}_{N+1}\rangle\!\rangle\langle\!\langle\tilde{V}^{[r]}_{N+1}|$ is trace-preserving—that is, (cf. Eq. (3.57)),

$$\| \, |\tilde{A}_N\rangle\!\rangle * |\tilde{V}_N\rangle\!\rangle * |\tilde{A}_{N-1}\rangle\!\rangle * \cdots * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \, \|^2 = \sum_r \| \, |\tilde{V}^{[r]}_{N+1}\rangle\!\rangle * |\tilde{A}_N\rangle\!\rangle * |\tilde{V}_N\rangle\!\rangle * |\tilde{A}_{N-1}\rangle\!\rangle * \cdots * |\tilde{V}_1\rangle\!\rangle * |\psi\rangle \, \|^2, \tag{3.77}$$

for all states $|\psi\rangle$ and operations A_k.

Similarly to Proposition 3.8, the probabilistic process matrix corresponding to a probabilistic quantum circuit with quantum control is then

$$W^{[r]} = \mathrm{Tr}_{\alpha_{N+1}} |w^{[r]}\rangle\langle w^{[r]}| \text{ with}$$
$$|w^{[r]}\rangle := \sum_{(k_1,\dots,k_N)} |V^{\to k_1}_{\emptyset,\emptyset}\rangle\rangle * |V^{\to k_2}_{\emptyset,k_1}\rangle\rangle * |V^{\to k_3}_{\{k_1\},k_2}\rangle\rangle * \cdots * |V^{\to k_N}_{\{k_1,\dots,k_{N-2}\},k_{N-1}}\rangle\rangle * |V^{\to F\,[r]}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\rangle. \tag{3.78}$$

The following proposition characterises quantum superinstruments with quantum control of causal order:

Proposition 3.13. *A quantum superinstrument with quantum control of causal order is represented by a set of positive semidefinite matrices $\{W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})\}_r$, whose sum $W := \sum_r W^{[r]}$ is the process matrix of a quantum circuit with quantum control of causal order (as characterised in Proposition 3.9).*

Conversely, any set of positive semidefinite matrices $\{W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})\}_r$ whose sum is the process matrix of a quantum circuit with quantum control of causal order represents a quantum superinstrument with quantum control of causal order.

The proof extends directly from that of Proposition 3.9.

Proof. Consider some QC-QC superinstrument as described above, with probabilistic process matrices as given in Eq. (3.78). It is straightforward to see that $\sum_r |w^{[r]}\rangle \otimes |r\rangle^{F'}$, with an additional Hilbert space $\mathcal{H}^{F'}$ with computational basis $\{|r\rangle\}_r$, is the process vector of a deterministic QC-QC with the last circuit operation $\sum_r |\tilde{V}^{[r]}_{N+1}\rangle\rangle \otimes |r\rangle$. Its process matrix $\mathrm{Tr}_{\alpha_{N+1}}\left[\sum_{rr'} |w^{[r]}\rangle\langle w^{[r']}| \otimes |r\rangle\langle r'|^{F'}\right]$ thus satisfies the QC-QC characterisation 3.9, and so does $\mathrm{Tr}_{\alpha_{N+1}F'}\left[\sum_{rr'} |w^{[r]}\rangle\langle w^{[r']}| \otimes |r\rangle\langle r'|^{F'}\right] = \sum_r W^{[r]} = W$.

Conversely, consider a set $\{W^{[r]}\}_r$ of positive semidefinite matrices that sum up to the process matrix of a QC-QC. Analogously to the proofs for QC-FO and QC-CC superinstruments, we define the matrix $W_{\text{ext}} := \sum_r W^{[r]} \otimes |r\rangle\langle r|^{F'} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}FF'})$, , which is the process matrix of a (deterministic) QC-QC as per Proposition 3.9, with the additional output space $\mathcal{H}^{F'}$. We can thus find an explicit implementation as per Proposition 3.8 of W_{ext}, i.e., we have that $W_{\text{ext}} = \mathrm{Tr}_{\alpha_{N+1}} |w_{\text{ext}}\rangle\langle w_{\text{ext}}|$, with

$$|w_{\text{ext}}\rangle := \sum_{k_1,\dots,k_N} |V^{\to k_1}_{\emptyset,\emptyset}\rangle\rangle * |V^{\to k_2}_{\emptyset,k_1}\rangle\rangle * |V^{\to k_3}_{\{k_1\},k_2}\rangle\rangle * \cdots * |V^{\to k_N}_{\{k_1,\dots,k_{N-2}\},k_{N-1}}\rangle\rangle * |V^{\to FF'}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\rangle \tag{3.79}$$

Each probabilistic process matrix $W^{[r]}$ can then be obtained as in Eq. (3.78), with $|V^{\to F\,[r]}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\rangle = |V^{\to FF'}_{\{k_1,\dots,k_{N-1}\},k_N}\rangle\rangle * |r\rangle^{F'}$. □

We then have that a given matrix $W^{[r]} \in \mathcal{L}(\mathcal{H}^{PA^{IO}_{\mathcal{N}}F})$ is the process matrix of a probabilistic QC-QC if and only it is an element of a quantum superinstrument with quantum control of causal order $\{W^{[r]}\}_r$, characterised as in Proposition 3.13 above.

Similarly to the example that we discussed above for the QC-CC case, the matrices $W_{(\mathcal{K}\setminus\{k_N\},k_N)}$ in Sec. 3.5 are probabilistic process matrices that constitute the quantum superinstrument $\{W_{(\mathcal{K}\setminus\{k_N\},k_N)}\}_{k_N}$. They arise when the quantum instrument at the end of the circuit measures the control system C'_N (while also transforming the output state of the last operation and the ancilla), that is, by taking $r \in \mathcal{N}$ and $|V^{\to F\,[r]}_{\mathcal{N}\setminus\{k_N\},k_N}\rangle\!\rangle = \delta_{rk_N}\,|V^{\to F}_{\mathcal{N}\setminus\{k_N\},k_N}\rangle\!\rangle$ in Eq. (3.76).

As another example, let us once again consider the quantum switch, with its process matrix description Eq. (3.71), and where the control qubit is measured at the end of the circuit in the basis $\{|+\rangle, |-\rangle\}$, with $|\pm\rangle := \frac{|1\rangle \pm |2\rangle}{\sqrt{2}}$. The circuit operations that constitute the corresponding quantum superinstrument as per Eq. (3.78) are $|V^{\to k_1}_{\emptyset,\emptyset}\rangle\!\rangle = |\mathbb{1}\rangle\!\rangle^{P_{\mathrm{t}}A^I_{k_1}} \otimes |k_1\rangle^{P_{\mathrm{c}}}$ and $|V^{\to k_2}_{\emptyset,k_1}\rangle\!\rangle = |\mathbb{1}\rangle\!\rangle^{A^O_{k_1}A^I_{k_2}}$ as in Eq. (3.70), and $|V^{\to F[\pm]}_{\{k_1\},k_2}\rangle\!\rangle = \frac{(\pm 1)^{k_2}}{\sqrt{2}}\,|\mathbb{1}\rangle\!\rangle^{A^O_{k_2}F_{\mathrm{t}}}$.

The superinstrument is therefore $\{W^{[+]}, W^{[-]}\}$, with $W^{[\pm]} = |w^{[\pm]}\rangle\langle w^{[\pm]}|$, and

$$|w^{[\pm]}\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle^{P_{\mathrm{c}}}\,|\mathbb{1}\rangle\!\rangle^{P_{\mathrm{t}}A^I_1}\,|\mathbb{1}\rangle\!\rangle^{A^O_1A^I_2}\,|\mathbb{1}\rangle\!\rangle^{A^O_2F_{\mathrm{t}}} \pm |2\rangle^{P_{\mathrm{c}}}\,|\mathbb{1}\rangle\!\rangle^{P_{\mathrm{t}}A^I_2}\,|\mathbb{1}\rangle\!\rangle^{A^O_2A^I_1}\,|\mathbb{1}\rangle\!\rangle^{A^O_1F_{\mathrm{t}}}\right) \tag{3.80}$$

We then have that $W = W^{[+]} + W^{[-]} = \mathrm{Tr}_{F_{\mathrm{c}}}\, W_{\mathrm{switch}}$ is indeed a QC-QC process matrix. In fact, it is even a QC-CC process matrix, since it is causally separable. Nevertheless, $\{W^{[+]}, W^{[-]}\}$ is not a QC-CC superinstrument; in order to realise it, the two causal orders need to "interfere" in the switch before the control qubit is measured.

Indeed, the matrices $W^{[\pm]}$ do not satisfy the additional constraints in Proposition 3.12. That is, $W^{[\pm]}$ cannot be decomposed as $W^{[\pm]} = W^{[\pm]}_{(1,2,F)} + W^{[\pm]}_{(2,1,F)}$, such that, with $W_{(1,2,F)} = W^{[+]}_{(1,2,F)} + W^{[-]}_{(1,2,F)}$ and $W_{(2,1,F)} = W^{[+]}_{(2,1,F)} + W^{[-]}_{(2,1,F)}$, we obtain a decomposition of $W = W_{(1,2,F)} + W_{(2,1,F)}$ as in Proposition 3.6. (This follows from the fact that $W^{[\pm]}$ are rank one projectors, and can therefore not be further decomposed into a (nontrivial) sum of positive semidefinite matrices, and neither $W^{[+]}$ nor $W^{[-]}$ satisfies individually the constraints on either $W_{(1,2,F)}$ or $W_{(2,1,F)}$ in Eq. (3.27)).

Note that the example we described here is precisely the superinstrument that we use in the discrimination task between commuting and anticommuting unitaries (cf. Section 1.6). Indeed, consider the case where we prepare the control qubit in the state $|+\rangle^{P_{\mathrm{c}}}$ and the target system in some fixed state $|\psi\rangle^{P_{\mathrm{t}}}$, take the input operations of the quantum superinstrument to be unitaries U_1 and U_2, and discard F_{t}. The output of the superinstrument

$$p^{[\pm]} := \mathrm{Tr}\left[W^{[\pm]} \cdot (|+\rangle\langle+|^T)^{P_{\mathrm{c}}} \otimes (|\psi\rangle\langle\psi|^T)^{P_{\mathrm{t}}} \otimes (|U_1\rangle\!\rangle\langle\!\langle U_1|^T)^{A^{IO}_1} \otimes (|U_2\rangle\!\rangle\langle\!\langle U_2|^T)^{A^{IO}_2} \otimes \mathbb{1}^{F_{\mathrm{t}}}\right] \tag{3.81}$$

is then simply the probability to obtain the outcome "+" or "−". We find that $p^{[+]} = 1$ and $p^{[-]} = 0$ if the two unitaries U_1 and U_2 commute, and, vice versa, $p^{[-]} = 1$ and $p^{[+]} = 0$ if they anticommute, allowing us to discriminate between the two cases with certainty.

We thus recover straightforwardly this known advantage of the quantum switch over causally separable processes. However, the characterisation of the full class of QC-QC and the corresponding superinstruments now also allow us to go beyond that simple, canonical example, and to search for new applications of physically realisable, causally nonseparable processes in a systematic way. In the next section, we present an illustrative example.

3.7 Applications

The discrimination between commuting and anticommuting unitaries considered above is an example of a "higher-order quantum computation" problem, i.e., a task with the aim of generating some desired outcome from some unknown, black-box input operations. Other examples of such tasks that have been studied before are, for instance, the cloning [56], the storage and retrieval [56], or the replication of the inverse or transpose [57, 62] of some undisclosed operation of which one or multiple copies are available.

Here, we are particularly interested in tasks for which QC-QC provide an advantage over QC-CC. In order to quantify how a given class of circuits performs for some task, one needs to find the circuit that maximises some figure of merit, such as for example the channel fidelity between the desired "target" channel and the output of the supermap, or the success probability of the task. The characterisation of circuits in terms of process matrices allows to conveniently perform this optimisation by semidefinite programming. This approach has been used, for instance, in Ref. [58] to compare superinstruments where the input operations are applied in parallel[21] with QC-FO superinstruments for certain specific tasks. In Ref. [62], where the exact, probabilistic inversion of an unknown unitary operation was studied, general quantum superinstruments[22] were also considered. The characterisations we gave here now enable us to include the new classes of QC-CC and QC-QC in the comparison, and, in particular, to identify tasks for which QC-QC perform better than QC-CC.

We now present a concrete example of a task for which we find indeed a higher success probability for indefinite causal orders, and in particular a gap between QC-CC and QC-QC. It is based on an equivalence determination task studied in [58], where one has two *reference boxes* that implement black-box qubit unitary operations U_1 and U_2, and a *target box* that implements either U_1 or U_2 with probability 1/2. The aim is to determine which of the two operations is implemented by the target box, while using each of the three boxes once. We denote the

[21] In that case, the constraint is that the positive semidefinite matrices $W^{[r]}$ constituting the superinstrument sum up to a matrix of the form $\sum_r W^{[r]} = W^I \otimes \mathbb{1}^{A_1^O A_2^O A_3^O}$ with $W^I \in \mathcal{L}(\mathcal{H}^{A_1^I A_2^I A_3^I})$ (see Ref. [58] for the detailed characterisation).

[22] These are defined as a set of positive semidefinite matrices $\{W^{[r]} \in \mathcal{L}(\mathcal{H}^{P A_{\mathcal{N}}^{IO} F})\}_r$, whose sum $W := \sum_r W^{[r]}$ is a (general) valid process matrix [62].

input and output spaces of the reference boxes as $\mathcal{H}^{A_i^I}$ and $\mathcal{H}^{A_i^O}$ (with $i = 1, 2$), those of the target box as $\mathcal{H}^{A_3^I}$ and $\mathcal{H}^{A_3^O}$, and the quantum superinstrument that we have at our disposal by $\{W^{[i]}\}_{i=1,2}$, with $W^{[i]} \in \mathcal{L}(\mathcal{H}^{A_1^{IO} A_2^{IO} A_3^{IO}})$, where P and F are trivial, and the outcome i of the superinstrument corresponds to our guess of which operation is implemented. The success probability (for some specific U_1 and U_2) is then

$$p_{U_1,U_2} = \frac{1}{2} \sum_{i=1,2} W^{[i]} * \Pi_i = \frac{1}{2} \sum_{i=1,2} \mathrm{Tr}(W^{[i]} \cdot \Pi_i^T) \tag{3.82}$$

with $\Pi_i = |U_1\rangle\!\rangle\langle\!\langle U_1| \otimes |U_2\rangle\!\rangle\langle\!\langle U_2| \otimes |U_i\rangle\!\rangle\langle\!\langle U_i|$.

For the case of completely unknown U_1 and U_2, Eq. (3.82) needs to be averaged over the Haar measure. Averaging numerically over many pairs of Haar random unitaries, and optimising the superinstrument in Eq. (3.82) by SDP, under the constraint that it belongs to the respective class whose performance we want to determine, we did not find any advantage of indefinite over definite causal orders for this particular task. We recovered the success probability of 0.875 found in Ref. [58] for superinstruments with an application of the boxes in parallel as well as for (general) QC-FO quantum superinstruments, and we found the same value for QC-QC superinstruments (and also for general quantum superinstruments).

We therefore considered the case where we have additional information about the unitary operations, and where they are drawn from some given, finite set $\mathcal{U} = \{U_1, \ldots, U_N\}$ with uniform probability. We considered various sets $\mathcal{U}$ of simple one-qubit gates and indeed found a gap between the different classes for some choices. Two examples exhibiting a significant difference between QC-CC and QC-QC, as well as between QC-QC and general indefinite causal order, are shown in table 3.1.

$\mathcal{U}$	PAR	QC-FO	QC-CC	QC-QC	IND
$\{\mathbb{1}, R_X, R_Y, S\}$	0.7738	0.7910	0.7910	0.7939	0.7942
$\{Y, R_X, R_Y, T\}$	0.7959	0.8080	0.8080	0.8093	0.8101

TABLE 3.1: Success probabilities for the equivalence determination of unitaries from a given set $\mathcal{U}$. Here, Y is the Pauli-Y gate, $R_X = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & -i \\ -i & 1 \end{smallmatrix}\right)$ and $R_Y = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ are Bloch sphere rotations by $\pi/2$ around the x and y axis, respectively, S is the phase gate $\left(\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right)$ and T is the phase shift gate $\left(\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{smallmatrix}\right)$. PAR stands for application of the boxes in parallel, QC-FO is the maximum of the three success probabilities for the possible fixed order configurations (i.e., with the target box before, between or after the reference boxes), QC-CC and QC-QC denote the class of superinstruments with classical and quantum control of causal order, respectively, and IND stands for general quantum superinstruments.

We also found advantages in other variants of the above equivalence determination task. For instance, we considered reference and target boxes that perform general quantum channels (CPTP maps) instead of unitary operations. In that case, one has $\Pi_i = C_1 \otimes C_2 \otimes C_i$ in Eq. (3.82). Averaging over a large number of random pairs of quantum channels (which we generated with *qetlab* [63]), we found success probabilities of 0.5666 for parallel application of the black boxes, 0.5667 in the fixed-order case, 0.5714 for QC-QC and 0.5715 for general indefinite causal order.

We also studied another variant of the task where the target box implements two different functions of the reference boxes, i.e., we have $U_i = f_1(U_1, U_2)$ or $U_i = f_2(U_1, U_2)$ (instead of $U_i = U_1$ or $U_i = U_2$), and we indeed found higher success probabilities for QC-QC for some choices (when averaging numerically over a large number of Haar random unitaries). Three examples are given in table 3.2.

$f_1(U_1, U_2)$	$f_2(U_1, U_2)$	PAR	QC-FO	QC-CC	QC-QC	IND
$U_1 U_2 U_1^\dagger$	$U_2 U_1 U_2^\dagger$	0.6169	0.6681	0.6681	0.6854	0.6868
$U_1 U_1$	$U_2 U_2$	0.6597	0.6814	0.6856	0.6926	0.6932
$U_1 U_2 U_1$	$U_1 U_2 U_2$	0.6403	0.7070	0.7070	0.7133	0.7138

TABLE 3.2: Success probabilities for the discrimination task between two functions $f_1(U_1, U_2)$ and $f_2(U_1, U_2)$ implemented by the target box.

We also started investigating tasks where one is given several copies of a unitary operation U, and the aim is to generate some function of U, such as for instance its inverse $U^\dagger$, or its transpose U^T. One can consider variants of these tasks where the desired transformation is realised approximately (i.e., one optimises the channel fidelity [57]) or perfectly, but only with some finite probability of success, which is to be optimised [62, 64]. One can furthermore consider variants where the input state of the circuit is provided only after the black-box operations have been processed (so-called *quantum learning protocols* [65, 66] or *delayed input state protocols* [62]). In some first numerical tests, we could not identify any gaps between QC-QC and QC-CC in these kind of tasks, although some gaps between general process matrices and QC-QC could be found. However, more detailed studies are needed, and it needs to be clarified whether, similarly to the example discussed above, advantages of QC-QC can be identified when additional information about the unitary operations is available.

3.8 Discussion

In this chapter, we introduced two new classes of circuits, those of quantum circuits with classical and quantum control. We described these circuits, and characterised the sets of process matrices they define, generalising similar results for quantum circuits with fixed causal order, or quantum combs.

Using similar techniques as for witnesses of causal nonseparability, one can straightforwardly verify whether some process matrix is in a given class or not. For instance, one can show that the classical switch does not have a fixed order, that the quantum switch does not have a classical control, or that the process matrix W^{ocb} considered in Sec. 1.2.3 or the tripartite "classical" example of Baumeler *et al.* [67] are not of the quantum control type (note that this also follows from the fact that these can violate causal inequalities, while we showed that QC-QCs cannot; in the case of W^{ocb}, the problem again reduces to a causal witness (cf. Sec. 1.2.3), since for $N = 2$ and trivial global past and future systems, QC-QCs reduce to simple classical mixtures of the two possible orders).

Our study here represents a "bottom up approach" where, starting from some basic assumptions, we constructively obtain a large class of process matrices with a physical interpretation. In particular, an important assumption that we made in our construction, and that allowed us to indeed obtain valid quantum supermaps, is that each input operation of the circuit is applied once and only once. We point out, however, that in the case of quantum control, with the time of the operations being controlled coherently, this assumption is not so clear, and, in this connection, there is some debate on whether these protocols are "realisations" or "simulations" of quantum processes with indefinite causal order. We will return to this point in Chapter 7, where we will show that the statement that each operation is applied once and only once can be made more precise by resorting to the notion of *time-delocalised subsystems* [38].

In the fixed-order case, one finds that the class of circuits than one obtains through the constructive approach matches with the most general process matrices that respect the corresponding fixed causal order. Here, we find a gap between what we obtain from our constructive ("bottom-up") approach, and what is obtained with an axiomatic ("top-down") approach by just imposing some consistency constraints, which leads to the class of general process matrices. This raises the question of where this gap comes from. Is it just due to a lack of imagination, or is there some deep underlying reason for it?

What other interesting properties does the set of QC-QC process matrices we identified here have? Could it be true, for instance, that it contains all the process matrices that cannot violate causal inequalities?

On the more applied side, an interesting question is which experimental platforms are most suitable in order to realise new examples of QC-QC, beyond the quantum switch, concretely in the lab. For instance, photonic implementations similarly to those of the switch, with spatially separate "boxes" realising the operations A_k, and the control system including the path, are conceivable.

Furthermore, we found that our approach here is very useful in order to more systematically investigate the question of what advantages can be obtained for quantum information processing with physically realisable processes beyond fixed order quantum circuits. Studying various tasks, and identifying those where quantum circuits with quantum control yield an advantage, will shed light on this question.

Chapter 4

Communication through coherent control of quantum channels

This chapter is based on the publication:
Alastair A. Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, and Cyril Branciard. Communication through coherent control of quantum channels, *arXiv:1810.09826 [quant-ph]* (Ref. [34]).

Secs. 4.1 to 4.4 and 4.7 correspond to the main text of the paper, with some minor modifications to avoid repetitions and to integrate the paper into the broader context of the thesis. Secs. 4.5 and 4.6 are taken from the appendix of [34]. My main contribution to this paper was to perform some initial calculations, together with Alastair Abbott and Cyril Branciard, which led to the main results of [34]. I also developed Sec. 4.5 and an earlier version of the proof of Eq. (4.9).

As mentioned in Sec. 1.6, it was found that the quantum switch enables some surprising classical and quantum channel capacity enhancement effects. When certain zero-capacity channels are inserted into the quantum switch, the resulting global transformation still allows information to be transmitted, in a way that is impossible if their causal order is fixed or classically controlled. More particularly, in a recent letter [29] it was shown that for two completely depolarizing channels, information about the initial state of the target system can be acquired after the passage through the switch, despite the fact that a completely depolarizing channel by itself (and, by extension, any causally ordered composition of two such channels) always outputs a completely mixed state. This communication advantage has also been confirmed experimentally [21, 22].

It is perhaps tempting to attribute this result to the indefinite causal order between the channels. In this chapter, we present the results from Ref. [34], where we showed that the same phenomenon can be observed in an even simpler setup, where one coherently controls between applying one of two identical completely depolarising channels to the target system. Such a situation involves no indefinite causal order; we argue that these results should therefore instead be understood as resulting from coherent control of quantum communication channels.

Motivated by this example, we then revisit more generally the notion of coherent control of arbitrary quantum channels by exploiting a control system to determine which channel is used to transmit a state rather than the order in which two communication channels are used.

It turns out that—in contrast to the quantum switch [3, 29]—the situation where channels are controlled coherently cannot be described in terms of a quantum supermap, and does therefore not fit into the framework we set up in Chapter 3. In particular, the action of the global channel in this case depends not only on the descriptions of the individual channels as completely positive trace-preserving (CPTP) maps—i.e., their standard description as "quantum channels" [30, 31, 68]—but also on more fine-grained information about their specific realisations, which includes, but goes beyond, relative phase information; in general, we find that, in order to completely describe the action of a coherently controlled quantum channel, one needs to specify a "transformation matrix" depending on the implementation of the channel, in addition to its description (e.g., in terms of Kraus operators). This additional information is in fact accessible in the output state of the joint control-target system and allows two different implementations of what is usually considered to be the same channel to therefore be differentiated. These results highlight the subtleties involved in describing "controlled channels", and show that, without specifying this extra information on the specific channel implementation the notion is in fact ill-defined.

4.1 Communication through the "depolarising quantum switch"

Let us first recall the case of the quantum switch with two completely depolarising channels, as studied in Ref. [29].

We consider once again the quantum switch, with a d-dimensional target system and a two-dimensional control system, which we initialise in a fixed state $|+\rangle^c = \frac{1}{\sqrt{2}}(|0\rangle^c + |1\rangle^c)$, as in Sec. 1.5. Viewed as a quantum supermap, the quantum switch then takes two input operations $\mathcal{C}_0$ and $\mathcal{C}_1$ (which we take to be CPTP maps here), and applies them to the target system in a superposition of orders. The result is a global channel $\mathcal{S}[\mathcal{C}_0, \mathcal{C}_1]$ from the global input target system to the joint control-target global output system. In Ref. [29], it was observed that, if the CPTP maps $\mathcal{C}_i$ are fully depolarising channels $\mathcal{N}_i$ (which map any initial target state ρ^t_{in} to the maximally mixed state $\frac{\mathbb{1}^t}{d}$), then the global channel $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$ maps ρ^t_{in} to the joint control-target state

$$\rho^{ct}_{\text{out}} = \frac{\mathbb{1}^c}{2} \otimes \frac{\mathbb{1}^t}{d} + \frac{1}{2}\left[|0\rangle\langle 1|^c + |1\rangle\langle 0|^c\right] \otimes \frac{1}{d^2}\rho^t_{\text{in}}\,, \tag{4.1}$$

which is *not* $\frac{\mathbb{1}^t}{d}$ but instead retains some dependence on ρ^t_{in}. Thus, information can propagate through the "depolarising quantum switch" despite this being impossible for the channels $\mathcal{N}_0$, $\mathcal{N}_1$, $\mathcal{N}_1 \circ \mathcal{N}_0$, and $\mathcal{N}_0 \circ \mathcal{N}_1$ individually. This surprising result arising in the presence of indefinite causal order has recently been subject to experimental confirmation [21, 22], and generalised to more complex setups that put more channels in a superposition of different orders [69].

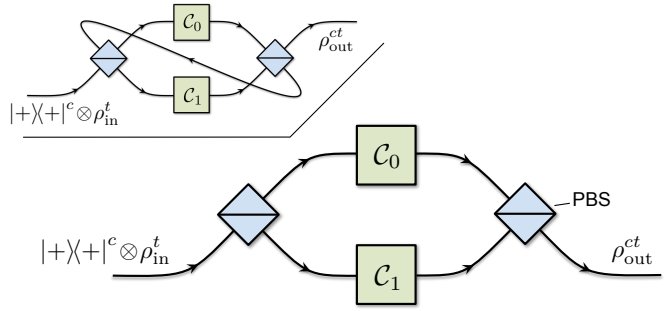

FIGURE 4.1: The inset shows a typical photonic implementation of the quantum switch [3, 25], in which the control qubit is encoded in the polarisation of a photon which is routed by polarising beamsplitters (PBS), and the target system is encoded in some internal degree of freedom of the photon (as e.g. in Refs. [19, 22]). In Ref. [34] we consider only the "first half" of the quantum switch process (main figure). This implements a coherent control between the two boxes implementing $\mathcal{C}_0$ and $\mathcal{C}_1$ that the target system traverses. As we show, the above diagram is in fact *ill-defined* since, when controlled coherently, the maps $\mathcal{C}_i$ do not fully determine the output state ρ^{ct}_{out}; see Fig. 4.2.

Note that, since the switch is a quantum supermap, the resulting channels $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$ (and more generally, $\mathcal{S}[\mathcal{C}_0, \mathcal{C}_1]$) depend only on the input CPTP maps—something that will no longer be true in the setup with coherently controlled channels considered below.

4.2 Communication through coherently-controlled depolarising channels

Remember from Sec. 1.5 that in a standard interferometric implementation of the quantum switch (such as in Refs. [16–22]), the target system is routed to the switched operations, which here correspond to communication channels, via some beamsplitters. In the following, we shall consider instead the state of the joint control-target system after traversing only half of such a quantum switch; that is, after the target system has passed, in a superposition, through the communication channels only a single time. This situation, a possible implementation of which is shown in Fig. 4.1, amounts to coherently controlling between applying the operations implementing $\mathcal{C}_0$ or $\mathcal{C}_1$ to the target system. By preparing the control qubit in the state $|+\rangle^c$, a "superposition" of the two operations is thus applied[1].

We note that in general, coherent control of completely unknown quantum operations is impossible. However, in interferometer-type situations, the setup provides additional information about the structure of the joint control-target Hilbert space that makes such control possible [70–73]; indeed, coherent control of unitary operations by such means has been demonstrated experimentally in many scenarios [74–78].

Let us consider, as in Ref. [29], the case where the two operations implement fully depolarising channels ($\mathcal{C}_i = \mathcal{N}_i$), and consider first the concrete case analysed in Ref. [29] and implemented

[1] Our calculations below generalise easily to any other initial state of the control (not necessarily $|+\rangle^c$).

experimentally (for a qubit target system) in Ref. [22] where these are realised by randomising over a set of d^2 orthogonal unitary operators $\{U_i\}_{i=0}^{d^2-1}$. For each channel, one then indeed has $\mathcal{N}_{0/1}(\rho_{\text{in}}^t) = \frac{1}{d^2}\sum_i U_i \rho_{\text{in}}^t U_i^\dagger = \frac{\mathbb{1}^t}{d}$.

For each random choice of unitary operators (U_i, U_j), the control-target system therefore undergoes the unitary evolution $|0\rangle\langle 0|^c \otimes U_i + |1\rangle\langle 1|^c \otimes U_j$ [2]. If the control qubit is initially in the state $|+\rangle^c$ and the target system is in some state $|\psi_{\text{in}}\rangle^t$, the joint system thus evolves to

$$|\Phi_{ij}\rangle^{ct} = \frac{1}{\sqrt{2}}\left(|0\rangle^c \otimes U_i\,|\psi_{\text{in}}\rangle^t + |1\rangle^c \otimes U_j\,|\psi_{\text{in}}\rangle^t\right). \tag{4.2}$$

Averaging over all choices of (U_i, U_j) one finds that the output state is

$$\rho_{\text{out}}^{ct} = \frac{1}{d^4}\sum_{i,j} |\Phi_{ij}\rangle\langle\Phi_{ij}|^{ct} = \frac{\mathbb{1}^c}{2} \otimes \frac{\mathbb{1}^t}{d} + \frac{1}{2}\left[|0\rangle\langle 1|^c + |1\rangle\langle 0|^c\right] \otimes T\rho_{\text{in}}^t T^\dagger \tag{4.3}$$

where $T := \frac{1}{d^2}\sum_i U_i$ and $\rho_{\text{in}}^t := |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t$. By linearity, Eq. (4.3) holds for arbitrary inputs ρ_{in}^t, and the setup thus induces the global channel $\mathcal{M}$ mapping $\rho_{\text{in}}^t \to \rho_{\text{out}}^{ct}$.

It is immediately clear that ρ_{out}^{ct} depends in general on ρ_{in}^t, and thus some information can be transmitted through the setup. If, on the other hand, one classically controls which channel is applied, no information can be transmitted. Indeed, if the initial state of the control qubit is diagonal, or if it decoheres, in the Pauli σ_z eigenbasis, then one can easily check that all dependence on ρ_{in}^t disappears in ρ_{out}^{ct}. Thus, the global channel $\mathcal{M}$ arising from coherently controlling between $\mathcal{N}_0$ and $\mathcal{N}_1$ provides a communication advantage over classical control. This mirrors that found using the quantum switch in Ref. [29], where it was seen to arise when applying $\mathcal{N}_0$ and $\mathcal{N}_1$ in an indefinite causal order. In the example above, however, there is no indefinite causal order and yet the effect remains, contradicting any possible intuition that it should be attributed to causal indefiniteness (a conclusion also recently reached in Ref. [79]).

It was further noted in Ref. [29] that if one traces out either the control or target system from the output of the depolarising quantum switch one obtains the completely mixed state, and thus information is transmitted solely in the correlations between the control and target states. In the present example, while it is still true that if the control is traced out the target system is left in the completely mixed state, if one traces out the target from Eq. (4.3) one obtains $\rho_{\text{out}}^c = \frac{1}{2}(\mathbb{1}^c + \text{Tr}[T\rho_{\text{in}}^t T^\dagger]\sigma_x^c)$, which still depends on ρ_{in}^t. Nevertheless, the control system itself never contains all of the information about the input target state that gets transmitted to ρ_{out}^{ct} (it only contains $\text{Tr}[T\rho_{\text{in}}^t T^\dagger]$, while ρ_{out}^{ct} contains $T\rho_{\text{in}}^t T^\dagger$). Note that in the setup of Fig. 4.1, just as in the depolarising quantum switch, the subspace on which each channel acts has nontrivial "spread" over both the target and control systems [38] and so the channels cannot strictly be said to act trivially on the control system. Indeed, in both cases the global channel entangles the two systems and gives a state containing some delocalised information about the target. This is

[2] Note that U_i and U_j (and the Kraus operators K_i, L_j considered later) must be written with respect to a common reference phase. We also assume that the arms of the circuit do not introduce any additional relative phases).

conceptually similar to the effect of quantum phase kickback associated with controlled unitary operations [30, 80], a connection further explored in Refs. [79, 81].

4.3 Dependence on channel implementation

The approach employed above of randomising over unitary channels is not, however, the only way to implement a fully depolarising channel. Recall that in general, a quantum channel $\mathcal{C}$ is defined as a CPTP map, and can be described in terms of a (non-unique) set of Kraus operators $\{K_i\}_i$ satisfying $\sum_i K_i^\dagger K_i = \mathbb{1}$, such that the output of the channel is given by $\mathcal{C}(\rho) = \sum_i K_i \rho K_i^\dagger$ for every density matrix ρ [30, 31, 68]. Note in particular that $\mathcal{C}$ may be applied to a subsystem in a subspace of some larger Hilbert space, as is the case both in the quantum switch and the scenario of Fig. 4.1. There, however, if the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ are not unitary—or not described, as previously considered, as a randomisation over unitary channels—it is *a priori* unclear how to determine the global channel mapping $\rho_{\text{in}}^t \to \rho_{\text{out}}^{ct}$ from the Kraus operators of $\mathcal{C}_0$ and $\mathcal{C}_1$.

One possible approach to doing so is to "purify" the channels via (independent) Stinespring dilations [82]. Any channel $\mathcal{C}$ with Kraus operators $\{K_i\}_i$ can indeed be extended to a unitary operation by introducing an environment in an initial state $|\varepsilon\rangle^e$ and considering the operation that acts on the system under consideration (in our case, the target) and the environment as $|\psi_{\text{in}}\rangle^t \otimes |\varepsilon\rangle^e \to \sum_i K_i |\psi_{\text{in}}\rangle^t \otimes |i\rangle^e := |\Phi_{\text{out}}\rangle^{te}$, where the ket vectors $|i\rangle^e$ are (normalised) orthogonal states of the environment. After tracing out the environment, we recover $\text{Tr}_e |\Phi_{\text{out}}\rangle\langle\Phi_{\text{out}}|^{te} = \sum_i K_i |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t K_i^\dagger = \mathcal{C}(|\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|^t)$, as required.

In the setup of Fig. 4.1 where the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ have Kraus operators $\{K_i\}_i$ and $\{L_j\}_j$, respectively, one may therefore purify the channels by introducing two, initially uncorrelated, environments with initial states $|\varepsilon_0\rangle^{e_0}$ and $|\varepsilon_1\rangle^{e_1}$. Note that the control qubit must then be seen as controlling the action of the purified unitary extensions of the channels not only on the target system, but also on the corresponding environments. This is nevertheless sensible in the interferometric picture of Fig. 4.1 where the channels may be seen as black boxes with "internal" environments that a photon traverses (in a superposition of "here" and "there").

Under these controlled, purified channels, the combined control-target-environments state evolves unitarily from $|+\rangle^c \otimes |\psi_{\text{in}}\rangle^t \otimes |\varepsilon_0\rangle^{e_0} \otimes |\varepsilon_1\rangle^{e_1}$ to

$$\frac{1}{\sqrt{2}} |0\rangle^c \otimes \sum_i K_i |\psi_{\text{in}}\rangle^t \otimes |i\rangle^{e_0} \otimes |\varepsilon_1\rangle^{e_1} + \frac{1}{\sqrt{2}} |1\rangle^c \otimes \sum_j L_j |\psi_{\text{in}}\rangle^t \otimes |\varepsilon_0\rangle^{e_0} \otimes |j\rangle^{e_1} . \tag{4.4}$$

After tracing out the environments, the resulting joint control-target state ρ_{out}^{ct} is found to be

$$\rho_{\text{out}}^{ct} = \frac{1}{2}[|0\rangle\langle 0|^c \otimes \mathcal{C}_0(\rho_{\text{in}}^t) + |1\rangle\langle 1|^c \otimes \mathcal{C}_1(\rho_{\text{in}}^t)] + \frac{1}{2}[|0\rangle\langle 1|^c \otimes T_0 \rho_{\text{in}}^t T_1^\dagger + |1\rangle\langle 0|^c \otimes T_1 \rho_{\text{in}}^t T_0^\dagger] \tag{4.5}$$

with $T_0 := \sum_i \langle\varepsilon_0|i\rangle K_i$ and $T_1 := \sum_j \langle\varepsilon_1|j\rangle L_j$.

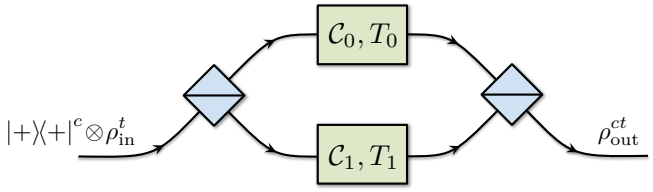

FIGURE 4.2: A corrected version of Fig. 4.1 in which the description of the two operations, implementing the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ on their respective subspaces, have been supplemented by the transformation matrices T_0 and T_1 needed to fully specify the output state ρ^{ct}_{out}.

The output state (4.3), obtained when $\mathcal{C}_0$ and $\mathcal{C}_1$ are depolarising channels implemented as a classical randomisation over d^2 orthogonal unitary operators U_i, is recovered by taking $K_i = \frac{1}{d}U_i$, $L_j = \frac{1}{d}U_j$, and the initial states of the environment to be $|\varepsilon_0\rangle^{e_0} = \sum_{i=0}^{d^2-1} \frac{1}{d} |i\rangle^{e_0}$, $|\varepsilon_1\rangle^{e_1} = \sum_{j=0}^{d^2-1} \frac{1}{d} |j\rangle^{e_1}$. Note, however, that a different choice of orthogonal unitary operations (even due to the addition of a relative phase between them, so that, taken individually, they would still implement the same local unitary channels) would have led to a different output state in Eq. (4.5). If we had instead taken the environments to initially be in the states $|0\rangle^{e_0}$ and $|0\rangle^{e_1}$ and chosen a set of orthogonal unitary operators such that $K_0 = L_0 = \frac{1}{d}\mathbb{1}$, we would have obtained Eq. (4.3) with $T = \frac{1}{d}\mathbb{1}$—which, incidentally, coincides with the state of Eq. (4.1) obtained in Ref. [29] as the output of the depolarising quantum switch. We nevertheless emphasise that Eq. (4.5) gives the output control-target state for any description of the channels in terms of Kraus operators.

The crucial observation here is that ρ^{ct}_{out} depends on the implementation of the channels $\mathcal{C}_0$ and $\mathcal{C}_1$ [83, 84]. The interferometric circuit in Fig. 4.1 is therefore not fully defined by the channels $\mathcal{C}_0$ and $\mathcal{C}_1$, or the Kraus operators chosen to represent them. This may appear surprising given that, in the usual paradigm, quantum channels are understood to be fully characterised by their (non-unique) Kraus representation [30, 31, 68]. However, one should note that such a description of a channel is unchanged under addition of any global phase. On the other hand, any such "global" phase applied by one of the channels in Fig. 4.1 is only applied to the corresponding arm of the interferometer and therefore, in the overall controlled circuit, becomes a "relative" phase with physical significance. In the case where $\mathcal{C}_0$ and $\mathcal{C}_1$ are unitary, the fact that Fig. 4.1 is only defined up to such a phase on the unitaries is well known [76, 80].

What we see here, however, is that the output of the interferometric circuit depends not only on any relative phases between (the Kraus operators of) the two channels, but also on a more detailed description of their implementations. More precisely, one requires some additional information encoded in the matrices T_0, T_1 introduced in Eq. (4.5) in order to fully specify the global channel $\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1] : \rho^t_{\text{in}} \to \rho^{ct}_{\text{out}}$ induced by the circuit; see Fig. 4.2. We call these the "transformation matrices" of the channel implementations. In the description above in terms of a Stinespring dilation, these depend not only on the set of Kraus operators used to decompose the channel, but also on how these are combined (with coefficients that depend on the environment states) to define $T_0 := \sum_i \langle\varepsilon_0|i\rangle K_i$ and $T_1 := \sum_j \langle\varepsilon_1|j\rangle L_j$.

We can in fact completely characterise the transformation matrices T obtainable from some realisation of any given channel $\mathcal{C}$, by deriving a general constraint expressed in terms of the Choi-Jamiolkowski representations of $\mathcal{C}$ and T. In the appendix of Ref. [36], we show that the set $\mathcal{T}_{\mathcal{C}}$ of all possible transformation matrices T of $\mathcal{C}$ is

$$\mathcal{T}_{\mathcal{C}} = \{T : |T\rangle\rangle \in \text{range}(C) \text{ and } \langle\langle T| C^{+} |T\rangle\rangle \leq 1\}. \tag{4.6}$$

where C is the Choi-Jamiolkowski matrix of $\mathcal{C}$, and C^{+} is its Moore-Penrose pseudoinverse.

Consider for example an identity channel $\mathcal{I}$, whose Choi-Jamiolkowski representation is $I = |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|$ (in any dimension); its range is the span of $|\mathbb{1}\rangle\rangle$ only, and its Moore-Penrose pseudoinverse is $I^{+} = \frac{|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|}{\langle\langle\mathbb{1}|\mathbb{1}\rangle\rangle^2} = \frac{|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|}{d^2}$. Eqs. (4.6) and (1.4) imply that

$$\mathcal{T}_{\mathcal{I}} = \{T = \alpha\mathbb{1} \,:\, \alpha \in \mathbb{C}, |\alpha| \leq 1\}. \tag{4.7}$$

Any such $T = \alpha\mathbb{1}$ with $|\alpha| \leq 1$ can indeed be obtained by taking for instance $\{K_i\}_i = \{K_0 = \mathbb{1}\}$ and $\langle\varepsilon_0|0\rangle = \alpha$. As one can see, even the identity channel does not define a unique transformation matrix. The freedom one has on its possible transformation matrices is not just due to a possible global phase (which would just restrict α above to $|\alpha| = 1$), but also to the possible coherent control of some operation $|\psi\rangle^t \otimes |\varepsilon\rangle^e \to |\psi\rangle^t \otimes |0\rangle^e$ that (while acting trivially on the target system) acts nontrivially on the environment. Note that Eq. (4.7) generalises straightforwardly to any unitary channel $\mathcal{U} : \rho \to U\rho U^{\dagger}$, whose possible transformation matrices are of the form $T = \alpha U$ with $|\alpha| \leq 1$.

For a d-dimensional fully depolarising channel $\mathcal{N}$, the Choi matrix is $N = \frac{1}{d}\mathbb{1}$; its range is the full Hilbert space $\mathcal{H}_I \otimes \mathcal{H}_O$, and its Moore-Penrose pseudoinverse is $N^{+} = d\,\mathbb{1}$. Noting that $\langle\langle T|T\rangle\rangle = \text{Tr}[T^{\dagger}T]$, Eq. (4.6) implies that

$$\mathcal{T}_{\mathcal{N}} = \Big\{T \,:\, \text{Tr}[T^{\dagger}T] \leq \frac{1}{d}\Big\}. \tag{4.8}$$

Any such T satisfying $\text{Tr}[T^{\dagger}T] \leq \frac{1}{d}$ can indeed be obtained by taking for instance the set of Kraus operators $\{\frac{1}{d}U_i\}_{i=0}^{d^2-1}$ (where the U_i's are again orthogonal unitary matrices) and $|\varepsilon\rangle^e$ such that $\langle\varepsilon_0|i\rangle = \langle\langle U_i|T\rangle\rangle = \text{Tr}[U_i^{\dagger}T]$.

Under the constraint (4.8), applied to both T_0 and T_1, Eq. (4.5) characterises all possible output states that one can obtain from the setup of Fig. 4.2, for any implementation of the channels

$\mathcal{C}_0, \mathcal{C}_1 = \mathcal{N}$.

4.4 Distinguishing different implementations of coherently-controlled channels

The dependence of the output of the circuit of Fig. 4.2 on the implementation of the channels means that it is also possible to differentiate between two distinct implementations of the same quantum channel with different transformation matrices.

Consider the case where $\mathcal{C}_0$ has a single, fixed implementation with a transformation matrix T_0, while $\mathcal{C}_1$ can have two different possible implementations, with $T_1 \neq T_1'$. The global channels $\mathcal{M}_{T_1} := \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1]$ and $\mathcal{M}_{T_1'} := \mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1']$ thus differ in general. If T_1 and T_1' are equally probable, then the maximal probability of successfully distinguishing these two channels—and thereby the two implementations of $\mathcal{C}_1$—is $\frac{1}{2}(1 + \mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}))$, where $\mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}) := \frac{1}{2}\|\mathcal{M}_{T_1} - \mathcal{M}_{T_1'}\|_\diamond$ is the diamond-norm distance between the two global channels [85]. In the relevant paper [34], we show that

$$\mathcal{D}(\mathcal{M}_{T_1}, \mathcal{M}_{T_1'}) \leq \frac{1}{2}\|T_1 - T_1'\|_2 \tag{4.9}$$

(where $\|\cdot\|_2$ is the spectral norm), and that this upper bound can be reached with $\mathcal{C}_0 = \mathcal{I}$ (the identity channel), $T_0 = \mathbb{1}$ by performing optimal state discrimination between the two output states of the global channel obtained for the input $\rho_{\text{in}}^t = |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|$ maximising $\langle\psi_{\text{in}}| (T_1 - T_1')^\dagger (T_1 - T_1') |\psi_{\text{in}}\rangle$.

It is indeed well-known, for instance, that the interferometric setup of Fig. 4.2 allows one to perfectly discriminate whether the lower arm applies the operation $|\psi\rangle^t \to |\psi\rangle^t$ or $|\psi\rangle^t \to -|\psi\rangle^t$ (the unitaries $\pm\mathbb{1}$), even though these both correspond to $\mathcal{C}_1 = \mathcal{I}$ (but with $T_1, T_1' = \pm\mathbb{1}$) on the relevant subspace. As another, perhaps more interesting, example consider the case where $\mathcal{C}_1 = \mathcal{N}$ is the fully depolarising channel, with the two possible transformation matrices $T_1^{(\prime)} = \pm\frac{1}{\sqrt{d}}|0\rangle\langle 0|$. We have $\frac{1}{2}\|T_1 - T_1'\|_2 = \frac{1}{\sqrt{d}}$, so that these two implementations of the depolarising channel can be distinguished with probability $\frac{1}{2}(1 + \frac{1}{\sqrt{d}})$ ($\simeq 0.85$ for $d = 2$).

4.5 Channel implementation independence for the full quantum switch and other quantum processes

In this section, we show explicitly that for the quantum switch, the transformation $\mathcal{S}[\mathcal{C}_0, \mathcal{C}_1]$ it induces has no dependence on the implementation of $\mathcal{C}_0$ and $\mathcal{C}_1$, and that this is more generally true in any setup in which each channel is always applied once and only once to the target system. In particular, it holds also for the quantum circuits with quantum control we considered in Chapter 3.

For the case of the full quantum switch, each of the two channels $\mathcal{C}_0$ and $\mathcal{C}_1$, with Kraus operators $\{K_i\}_i$ and $\{L_j\}_j$, is applied once and only once on the target system. Considering a purified version of the channels via a Stinespring dilation, as described in Sec. 4.3, the state at the output of the interferometer (see the inset of Fig. 4.1) reads

$$\frac{1}{\sqrt{2}}\left|0\right\rangle^c \otimes \sum_{i,j} L_j K_i \left|\psi_{\text{in}}\right\rangle^t \otimes \left|i\right\rangle^{e_0} \otimes \left|j\right\rangle^{e_1} + \frac{1}{\sqrt{2}}\left|1\right\rangle^c \otimes \sum_{i,j} K_i L_j \left|\psi_{\text{in}}\right\rangle^t \otimes \left|i\right\rangle^{e_0} \otimes \left|j\right\rangle^{e_1}. \tag{4.10}$$

In contrast to the output state (4.4) for the circuit of Fig. 4.1, no terms appear in which either environment is untouched and remains in its initial state. After tracing out the environments, one obtains

$$\begin{aligned}\rho^{ct}_{\text{out}} =& \frac{1}{2}\Big(\left|0\right\rangle\!\left\langle 0\right|^c \otimes \mathcal{C}_1 \circ \mathcal{C}_0(\rho^t_{\text{in}}) + \left|1\right\rangle\!\left\langle 1\right|^c \otimes \mathcal{C}_0 \circ \mathcal{C}_1(\rho^t_{\text{in}}) \\ &+ \left|0\right\rangle\!\left\langle 1\right|^c \otimes \sum_{i,j} L_j K_i \rho^t_{\text{in}} L_j^\dagger K_i^\dagger + \left|1\right\rangle\!\left\langle 0\right|^c \otimes \sum_{i,j} K_i L_j \rho^t_{\text{in}} K_i^\dagger L_j^\dagger \Big),\end{aligned} \tag{4.11}$$

which depends neither on the initial state of the environments, nor on the sets of Kraus operators chosen to describe each channel. Indeed, for any other Kraus representations $\{M_r\}_r$ of $\mathcal{C}_0$ and $\{N_s\}_s$ of $\mathcal{C}_1$, one has $K_i = \sum_r u_{ir} M_r$ and $L_j = \sum_s v_{js} N_s$, where u_{ir} and v_{js} are the elements of unitary matrices [30]. We thus obtain

$$\begin{aligned}\sum_{i,j} L_j K_i \rho^t_{\text{in}} L_j^\dagger K_i^\dagger &= \sum_{i,j} \sum_{r,r',s,s'} u_{ir} u^*_{ir'} v_{js} v^*_{js'} N_s M_r \rho^t_{\text{in}} N_{s'}^\dagger M_{r'}^\dagger \\ &= \sum_{r,r',s,s'} \delta_{r,r'} \delta_{s,s'} N_s M_r \rho^t_{\text{in}} N_{s'}^\dagger M_{r'}^\dagger = \sum_{r,s} N_s M_r \rho^t_{\text{in}} N_s^\dagger M_r^\dagger\end{aligned} \tag{4.12}$$

(where δ is the Kronecker delta), and analogously for the term $\sum_{i,j} K_i L_j \rho^t_{\text{in}} K_i^\dagger L_j^\dagger$.

More generally, consider a combination of multiple channels $\mathcal{C}_0, \ldots, \mathcal{C}_N$ with Kraus operators $\{K^{(0)}_{i_0}\}_{i_0}, \ldots, \{K^{(N)}_{i_N}\}_{i_N}$, and assume that for any possible initial state $\left|\Psi_{\text{in}}\right\rangle$ sent through the setup, each channel is applied once and only once (not necessarily in a definite order). Considering a Stinespring dilation of the channels with environment initial states $\left|\varepsilon_0\right\rangle^{e_0}, \ldots, \left|\varepsilon_N\right\rangle^{e_N}$, this means that the joint initial state evolves as

$$\left|\Psi_{\text{in}}\right\rangle \otimes \left|\varepsilon_0\right\rangle^{e_0} \otimes \cdots \otimes \left|\varepsilon_N\right\rangle^{e_N} \to \sum_{i_0,\ldots,i_N} F(K^{(0)}_{i_0}, \ldots, K^{(N)}_{i_N}) \left|\Psi_{\text{in}}\right\rangle \otimes \left|i_0\right\rangle^{e_0} \otimes \cdots \otimes \left|i_N\right\rangle^{e_N}, \tag{4.13}$$

where each $F(K^{(0)}_{i_0}, \ldots, K^{(N)}_{i_N})$ is an operator composed as a sum of product terms in which each $K^{(\ell)}_{i_\ell}$ appears once and only once, in possibly different orders (e.g., for the quantum switch: $\left|\Psi_{\text{in}}\right\rangle = \left|+\right\rangle^c \otimes \left|\psi_{\text{in}}\right\rangle^t$ and $F(K_i, L_j) = \left|0\right\rangle\!\left\langle 0\right|^c \otimes L_j K_i + \left|1\right\rangle\!\left\langle 1\right|^c \otimes K_i L_j$). For any such transformation, a similar calculation as for the full quantum switch can be conducted, which shows that after tracing out the environments, the final output state does not depend on the choice of Kraus operators, nor on the initial states of the environments.

As is straightforward to see, the global transformation induced by a quantum circuit with quantum control (cf. Eq. (3.58)) satisfies (4.13), and does thus not depend on the implementation.

4.6 Holevo information of the coherently controlled depolarising channels

In Ref. [29], the authors quantified precisely how much classical information can be transmitted by a single use of the depolarising quantum switch (i.e., its Holevo information [86, 87]). In the appendix of [36], we presented a lower bound for the Holevo information of the global channel $\mathcal{M}$ defined above. We found that significantly more information can be transmitted by this setup than with the full depolarising quantum switch.

The Holevo information of a channel $\mathcal{C}$ quantifies how much classical information can be transmitted through a single use of $\mathcal{C}$ from a party A to another party B. It is defined as $\chi(\mathcal{C}) := \max_{\{p_a,\rho_a\}} I(A;B)_\nu$, where $I(A;B)_\nu$ is the quantum mutual information calculated on the state $\nu := \sum_a p_a \left|a\right\rangle\!\left\langle a\right|_A \otimes \mathcal{C}(\rho_a)_B$ [86, 87] (i.e., $I(A;B)_\nu = H(A)_\nu + H(B)_\nu - H(AB)_\nu$, where $H(X)_\nu$ is the von Neumann entropy of the system $X \in \{A, B, AB\}$ in the state ν). $\chi(\mathcal{C})$ provides a lower bound for the classical capacity of a quantum channel $\mathcal{C}$.

In Ref. [29] it was shown that the Holevo information of the global channel $\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]$, mapping ρ^t_{in} to ρ^{ct}_{out} (see inset of Fig. 4.1), with $\rho^{ct}_{\text{out}} = \mathcal{S}[\mathcal{N}_0, \mathcal{N}_1](\rho^t_{\text{in}})$ given by Eq. (4.1), is $\chi(\mathcal{S}[\mathcal{N}_0, \mathcal{N}_1]) = -\frac{3}{8} - \frac{5}{8}\log_2\frac{5}{8} \simeq 0.05$ for a qubit target system, while a more general formula for any dimension d was also given.

For the global channel $\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1]$ induced by the situation of Fig. 4.2, which maps ρ^t_{in} to ρ^{ct}_{out} according to Eq. (4.5), a lower bound on the Holevo information $\chi(\mathcal{M}[\mathcal{C}_0, T_0, \mathcal{C}_1, T_1])$ can (for a given pair of channels and transformation matrices) be easily obtained by simply taking $I(A;B)_\nu$, for any particular choice of the weighted ensemble $\{p_a, \rho_a\}$. For two fully depolarising channels $\mathcal{N}_0, \mathcal{N}_1$ with $T_0 = T_1 = \frac{1}{\sqrt{d}}\left|0\right\rangle\!\left\langle 0\right|$ for instance (which indeed satisfies $\text{Tr}[T_i^\dagger T_i] \leq \frac{1}{d}$ as required by Eq. (4.8)), taking $\{p_0 = \frac{3}{5}, \rho_0 = \left|0\right\rangle\!\left\langle 0\right|, p_1 = \frac{2}{5}, \rho_1 = \left|1\right\rangle\!\left\langle 1\right|\}$ gives the lower bound $\chi(\mathcal{M}[\mathcal{N}_0, T_0, \mathcal{N}_1, T_1]) \geq \frac{1}{d}\log_2\frac{5}{4}$, which is a significant increase over that obtained by the depolarising quantum switch. For $d = 2$ this indeed gives $\chi(\mathcal{M}[\mathcal{N}_0, T_0, \mathcal{N}_1, T_1]) \gtrsim 0.16$; for larger d this bound decreases as the dimension increases, but less rapidly than the (exact) Holevo information obtained in Ref. [29]. It remains an open question whether our lower bound is tight, both for the transformation matrices T_i as well as for any other transformation matrices for two completely depolarising channels.

We note that the fact that the lower bound obtained here exceed the Holevo information for the depolarising quantum switch obtained in Ref. [29] is perhaps not so surprising given the differences between the scenario in Fig. 4.2 and that of the quantum switch. Indeed, in the scenario we consider, the target system only goes through the depolarising channels (in a superposition) a single time, while in the quantum switch the target system always goes through

both channels (in a superposition of different orders). Thus, one may intuitively expect the target system to be less "degraded" in the scenario considered here.

While we focused on depolarising channels to illustrate the ability for two coherently controlled zero-capacity channels to transmit information, this is not the only case where one should intuitively expect no communication to be possible. Indeed, any constant channel has zero capacity [86, 87] and similar conclusions can be drawn for any such channel. Furthermore, while this situation allows for the communication of *classical* information, we note that the authors of Ref. [29] also investigated the transmission of *quantum* information through a quantum switch that puts two dephasing channels in a superposition of orders [48] (see also Ref. [49]). In fact the advantage found there with the quantum switch is also present in the analogous scenario of Fig. 4.2; see [34].

4.7 Discussion

Coherent control of quantum channels was previously shown to be a resource for communicating through noisy channels in the technique of "error filtration" [88]. Our analysis, following that of Ref. [29], shows how it provides more general communication advantages, increasing the capacity of transmission in the absence of postselection and even in the extreme case of completely depolarising channels.

Our analysis illuminated the fact that the global transformation implemented by the circuit in Fig. 4.2 depends on the implementation of whatever channels are used. This stands in contrast to the usual paradigm of quantum channels, where a channel is defined as a CPTP map, and where all descriptions in terms of Kraus operators, or all purifications of a quantum channel, are equivalent [30, 31, 68]. Although such a description suffices if a channel is only ever used in isolation, by exploiting quantum control (something possible when the channel is supplied as a "black-box" or a usable communication channel [70]) it is in fact possible to extract information about how a channel is implemented, opening up the possibility to use coherent control as a tool for, e.g., error correction [89], quantum channel security [90] and characterisation [84].

Our results thus show that the notions of coherently controlling quantum channels—and, by extension, their actions when composed in circuits—is, by itself, ill-defined. Nevertheless, the setups in Figs. 4.1–4.2 that we have considered are perfectly realisable experimentally; indeed, they are less demanding than implementations of the full quantum switch [16–22]. Our observations here add to the call (e.g., in Ref. [70] for the control of unknown unitaries) for a generalisation of the standard paradigm of quantum circuits to describe experimentally conceivable situations, that would include the possibility for operations to be quantum-controlled (a general quantum "if statement"), or more generally to be applied on subspaces only. In the situation we considered, we saw that (generalised) quantum channels could not be defined only by the CPTP maps they induce, but also required one to specify the "transformation matrices" T introduced above. We

expect that this approach can be used for more general situations than the one investigated here, and leave its possible generalisation for future work.

Chapter 5

Genuinely multipartite noncausality

This chapter is based on the publication:
Alastair A. Abbott, Julian Wechs, Fabio Costa, and Cyril Branciard. Genuinely multipartite noncausality, *Quantum 1, 39 (2017)* (Ref. [35]).

It corresponds to Secs. 1 to 3.2 of the published paper, with some modifications to avoid repetitions and to fit into the broader context of the thesis. Some relevant technical details about the characterisation of the 2-causal polytope in the tripartite "lazy scenario" and the violations of 2-causal inequalities with process matrix correlations that were not given explicitly in the published paper have been added. My main contribution to this work was to perform the analysis of the tripartite "lazy scenario" described in Sec. 5.2, in parallel to the other coauthors in order to compare the results, and to write up Section 2 of the published manuscript [35].

In Section 1.4.2, we defined *causal correlations* as the most general multipartite correlations that are compatible with a definite causal order between the parties. Such correlations were studied and characterised in detail in Refs. [14, 15], and include those compatible with probabilistic or dynamical causal orders.

When analysing noncausal correlations in a multipartite setting, however, a natural question arises: is the noncausality of these correlations a truly multipartite phenomenon, or can it be reduced to a simpler one, that involves fewer parties? Indeed, there can be situations in which no overall causal order exists, but where there still is a ("coarse-grained") causal order between certain subsets of parties, obtained by grouping certain parties together. The correlations that can be established in such situations are more general than causal correlations, but nevertheless restricted due to the existence of this partial causal ordering. The goal of this chapter is precisely to address this question, and provide criteria to justify whether one really deals with *genuinely multipartite noncausality* or not.

To make things more precise, let us give an example. Consider three parties A and B and C. Each party receives an input x, y, z and returns an output a, b, c respectively. The correlations shared by A, B and C are described by the conditional probability distribution $P(a, b, c|x, y, z)$.

In that tripartite scenario, a causal correlation is defined as one that is compatible with one party acting first—which one it is may again be chosen probabilistically—and such that whatever happens with that first party, the reduced bipartite correlation shared by the other two parties, conditioned on the input and output of the first party, is causal (see Definition 1.3).

In contrast, a *noncausal* tripartite correlation $P(a, b, c|x, y, z)$ *cannot* for instance be decomposed as

$$P(a, b, c|x, y, z) = P(a|x)\, P_{x,a}(b, c|y, z) \tag{5.1}$$

with bipartite correlations $P_{x,a}(b, c|y, z)$ that are causal for each x, a. Nevertheless, such a decomposition may still be possible for a tripartite noncausal correlation if one does not demand that (all) the bipartite correlations $P_{x,a}(b, c|y, z)$ are causal. Without this constraint, the correlation (5.1) is thus compatible with the "coarse-grained" causal order $A \prec \{B, C\}$, if B and C are grouped together to define a new "effective party" and act "as one". This illustrates that although a multipartite correlation may be noncausal, there might still exist some definite causal order between certain subsets of parties; the intuition that motivates the work we presented in Ref. [35] is that such a correlation would therefore not display genuinely multipartite noncausality. If we want to identify the idea of noncausality as a *genuinely N-partite* phenomenon, we should exclude such correlations, and characterise correlations for which no subset of parties can have a definite causal relation to any other subset. This idea was already suggested in Ref. [14]; in Ref. [35] we defined the concept precisely.

In Sec. 5.1, we introduce the notion of *genuinely N-partite noncausal correlations* in opposition to what we call *2-causal correlations*, which can be established whenever two separate groups of parties can be causally ordered; we furthermore show how such correlations can be characterised via so-called *2-causal inequalities*. In Sec. 5.2, as an illustration we analyse in detail the simplest nontrivial tripartite scenario where these concepts make sense. We present explicit 2-causal inequalities for that scenario and investigate their violations in the process matrix framework.

5.1 Genuinely N-partite noncausal correlations

The general multipartite scenario that we consider in this chapter is the one we introduced in Section 1.4.1. Let us briefly recall some of the notations. A finite number $N \geq 1$ of parties A_k each receive an input x_k from some finite set (which can in principle be different for each party) and generate an output a_k that also belongs to some finite set (and which may also differ for each input). The vectors of inputs and outputs are denoted by $\vec{x} = (x_1, \ldots, x_N)$ and $\vec{a} = (a_1, \ldots, a_N)$. The correlations between the N parties are given by the conditional probability distribution $P(\vec{a}|\vec{x})$. For some (nonempty) subset $\mathcal{K} = \{k_1, \ldots, k_{|\mathcal{K}|}\}$ of $\mathcal{N} := \{1, \ldots, N\}$, we denote by $\vec{x}_{\mathcal{K}} = (x_{k_1}, \ldots, x_{k_{|\mathcal{K}|}})$ and $\vec{a}_{\mathcal{K}} = (a_{k_1}, \ldots, a_{k_{|\mathcal{K}|}})$ the vectors of inputs and outputs of the parties in $\mathcal{K}$; with this notation, $\vec{x}_{\mathcal{N}\setminus\mathcal{K}}$ and $\vec{a}_{\mathcal{N}\setminus\mathcal{K}}$ (or simply $\vec{x}_{\mathcal{N}\setminus k}$ and $\vec{a}_{\mathcal{N}\setminus k}$ for a singleton $\mathcal{K} = \{k\}$) denote the vectors of inputs and outputs of all parties that are not in $\mathcal{K}$. For simplicity

we will identify the parties' names with their labels, so that $\mathcal{N} = \{1, \ldots, N\} \equiv \{A_1, \ldots, A_N\}$, and similarly for any subset $\mathcal{K}$.

The assumption that the parties in such a scenario are embedded into a well-defined causal structure restricts the correlations that they can establish, and leads to our definition 1.3 of causal correlations—which, for clarity, we shall often call *fully causal* in this chapter.

Let us now characterise the correlations incompatible with an overall causal order, but still compatible with a ("coarse-grained") causal order between certain subsets. Note that if several different nonempty subsets do have definite causal relations to each other, then clearly there will be two subsets having a definite causal relation between them—one can consider the subset that comes first and group the remaining subsets together into the complementary subset, which then comes second. We shall for now consider partitions of $\mathcal{N}$ into just two (nonempty) subsets $\mathcal{K}$ and $\mathcal{N}\backslash\mathcal{K}$, and we thus introduce the following definition:

Definition 5.1 (2-causal correlations)**.** An N-partite correlation (for $N \geq 2$) is said to be 2-causal if and only if it can be decomposed in the form

$$P(\vec{a}|\vec{x}) = \sum_{\emptyset \subsetneq \mathcal{K} \subsetneq \mathcal{N}} q_{\mathcal{K}}\, P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}}) \tag{5.2}$$

where the sum runs over all nonempty strict subsets $\mathcal{K}$ of $\mathcal{N}$, with $q_{\mathcal{K}} \geq 0$ for each $\mathcal{K}$, $\sum_{\mathcal{K}} q_{\mathcal{K}} = 1$, and where (for each $\mathcal{K}$) $P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})$ is a valid probability distribution for the parties in $\mathcal{K}$ and (for each $\mathcal{K}, \vec{x}_{\mathcal{K}}, \vec{a}_{\mathcal{K}}$) $P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}})$ is a valid probability distribution for the remaining $N-|\mathcal{K}|$ parties.

For $N = 2$, the above definition reduces to the standard definition of bipartite causal correlations [4], which is equivalent to Definition 1.3 above. In the general multipartite case, it can be understood in the following way: each individual summand $P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}})$ for each bipartition $\{\mathcal{K}, \mathcal{N}\backslash\mathcal{K}\}$ describes correlations compatible with all the parties in $\mathcal{K}$ acting before all the parties in $\mathcal{N}\backslash\mathcal{K}$, since the choice of inputs for the parties in $\mathcal{N}\backslash\mathcal{K}$ does not affect the outputs for the parties in $\mathcal{K}$. The convex combination in Eq. (5.2) then takes into account the possibility that the subset $\mathcal{K}$ acting first can be chosen randomly.[1]

For correlations that are not 2-causal, we introduce the following terminology:

Definition 5.2 (Genuinely N-partite noncausal correlations)**.** An N-partite correlation that is not 2-causal is said to be genuinely N-partite noncausal.

[1] One can easily see that it is indeed sufficient to consider just one term per bipartition $\{\mathcal{K}, \mathcal{N}\backslash\mathcal{K}\}$ in the sum (5.2). That is, for some given $\mathcal{K}$, some correlations $P'(\vec{a}|\vec{x}) = P'_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P'_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}})$ and $P''(\vec{a}|\vec{x}) = P''_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P''_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}})$, and some weights $q', q'' \geq 0$ with $q' + q'' = 1$, the convex mixture $P(\vec{a}|\vec{x}) = q'P'(\vec{a}|\vec{x}) + q''P''(\vec{a}|\vec{x})$ is also of the same form $P(\vec{a}|\vec{x}) = P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\backslash\mathcal{K}}|\vec{x}_{\mathcal{N}\backslash\mathcal{K}})$ (with $P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}}) = q'P'_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}}) + q''P''_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})$ and $P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}} = P(\vec{a}|\vec{x})/P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})$). This already implies, in particular, that 2-causal correlations form a convex set.

Thus, genuinely N-partite noncausal correlations are those for which it is impossible to find any definite causal relation between any two (complementary) subsets of parties, even when taking into consideration the possibility that the subset acting first may be chosen probabilistically.

As shown in Ref. [12] for the bipartite case and in Refs. [14, 15] for the general N-partite case, any fully causal correlation can be written as a convex combination of deterministic fully causal correlations. As the number of such deterministic fully causal correlations is finite (for finite alphabets of inputs and outputs), they correspond to the extremal points of a convex polytope—the *(fully) causal polytope*. The facets of this polytope are given by linear inequalities, which define so-called *(fully) causal inequalities*.

As it turns out, the set of 2-causal correlations can be characterised as a convex polytope in the same way:

Theorem 5.3. *The set of 2-causal correlations forms a convex polytope, whose (finitely many) extremal points correspond to deterministic 2-causal correlations.*

Proof. For a given nonempty strict subset $\mathcal{K}$ of $\mathcal{N}$, $P_{\mathcal{K}}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\, P_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}})$ defines an "effectively bipartite" correlation, that is, a bipartite correlation between an effective party $\mathcal{K}$ with input $\vec{x}_{\mathcal{K}}$ and output $\vec{a}_{\mathcal{K}}$ and an effective party $\mathcal{N}\setminus\mathcal{K}$ with input $\vec{x}_{\mathcal{N}\setminus\mathcal{K}}$ and output $\vec{a}_{\mathcal{N}\setminus\mathcal{K}}$, which are formed by grouping together all parties in the respective subsets. That effectively bipartite correlation is compatible with the causal order[2] $\mathcal{K} \prec \mathcal{N}\setminus\mathcal{K}$. As mentioned above, the set of such correlations forms a convex polytope whose extremal points are deterministic, effectively bipartite causal correlations [12]—which, according to Definition 5.1, define deterministic 2-causal N-partite correlations.

Eq. (5.1) then implies that the set of 2-causal correlations is the convex hull of all such polytopes for each nonempty strict subset $\mathcal{K}$ of $\mathcal{N}$; it is thus itself a convex polytope, whose extremal points are indeed deterministic 2-causal correlations. □

As any fully causal correlation is 2-causal, but not vice versa, the fully causal polytope is a strict subset of what we shall call the *2-causal polytope* (see Fig. 5.1). Every vertex of the 2-causal polytope corresponds to a deterministic function $\vec{\alpha}$ that assigns a list of outputs $\vec{a} = \vec{\alpha}(\vec{x})$ to the list of inputs $\vec{x}$, such that the corresponding probability distribution $P^{\text{det}}_{\vec{\alpha}}(\vec{a}|\vec{x}) = \delta_{\vec{a},\vec{\alpha}(\vec{x})}$ is 2-causal, and thus satisfies Eq. (5.2). Since $P^{\text{det}}_{\vec{\alpha}}(\vec{a}|\vec{x})$ can only take values 0 or 1, there is only one term in the sum in Eq. (5.2), and it can be written such that there is a single (nonempty) strict subset $\mathcal{K}$ that acts first. That is, $\vec{\alpha}$ is such that the outputs $\vec{a}_{\mathcal{K}}$ of the parties in $\mathcal{K}$ are determined exclusively by their inputs $\vec{x}_{\mathcal{K}}$, while the outputs $\vec{a}_{\mathcal{N}\setminus\mathcal{K}}$ of the remaining parties are

[2] The notation $\mathcal{K}_1 \prec \mathcal{K}_2$ (or simply $A_{k_1} \prec A_{k_2}$ for singletons $\mathcal{K}_j = \{A_{k_j}\}$), formally means that the correlation under consideration satisfies $P(\vec{a}_{\mathcal{K}_1}|\vec{x}) = P(\vec{a}_{\mathcal{K}_1}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}_2})$. It will also be extended to more subsets, with $\mathcal{K}_1 \prec \mathcal{K}_2 \prec \cdots \prec \mathcal{K}_m$ meaning that $P(\vec{a}_{\mathcal{K}_1\cup\cdots\cup\mathcal{K}_j}|\vec{x}) = P(\vec{a}_{\mathcal{K}_1\cup\cdots\cup\mathcal{K}_j}|\vec{x}_{\mathcal{N}\setminus(\mathcal{K}_{j+1}\cup\cdots\cup\mathcal{K}_m)})$ for all $j = 1, \ldots, m-1$.

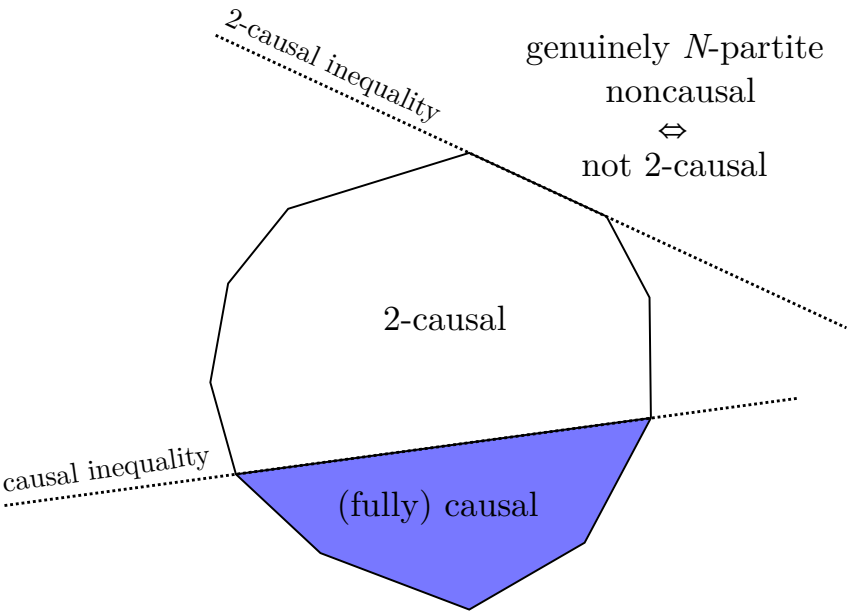

FIGURE 5.1: Sketch of the fully causal and 2-causal polytopes (the shading should be interpreted as indicating that the latter contains the former). The vertices of the polytopes correspond to deterministic fully causal and 2-causal correlations, and their facets correspond to causal and 2-causal inequalities, respectively. Correlations that are outside of the fully causal polytope are simply noncausal; correlations that are outside of the 2-causal polytope are genuinely N-partite noncausal.

determined by all inputs $\vec{x}$. The facets of the 2-causal polytope are linear inequalities that are satisfied by all 2-causal correlations; we shall call these *2-causal inequalities* (see Fig. 5.1).

5.2 Analysis of the tripartite "lazy scenario"

In this section we analyse in detail, as an illustration, the polytope of 2-causal correlations for the simplest nontrivial scenario with more than two parties. In Ref. [14] it was shown that this scenario is the so-called tripartite "lazy scenario", in which each party A_k receives a binary input x_k, has a single constant output for one of the inputs, and a binary output for the other. By convention we consider that for each k, on input $x_k = 0$ the output is always $a_k = 0$, while for $x_k = 1$ we take $a_k \in \{0, 1\}$. The set of fully causal correlations was completely characterised for this scenario in Ref. [14], which will furthermore permit us to compare the noncausal and genuinely tripartite noncausal correlations in this concrete example.

As is standard (and as we did in the introduction), we will denote here the three parties by A, B, C, their inputs x, y, z, and their outputs a, b and c. Furthermore, we will denote the complete tripartite probability distribution by P_{ABC} [i.e., $P_{ABC}(abc|xyz) := P(abc|xyz)$] and the marginal distributions for the indicated parties by P_{AB}, P_A, etc. [e.g., $P_{AB}(ab|xyz) = \sum_c P_{ABC}(abc|xyz)$].

5.2.1 Characterisation of the polytope of 2-causal correlations

We characterise the polytope of 2-causal correlations in much the same way as the polytope of fully causal correlations was characterised in Ref. [14].

As shown in Ref. [14], the minimum number of parameters needed to specify any probability $P_{ABC}(a,b,c|x,y,z)$ in the tripartite "lazy" scenario—and thus the dimension of the probability space under consideration—is $\sum_{n=0}^{3}\binom{3}{n}(2^n-1)=19$. (The sum runs over all possible combinations of ones and zeros for the inputs x, y, z. For n ones in the combination, we have 2^n possible outputs, and one of the corresponding probabilities is fixed because of normalisation.)

A suitable parametrisation of this probability space, also used in Ref. [14], is given by

$$\begin{aligned}\vec{P} = (&P_A(1|100), P_B(1|010), P_C(1|001),\\ &P_{AB}(10|110), P_{AB}(01|110), P_{AB}(11|110),\\ &P_{BC}(10|011), P_{BC}(01|011), P_{BC}(11|011),\\ &P_{AC}(01|101), P_{AC}(10|101), P_{AC}(11|101),\\ &P_{ABC}(100|111), P_{ABC}(010|111), P_{ABC}(001|111), P_{ABC}(110|111),\\ &\quad P_{ABC}(101|111), P_{ABC}(011|111), P_{ABC}(111|111))\,. \end{aligned} \tag{5.3}$$

To find the vertices of the polytope, we need to determine all deterministic 2-causal probability distributions P_{ABC}, i.e., those which admit a decomposition of the form (5.2) with (because they are deterministic) a single term in the sum, corresponding to a single group of parties acting first. The corresponding deterministic functions $\vec{\alpha}$ can be specified as a table

x	y	z	a	b	c
0	0	0	0	0	0
1	0	0	λ_1	0	0
0	1	0	0	λ_5	0
0	0	1	0	0	λ_9
1	1	0	λ_2	λ_6	0
0	1	1	0	λ_7	λ_{10}
1	0	1	λ_3	0	λ_{11}
1	1	1	λ_4	λ_8	λ_{12}

with parameters $\lambda_1,\cdots,\lambda_{12}\in\{0,1\}$, which satisfy additional constraints due to the compatibility with a causal order between some complementary subsets of $\{A,B,C\}$. For instance, if $\{A\}\prec\{B,C\}$, the output a of A is determined exclusively by the input x of A, and one therefore has $\lambda_1=\lambda_2=\lambda_3=\lambda_4$. If $\{B,C\}\prec\{A\}$, the outputs of B and C depend only on the inputs of B and C, and therefore one has $\lambda_5=\lambda_6$, $\lambda_9=\lambda_{11}$, $\lambda_7=\lambda_8$ and $\lambda_{10}=\lambda_{12}$. Similar constraints arise for the other partitions of $\{A,B,C\}$ into two complementary subsets.

It is straightforward to list all such deterministic functions that satisfy these constraints, and one finds that there are 1 520 of them. The corresponding correlations, or vertices of the polytope,

are then, in the parametrisation we chose above,

$$\vec{P}_{\vec{\alpha}}^{\,\mathrm{det}} = \Big(\lambda_1, \lambda_5, \lambda_9, \lambda_2(1-\lambda_6), (1-\lambda_2)\lambda_6, \lambda_2\lambda_6, \lambda_7(1-\lambda_{10}), (1-\lambda_7)\lambda_{10}, \lambda_7\lambda_{10},$$
$$(1-\lambda_3)\lambda_{11}, \lambda_3(1-\lambda_{11}), \lambda_3\lambda_{11}, \lambda_4(1-\lambda_8)(1-\lambda_{12}), (1-\lambda_4)\lambda_8(1-\lambda_{12}),$$
$$(1-\lambda_4)(1-\lambda_8)\lambda_{12}, \lambda_4\lambda_8(1-\lambda_{12}), (1-\lambda_4)\lambda_8\lambda_{12}, \lambda_4(1-\lambda_8)\lambda_{12}, \lambda_4\lambda_8\lambda_{12}\Big) \tag{5.4}$$

With the full set of vertices at hand, we can now determine the facets of the polytope, which in turn correspond to tight 2-causal inequalities, by solving the convex hull problem. For that purpose, the software CDD [91] can be used. We find that the polytope has 21 154 facets, each corresponding to a 2-causal inequality, the violation of which would certify genuinely tripartite noncausality. Many inequalities, however, can be obtained from others by either relabelling outputs or permuting parties, and as a result it is natural to group the inequalities into equivalence classes, or "families", of inequalities. Taking this into account, we find that there are 476 families of facet-inducing 2-causal inequalities, 3 of which are trivial, as they simply correspond to positivity constraints on the probabilities (and are thus satisfied by any valid probability distribution). While the 2-causal inequalities all detect genuinely N-partite noncausality, it is interesting to note that all except 22 of them can be saturated by fully causal correlations (and all but 37 even by correlations compatible with a fixed causal order).

We provide the complete list of these inequalities, organised by their symmetries and the types of distribution required to saturate them, in the supplementary material of the published paper [92], and will analyse in more detail a few particularly interesting examples in what follows. First, however, it is interesting to note that only 2 of the 473 nontrivial facets are also facets of the (fully) causal polytope for this scenario (one of which is Eq. (5.9) analysed below), and hence the vast majority of facet-inducing inequalities of the causal polytope do not single out genuinely tripartite noncausal correlations. Moreover, none of the 2-causal inequalities we obtain here differ from facet-inducing fully causal inequalities only in their bound, and, except for the aforementioned cases, our 2-causal inequalities thus represent novel inequalities.

Of the nontrivial 2-causal inequalities, those that display certain symmetries between the parties are particularly interesting since they tend to have comparatively simple forms and often permit natural interpretations (e.g., as causal games [4, 12]).

For example, three nontrivial families of 2-causal inequalities have forms (i.e., certain versions of the inequality within the corresponding equivalence class) that are completely symmetric under permutations of the parties. One of these is the inequality

$$\begin{aligned} I_1 = {} & [P_A(1|100) + P_B(1|010) + P_C(1|001)] \\ & + [P_{AB}(11|110) + P_{BC}(11|011) + P_{AC}(11|101)] \\ & - P_{ABC}(111|111) \geq 0, \end{aligned} \tag{5.5}$$

which can be naturally expressed as a causal game. Indeed, it can be rewritten as

$$P(\tilde{a}\tilde{b}\tilde{c} = xyz) \;\leq\; 3/4\,, \tag{5.6}$$

where $\tilde{a} = 1$ if $x = 0$, $\tilde{a} = a$ if $x = 1$ (i.e., $\tilde{a} = xa \oplus x \oplus 1$, where $\oplus$ denotes addition modulo 2), and similarly for $\tilde{b}$ and $\tilde{c}$, and where it is implicitly assumed that all inputs occur with the same probability. This can be interpreted as a game in which the goal is to collaborate such that the product of the nontrivial outputs (i.e., those corresponding to an input 1) is equal to the product of the inputs, and where the former product is taken to be 1 if all inputs are 0 and there are therefore no nontrivial outputs (in which case the game will always be lost). The probability of success for this game can be no greater than 3/4 if the parties share a 2-causal correlation. This bound can easily be saturated by a deterministic, even fully causal, distribution: if every party always outputs 0 then the parties will win the game in all cases, except when the inputs are all 0 or all 1.

Another party-permutation-symmetric 2-causal inequality is the following:

$$\begin{aligned} I_2 = 1 + 2\big[&P_A(1|100) + P_B(1|010) + P_C(1|001)\big] \\ -\big[&P_{AB}(11|110) + P_{BC}(11|011) + P_{AC}(11|101)\big] \geq 0, \end{aligned} \tag{5.7}$$

whose interpretation can be made clearer by rewriting it as

$$\begin{aligned} &P_A(1|100) + P_B(1|010) - P_{AB}(11|110) \\ &+ P_B(1|010) + P_C(1|001) - P_{BC}(11|011) \\ &+ P_A(1|100) + P_C(1|001) - P_{AC}(11|101) \geq -1. \end{aligned} \tag{5.8}$$

The left-hand side of this inequality is simply the sum of three terms corresponding to conditional "lazy guess your neighbour's input" (LGYNI) inequalities [12, 14], one for each pair of parties (of the form as in Eq. (1.36) in Sec. 1.4.2). That is, each of the three terms corresponds to a causal game where two parties need to guess each other's input when their respective input is 1, conditioned on the remaining party having input 0. The negative bound on the right-hand side accounts for the fact that any pair of parties that are grouped together in a bipartition may maximally violate the LGYNI inequality between them (and thus reach the minimum algebraic bound -1). This inequality can be interpreted as a "scored game" (as opposed to a "win-or-lose game") in which each pair of parties scores one point if they win their respective bipartite LGYNI game and the third party's input is 0, and where the goal of the game is to maximise the total score, given by the sum of all three pairs' individual scores. The best average score (when the inputs are uniformly distributed) for a 2-causal correlation is 5/4, corresponding to the 2-causal bounds of 0 in Eq. (5.7) and -1 in Eq. (5.8).[3] It is also clear from the form of Eq. (5.8) that for fully causal correlations the left-hand side is lower-bounded by 0. This inequality is thus

[3]The bound of these inequalities, and the best average score of the corresponding game, can be reached by a 2-causal strategy in which one party, say A, has a fixed causal order with respect to the other two parties grouped together, who share a correlation maximally violating the corresponding LGYNI inequality. For example, the

amongst the 22 facet-inducing 2-causal inequalities that cannot be saturated by fully causal distributions.

In addition to the inequalities that are symmetric under any permutation of the parties, there are four further nontrivial families containing 2-causal inequalities which are symmetric under cyclic exchanges of parties. One interesting such example is the following:

$$I_3 = 2 + [P_A(1|100) + P_B(1|010) + P_C(1|001)] - [P_A(1|101) + P_B(1|110) + P_C(1|011)] \geq 0. \tag{5.9}$$

This inequality can again be interpreted as a causal game in the form (where we again implicitly assume a uniform distribution of inputs for all parties)

$$P\big(x(y \oplus 1)(a \oplus z) = y(z \oplus 1)(b \oplus x) = z(x \oplus 1)(c \oplus y) = 0\big) \;\leq\; 7/8\,, \tag{5.10}$$

where the goal of the game is for each party, whenever they receive the input 1 and their right-hand neighbour has the input 0, to output the input of their left-hand neighbour (with C being considered, in a circular manner, to be to the left of A).[4] This inequality is of additional interest as it is one of the two nontrivial inequalities which is also a facet of the standard causal polytope for this scenario. (The second such inequality, which lacks the symmetry of this one, is presented in the supplementary material of the published paper [92].)

In the published paper [35], we also provide N-partite generalisations of some of the 2-causal inequalities presented here.

5.2.2 Violations of 2-causal inequalities by process matrix correlations

One of the major sources of interest in causal inequalities is the potential to violate them with process matrix correlations. In Ref. [14] it was shown that all the nontrivial fully causal inequalities for the tripartite lazy scenario can be violated by process matrices. However, for most inequalities violation was found to be possible using process matrices $W^{\{A,B\}\prec C}$ that are compatible with C acting last, which means the correlations they produced were necessarily 2-causal. It is therefore interesting to see whether process matrices are capable of violating 2-causal inequalities in general, and thus of exhibiting genuinely N-partite noncausality.

To look for violations of the 2-causal inequalities, we follow the same approach as in Refs. [12, 14]. Specifically, we write the 2-causal inequality that we want to test in the form $I(P(a,b,c|x,y,z)) \geq 0$, and we search for the quantum instruments $\{M^{A_{IO}}_{a|x}\}$, $\{M^{B_{IO}}_{b|y}\}$, $\{M^{C_{IO}}_{c|xz}\}$ and the (valid) process matrix $W \in \mathcal{W} \subset \mathcal{L}(\mathcal{H}^{A_{IO}} \otimes \mathcal{H}^{B_{IO}} \otimes \mathcal{H}^{C_{IO}})$ that minimise $I(P(a,b,c|x,y,z))$, where the

distribution $P(abc|xyz) = \delta_{a,0}\,\delta_{b,yz}\,\delta_{c,yz}$, where δ is the Kronecker delta function, is compatible with the order $A \prec \{B,C\}$ (or with $\{B,C\} \prec A$) and saturates Eqs. (5.7) and (5.8).

[4] The bound of 7/8 on the probability of success can, for instance, be reached by the fully causal (and hence 2-causal) distribution $P(abc|xyz) = \delta_{a,x}\,\delta_{b,xy}\,\delta_{c,yz}$, compatible with the order $A \prec B \prec C$, which wins the game in all cases except when $(x,y,z) = (1,0,0)$.

correlations $P(a,b,c|x,y,z)$ are calculated via the generalised Born rule (1.31), which in this tripartite case, reads

$$P(a,b,c|x,y,z) = \mathrm{Tr}\left[(M^{A_{IO}}_{a|x})^T \otimes (M^{B_{IO}}_{b|y})^T \otimes (M^{C_{IO}}_{c|z})^T \cdot W\right]. \tag{5.11}$$

For a value $I(P(a,b,c|x,y,z)) < 0$, we have found a violation of the inequality.

For some given, fixed instruments, finding the process matrix that minimises $I(P(a,b,c|x,y,z))$ is a semidefinite optimisation problem. Likewise, when the instruments of all but one parties, as well as the process matrix, are fixed, the instruments of the remaining party that minimise $I(P(a,b,c|x,y,z))$ can be determined by semidefinite optimisation. We therefore perform an interative "see-saw" algorithm, as in Refs. [12, 14], where we start with some random initial instruments, and then optimise the process matrix and the instruments of each party in turn. We continue this iterative procedure until the algorithm converges. The minimum thus obtained is not guaranteed to be a global one, but one can repeat the algorithm many times with different initial random instruments and thus obtain a bound on the optimal violation.

Here, we were able to find process matrices violating all but 2 of the 473 nontrivial families of tight 2-causal inequalities (including Eqs. (5.5) and (5.9) above) using qubits, and in all cases where a violation was found, the best violation was given by the instruments

$$\begin{gathered}\{M_{0|0} = |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|\} \\ \{M_{0|1} = |0\rangle\langle 0| \otimes |1\rangle\langle 1|\,, M_{1|1} = |1\rangle\langle 1| \otimes |0\rangle\langle 0|\}\end{gathered} \tag{5.12}$$

for all three parties (that also provided similar results in Ref. [14]). We similarly found that 284 families of these 2-causal inequalities (including Eq. (5.9)) could be violated by completely "classical" process matrices[5] (i.e., process matrices that are diagonal in the computational basis), a phenomenon that is not present in the bipartite scenario where classical processes are necessarily causal [4].

While the violation of 2-causal inequalities is again rather ubiquitous, the existence of two inequalities for which we found no violation is curious. One of these inequalities is precisely Eq. (5.7), and its decomposition in Eq. (5.8) into three LGYNI inequalities helps provide an explanation. In particular, the seemingly best possible violation of a (conditional) LGYNI inequality using qubits is approximately 0.2776 [12, 14], whereas it is clear that a process matrix violating Eq. (5.8) must necessarily violate a conditional LGYNI inequality between one pair of parties by at least 1/3. Moreover, in Ref. [12] it was reported that no better violation was found using three- or four-dimensional systems, indicating that Eq. (5.8) can similarly not be violated by such systems. It nonetheless remains unproven whether such a violation is indeed impossible, and the convex optimisation problem for three parties quickly becomes intractable for higher dimensional systems, making further numerical investigation difficult. The second inequality

[5] Incidentally, exactly the same number of families of fully causal inequalities were found to be violable with classical process matrices in Ref. [14]. It remains unclear whether this is merely a coincidence or the result of a deeper connection.

for which no violation was found can similarly be expressed as a sum of three different forms (i.e., relabellings) of a conditional LGYNI inequality, and a similar argument thus explains why no violation was found. Recall that, as they can be expressed as a sum of three conditional LGYNI inequalities with a negative 2-causal bound, these two 2-causal inequalities cannot be saturated by fully causal distributions; it is interesting that the remaining inequalities that require noncausal but 2-causal distributions to saturate can nonetheless be violated by process matrix correlations.

5.3 Discussion

The question whether some concept is genuinely multipartite arises naturally in scenarios involving multiple parties. In Ref. [35], we aimed to clarify when noncausal correlations can be considered to be a genuinely multipartite resource. In addressing this task, we first proposed a criterion to decide whether a given correlation shared by N parties is "genuinely N-partite noncausal"—i.e., its noncausality is indeed a genuinely N-partite resource—or not.

In particular, we introduced "2-causal" correlations, which are the most general correlations that are not genuinely N-partite noncausal. We showed that 2-causal correlations form a polytope, whose vertices are deterministic 2-causal correlations and whose facets define 2-causal inequalities. We completely characterised the 2-causal polytope for the simplest nontrivial tripartite scenario and found that almost all of the 473 nontrivial classes of 2-causal inequalities can be violated by process matrix correlations. However, we were unable to find any violation for 2 of those inequalities; this stands in contrast to previous studies of causal inequalities, where violations with process matrices were always found[6] [4, 12, 14, 47, 94]. Although it remains to be confirmed whether this is simply a failure of the search method we used, we provided some intuition why such a violation would in fact be a surprise.

In this chapter we only discussed correlations that can or cannot arise given a definite causal order between two subsets of parties. It makes sense to consider more refined definitions that discriminate, among noncausal correlations, to what extent and in which way they represent a genuinely multipartite resource. In Ref. [35], we addressed the question of how to refine the definition of genuinely multipartite noncausal correlations. We proposed two possible generalisations of the notion of 2-causal correlations, which we called *M-causal* and *size-S-causal correlations*, respectively. The first criterion is based on the largest number M of subsets that can be causally ordered while reproducing some given correlation, while the second criterion instead looks at how large the subsets that can be causally ordered are. We refer to Ref. [35] for an in-depth discussion.

[6] At least for standard causal inequalities that bound probabilities directly; for *entropic* causal inequalities, which only provide a relaxed characterisation of the set of causal correlations, no violations were found so far [93]. It would nevertheless also be interesting to investigate how genuinely multipartite noncausality can be characterised with the entropic approach.

Finally, in this chapter we only considered correlations from a fully theory- and device-independent perspective. As already mentioned at the end of Chapter 2, it would be interesting to develop similar notions within specific physical theories like the process matrix framework, and develop a notion of genuinely multipartite causal nonseparability.

Chapter 6

Anomalous weak values without post-selection

This chapter is based on the publication:
Alastair A. Abbott, Ralph Silva, Julian Wechs, Nicolas Brunner, and Cyril Branciard. Anomalous Weak Values Without Post-Selection, *Quantum 3, 194 (2019).* (Ref. [36])

The illustrative example in Sec. 6.3 is taken from Sec. 2 of the published paper. Sec. 6.4 corresponds to Sec. 3 of the published paper, with some of the calculations from the Appendix of Ref. [36] reproduced in a somewhat simplified form in the text here. My main contribution to this paper was to find a circuit implementation (cf. Chapter 3) for an initial example of a bipartite causally ordered process matrix generating an anomalous weak value. Based on this circuit implementation, we then established the simpler illustrative example presented here, and the main result that anomalous weak values without post-selection can arise in sequential weak measurements.

The main tool that we use in this thesis to study the causal structure between quantum events is the process matrix formalism. There are, however, various other approaches that have been developed to study questions revolving around the role of time and causality in quantum theory. One such framework, aimed at formulating quantum theory in a more time-symmetric way, is that of multiple-time states and -measurements [5]. In this approach, one considers situations consisting of multiple preparation and measurement stages, and one describes them via so-called *multiple-time states*, which are a new type of quantum state adapted to that kind of scenarios—i.e, an object associated with different times that contains all the relevant information we have about the physical system. In particular, one studies two-time states, which describe pre- and post-selected quantum systems that are prepared in some given initial state, and that also end up in some given final state at a later time.

In the context of such pre- and post-selected quantum systems, one often studies weak measurements. In order to weakly measure an observable on a quantum system, one lets it interact

weakly with another quantum system, the "pointer", which is then subjected to a standard, strong measurement. As a result of the weak coupling, the inevitable perturbation of the system due to the measurement is small, but only little information about the system is obtained. One can then associate a quantity called *weak value* to the measured observable, which, broadly speaking, is some analogue of the expectation value occuring specifically in weak measurements, and whose real part is the expectation value of the pointer position. Interestingly, the weak value can lie outside of the observable's spectrum, a phenomenon that is usually referred to as an "anomalous weak value".

In this chapter, we present a result that was obtained when studying questions concerning the link between the process matrix formalism and two-time states. It is generally believed that anomalous weak values are possible only when a non-trivial post-selection is performed. Here we show, however, that this is not the case in general, and that in sequential weak measurements (i.e., subsequent couplings of observables to different pointers, which are then jointly measured), an anomalous weak value can be obtained without post-selection, i.e., without discarding any data.

After giving some background information about pre- and post-selected quantum systems, as well as weak measurements, we present a simple example of that effect. We consider a quantum system that is subjected to a sequential weak measurement of two projection observables, and show that the expectation value of the product of the two pointer positions can become negative. We then show that this result can be associated to an anomalous weak value obtained without post-selection, and analyse the phenomenon more generally for arbitrary observables.

6.1 The two-time state formalism

First, let us briefly introduce pre-and post-selected quantum systems, and their description in terms of two-time states.

The most basic scenario described by a two-time state is the following: A quantum system is prepared (or "pre-selected") in a state $|\psi\rangle$ at some time t_1. In the time interval between t_1 and some final time t_2, Alice can perform arbitrary quantum experiments on the system. Finally, at t_2, an observable $\hat{Q}$ is measured, and the system is subjected to a post-selection onto one of the measurement outcomes q, which is the eigenvalue (assumed to be non-degenerate) corresponding to some eigenstate $|\phi\rangle$ of $\hat{Q}$. That is, the experiment is discarded whenever the measurement of $\hat{Q}$ does not yield q, and kept only if the outcome q occurs and the system ends up in the desired final state $|\phi\rangle$.

In such a situation, all the information (in addition to the operation of Alice) that we need to describe the experiment performed by Alice between t_1 and t_2 is encoded in the initial and final states $|\psi\rangle$ and $|\phi\rangle$. For instance, suppose that Alice acts on the system with a unitary U_1, and then performs a projective measurement with projectors $\{P_a\}_a$, followed by another unitary U_2.

The probability for Alice to obtain a specific measurement outcome a is then given by

$$P(a) = \frac{|\langle\phi| U_2 P_a U_1 |\psi\rangle|^2}{\sum_{a'} |\langle\phi| U_2 P_{a'} U_1 |\psi\rangle|^2}, \tag{6.1}$$

That is, the probability of obtaining a, *given that* the measurement of $\hat{Q}$ yields q, is the probability of obtaining a and then q, divided by the overall probability of obtaining q.

To capture all the information contained in the pre- and post-selected state in a convenient way, one can now introduce a new mathematical object called a *two-time state*, which consists of both a "forward-evolving" part $|\psi\rangle_{t_1}$, living in some space of ket vectors $\mathcal{H}^{\uparrow}_{t_1}$, and a "backward-evolving" part ${}_{t_2}\langle\phi|$, living in a space of bra vectors $\mathcal{H}^{\downarrow}_{t_2}$. It is defined as [5]

$$\Psi = {}_{t_2}\langle\phi| \bullet |\psi\rangle_{t_1} \in \mathcal{H} := \mathcal{H}^{\downarrow}_{t_2} \otimes \mathcal{H}^{\uparrow}_{t_1}. \tag{6.2}$$

The dot $\bullet$ denotes an open slot, into which Alice's operation will be inserted.

The basic situation considered here gives rise to a *pure product* two-time state. One can now add several layers of complexity, and consider superpositions of such two-time states, which are generally entangled between the forward-evolving and backward-evolving states, as well as mixed two-time states [6] or multipartite ones, in which, for instance, both Alice and Bob can perform experiments between the pre- and post-selection.

The link between two-time states and process matrices was established in Ref. [7] (and also mentioned in Ref. [8]) where it was shown that process matrices are equivalent to a particular subset of two-time states. To each process matrix, one can associate a two-time state that generates the same probabilities for any operations performed by Alice and Bob, and that has the particular property that the probabilities are *linear* in the operations. This is not true for all two-time states—in general, a denominator that depends on the measurements chosen by Alice and Bob appears in the probability rule and makes it nonlinear (similarly to the one in Eq. (6.1)).

Vice-versa, for any linear two-time state, one can find a corresponding process matrix.

6.2 Weak values and measurements

After this overview of pre- and post-selected quantum systems, we will now introduce the concept of weak measurements. We will then show that, when the two notions are combined, particularly interesting phenomena can occur.

Quantum systems are modified when they are subjected to a measurement, and there is a trade-off between information-gain and disturbance of the measured system. The general idea of a *weak measurement* is to study the regime where the system is perturbed as weakly as possible, while only little information about the system is obtained.

Weak measurements are studied within the von Neumann scheme of quantum measurements. In this model, quantum measurements consist of two steps. First, the system to be measured is coupled via a joint unitary operation to an ancillary quantum system, the "pointer", which represents the measurement device. Second, a strong measurement of the pointer is performed.

More specifically, say that we want to measure an observable $\hat{A}$ with eigenbasis $\{|a_k\rangle\}_k$ and the corresponding eigenvalues a_k, and the system to be measured is in a state $|\psi\rangle = \sum_k c_k |a_k\rangle$. We typically take the pointer to be initially in a state $|\varphi(0)\rangle$, where

$$|\varphi(a)\rangle \coloneqq \int_{-\infty}^{\infty} \mathrm{d}x \, \left(\frac{1}{2\pi\sigma^2}\right)^{1/4} \exp\left[-\frac{(x-a)^2}{4\sigma^2}\right] |x\rangle \,, \tag{6.3}$$

with $\{|x\rangle\}_x$ a continuous eigenbasis of the pointer position $\hat{x}$. That is, the initial state of the pointer is a Gaussian wave packet centred at a position $x = 0$ with spread (i.e., standard deviation) σ.

The standard choice for the unitary that couples the two systems during the "pre-measurement" step is $U = e^{-i\hat{H}\Delta t}$, where the Hamiltonian $\hat{H}$ is of the form $\hat{H} = \gamma \hat{A} \otimes \hat{p}$, with $\hat{p}$ denoting the momentum operator acting on the pointer, and γ a coupling constant [95].

The resulting state of the system and pointer after the pre-measurement reads

$$e^{-ig\hat{A}\hat{p}} |\psi\rangle |\varphi(0)\rangle = \sum_k \left(|a_k\rangle\langle a_k| \otimes e^{-iga_k\hat{p}}\right) |\psi\rangle |\varphi(0)\rangle = \sum_k c_k |a_k\rangle \otimes |\varphi(ga_k)\rangle \tag{6.4}$$

where tensor products are implicit, $\hbar = 1$ and $g = \gamma\Delta t$, and where we used that an operator of the form $e^{-i\alpha\hat{p}}$ acts as a displacement operator onto the pointer state, such that $e^{-i\alpha\hat{p}} |\varphi(a)\rangle = |\varphi(a+\alpha)\rangle$.

Whether the measurement is strong or weak depends on the overlap of the pointer states in Eq. (6.4), and therefore on the parameters g and σ. If g is sufficiently small compared to the spread of the pointer σ, the pointer states corresponding to the different eigenvalues of $\hat{A}$ have a large overlap, which means that little information can be extracted from a single measurement—however, the perturbation of the measured system due to the measurement is small.

Weak measurements are often studied in conjunction with pre- and post-selected quantum systems, as introduced in Section 6.1, since particularly interesting effects can arise in that case [96]. Suppose that we want to perform a weak measurement in the two-time state scenario described above, with a pre-selected state $|\psi\rangle$ and a post-selection onto the state $|\phi\rangle$ of the system. Assuming that we are indeed in the weak measurement regime, we can approximate the global state after the coupling (Eq. (6.4)) by

$$e^{-ig\hat{A}\hat{p}} |\psi\rangle |\varphi(0)\rangle \approx (\mathbb{1} - ig\hat{A}\hat{p}) |\psi\rangle |\varphi(0)\rangle \tag{6.5}$$

For simplicity we will henceforth choose units so that $g = 1$, as only the ratio between the pointer spread and g is important. The strength of the measurement will then be controlled solely by the pointer spread σ, and the validity of the weak regime will depend only on this being sufficiently large.[1]

Next, the system is post-selected onto the state $|\phi\rangle$. The final state of the pointer is then (up to normalisation)

$$\begin{aligned}\langle\phi|\,(\mathbb{1} - i\hat{A}\hat{p})\,|\psi\rangle\,|\varphi(0)\rangle &= \langle\phi|\psi\rangle\,(\mathbb{1} - iA^{\phi}_{\psi}\,\hat{p})\,|\varphi(0)\rangle \\ &\approx \langle\phi|\psi\rangle\; e^{-iA^{\phi}_{\psi}\,\hat{p}}\,|\varphi(0)\rangle\,, \end{aligned} \tag{6.6}$$

where

$$A^{\phi}_{\psi} := \frac{\langle\phi|\hat{A}|\psi\rangle}{\langle\phi|\psi\rangle} \tag{6.7}$$

is the so-called *weak value* of the observable $\hat{A}$ given the pre-selection in the state $|\psi\rangle$ and post-selection in the state $|\phi\rangle$ [96].

The mean position of the pointer is thus displaced—via the displacement operator $e^{-iA^{\phi}_{\psi}\,\hat{p}}$, which generates the (possibly unnormalised[2]) state $|\varphi(A^{\phi}_{\psi})\rangle = e^{-iA^{\phi}_{\psi}\,\hat{p}}\,|\varphi(0)\rangle$—to

$$\langle\hat{x}\rangle \approx \frac{\langle\varphi(A^{\phi}_{\psi})|\,\hat{x}\,|\varphi(A^{\phi}_{\psi})\rangle}{\langle\varphi(A^{\phi}_{\psi})|\varphi(A^{\phi}_{\psi})\rangle} = \mathrm{Re}(A^{\phi}_{\psi}). \tag{6.8}$$

Notably, the real part of the weak value can become very large when the pre- and post-selected states are almost orthogonal, i.e. $|\langle\phi|\psi\rangle| \ll 1$. In this case, the pointer is, on average, shifted by a large amount. In particular, $\mathrm{Re}(A^{\phi}_{\psi})$ can be outside of the interval $[\lambda_{\min}(\hat{A}), \lambda_{\max}(\hat{A})]$ (where $\lambda_{\min(\max)}(\hat{A}) = \min(\max)_k\,\lambda_k(\hat{A})$ and λ_k denotes the kth eigenvalue of an observable), i.e. it can be outside of the (convex hull of the) spectrum of $\hat{A}$.

In that case, the pointer's mean position moves beyond where it could have reached under simple weak measurements on an arbitrary pre-selected state without any post-selection. Indeed, in the absence of post-selection the expectation value of the pointer after the coupling (Eq. (6.4)) is (now with exact equalities)

$$\begin{aligned}\langle\hat{x}\rangle &= \langle\psi|\,\langle\varphi(0)|\,e^{i\hat{A}\hat{p}}\,(\mathbb{1}\otimes\hat{x})\,e^{-i\hat{A}\hat{p}}\,|\psi\rangle\,|\varphi(0)\rangle \\ &= \langle\psi|\,\langle\varphi(0)|\,\Big(\mathbb{1}\otimes\hat{x} + \hat{A}\otimes\mathbb{1}\Big)\,|\psi\rangle\,|\varphi(0)\rangle \\ &= \langle\psi|\,\hat{A}\,|\psi\rangle \quad\in\; [\lambda_{\min}(\hat{A}), \lambda_{\max}(\hat{A})]. \end{aligned} \tag{6.9}$$

[1] In the published paper [36], more precise conditions are given for the weak regime to be satisfied, and for the approximations made below to be valid. Here, these conditions are $\sigma \gg |a_k|\ \forall\, k$, and $\sigma \gg |A^{\phi}_{\psi}|$, where A^{ϕ}_{ψ} is the weak value defined in Eq. (6.7).

[2] $|\varphi(a)\rangle$ in Eq. (6.3) is properly normalised for $a \in \mathbb{R}$, while for a complex value of a, its norm is $e^{\frac{\mathrm{Im}(a)^2}{4\sigma^2}}$. The mean position of the pointer in the state $|\varphi(a)\rangle$ (possibly after renormalisation) is $\langle\hat{x}\rangle = \mathrm{Re}(a)$ and its variance is $\langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2 = \sigma^2$ [36].

Note that $\langle \hat{x} \rangle$, both with and without post-selection, can be determined experimentally by performing sufficiently many measurements, despite the large variance of the pointer (indeed, to obtain a given accuracy the number of measurements required scales proportionally to σ^2 [36]).

The definition (6.7) of a weak value can be generalised to post-selections on a given result for any general quantum measurement [97, 98], where the initial state is not necessarily pure, and where post-selection is conditioned on a given result of an arbitrary POVM measurement on the system, rather than a projective measurement. In that case, the weak value of the observable $\hat{A}$, given the pre-selection in the state ρ and the post-selection by the POVM element E, is defined as

$$A_\rho^E := \frac{\mathrm{Tr}(E\hat{A}\rho)}{\mathrm{Tr}(E\rho)}, \tag{6.10}$$

and the mean position of the pointer (when the post-selection is successful) is, as in Eq. (6.8), $\langle \hat{x} \rangle \approx \mathrm{Re}(A_\rho^E)$.

In particular, a trivial, deterministic measurement of the identity operator $\mathbb{1}$ (i.e., taking $E = \mathbb{1}$ in Eq. (6.10)) amounts to performing no post-selection. This allows one to also consider a *weak value with no post-selection*, defined as[3]

$$A_\psi^{\mathbb{1}} := \langle \psi | \hat{A} | \psi \rangle \,. \tag{6.11}$$

With this definition, Eq. (6.9) gives $\langle \hat{x} \rangle = A_\psi^{\mathbb{1}} = \mathrm{Re}(A_\psi^{\mathbb{1}})$: we recover the same relation as in Eq. (6.8), although now $A_\psi^{\mathbb{1}}$ is restricted to lie in $[\lambda_{\min}(\hat{A}), \lambda_{\max}(\hat{A})]$ since here it is simply equal to the expectation value of $\hat{A}$.

The phenomenon of a weak value outside the spectrum of $\hat{A}$ is referred to as an "anomalous weak value" [96, 100, 101] since it conflicts with our classical intuition, which would lead us to expect $\langle \hat{x} \rangle$ to lie within the range of the spectrum of $\hat{A}$. Given the analysis above, it is rather natural to attribute the origin of anomalous weak values to the presence of post-selection, an opinion that indeed seems to be widely shared in the community.

6.3 An anomalous weak value without post-selection

Contrary to what one might expect from the discussion in the last section, anomalous weak values can in fact be observed in the absence of post-selection and without discarding any outcomes. This effect was first found when studying whether anomalous weak values can arise in scenarios described by process matrices, with the parties performing weak measurements—a question that arises naturally in the context of the above-mentioned correspondence between two-time states and process matrices. In particular, one might conjecture that physically meaningful process matrices, such as the deterministic circuits considered in Chapter 3, cannot give rise to anomalous

[3] Note that the weak values $A_\psi^{\mathbb{1}}$ (in the absence of post-selection) and A_ψ^{ψ} (when one post-selects on the initial state, a situation called "re-selection" and studied in Ref. [99]) coincide. However, we emphasise that these correspond to different physical situations; in particular, without post-selection no data is discarded.

weak values. As it turned out, there is a counterexample to this conjecture, and physically realisable process matrices can indeed generate anomalous weak values. More specifically, we found that in a situation where two successive weak measurements of projection observables (whose spectrum is $\{0, 1\}$) are performed on a quantum system, associated with two different pointers, the average of the product of the pointer positions, as well as the corresponding weak value, can become negative.

We will now illustrate that effect with a simple example. Let us consider a qubit system initially prepared in the state $|0\rangle$, undergoing a sequence of two weak von Neumann measurements of the projection observables $|\psi_j\rangle\langle\psi_j|$ $(j = 1, 2)$. That is, the observables are successively coupled to the two pointers, which are then jointly measured (see Fig. 6.1).

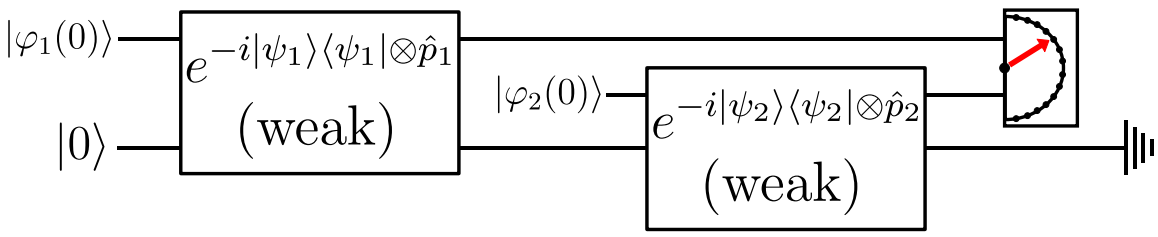

FIGURE 6.1: We consider a sequential weak measurement of two projection observables, with two different pointers that are measured jointly.

The states $|\psi_j\rangle$ and their orthogonal states $|\psi_j^\perp\rangle$ are defined as

$$\begin{aligned} |\psi_j\rangle &= \frac{1}{2}|0\rangle - (-1)^j \frac{\sqrt{3}}{2}|1\rangle\,, \\ |\psi_j^\perp\rangle &= \frac{\sqrt{3}}{2}|0\rangle + (-1)^j \frac{1}{2}|1\rangle\,. \end{aligned} \tag{6.12}$$

To each measurement is associated a pointer in the state $|\varphi_j(x_j)\rangle$, where x_j is the mean position of the pointer wavefunction. The two pointers are initially independent, and both centred at $x_j = 0$. The initial state of the system and pointers is therefore

$$|\Psi_0\rangle = |0\rangle\,|\varphi_1(0)\rangle\,|\varphi_2(0)\rangle\,. \tag{6.13}$$

Following the von Neumann measurement procedure described earlier with interaction Hamiltonians $\hat{H}_j = \gamma_j\,|\psi_j\rangle\langle\psi_j|\,\hat{p}_j$, the average post-measurement position of the corresponding pointer is (with appropriate units so that $\gamma_j\Delta t_j = 1$ as before) $x_j = 1$ if the state of the system is $|\psi_j\rangle$; if the state is $|\psi_j^\perp\rangle$ then the pointer does not move. The state of the system and pointers after the interaction with the first pointer is thus

$$\begin{aligned} |\Psi_1\rangle &= \left(|\psi_1\rangle\langle\psi_1|\,e^{-i\hat{p}_1} + |\psi_1^\perp\rangle\langle\psi_1^\perp|\,\mathbb{1}_1\right)\mathbb{1}_2\,|\Psi_0\rangle \\ &= \frac{1}{2}|\psi_1\rangle\,|\varphi_1(1)\rangle\,|\varphi_2(0)\rangle + \frac{\sqrt{3}}{2}|\psi_1^\perp\rangle\,|\varphi_1(0)\rangle\,|\varphi_2(0)\rangle\,. \end{aligned} \tag{6.14}$$

After interacting with the second pointer, it evolves to

$$
\begin{aligned}
|\Psi_2\rangle &= \left(|\psi_2\rangle\langle\psi_2|\, e^{-i\hat{p}_2} + |\psi_2^{\perp}\rangle\langle\psi_2^{\perp}|\, \mathbb{1}_2\right) \mathbb{1}_1\ |\Psi_1\rangle \\
&= -\frac{1}{4}\,|\psi_2\rangle\,|\varphi_1(1)\rangle\,|\varphi_2(1)\rangle + \frac{3}{4}\,|\psi_2\rangle\,|\varphi_1(0)\rangle\,|\varphi_2(1)\rangle \\
&\quad + \frac{\sqrt{3}}{4}|\psi_2^{\perp}\rangle|\varphi_1(1)\rangle|\varphi_2(0)\rangle + \frac{\sqrt{3}}{4}|\psi_2^{\perp}\rangle|\varphi_1(0)\rangle|\varphi_2(0)\rangle\,.
\end{aligned}
\tag{6.15}
$$

Tracing out the system, one finds that the joint pointer state is

$$
\eta_{12} = |\Phi_1^{(1)}\rangle\langle\Phi_1^{(1)}| \otimes |\phi_2(1)\rangle\langle\phi_2(1)| + |\Phi_1^{(0)}\rangle\langle\Phi_1^{(0)}| \otimes |\phi_2(0)\rangle\langle\phi_2(0)|\,, \tag{6.16}
$$

where

$$
\begin{aligned}
|\Phi_1^{(1)}\rangle &= \frac{1}{4}(\,|\varphi_1(1)\rangle - 3\,|\varphi_1(0)\rangle\,), \\
|\Phi_1^{(0)}\rangle &= \frac{\sqrt{3}}{4}(\,|\varphi_1(1)\rangle + |\varphi_1(0)\rangle\,)
\end{aligned}
\tag{6.17}
$$

are (generally unnormalised) states of the first pointer. The norms of these states, and thus the weight of each state in the mixture η_{12}, depend on the strength of the first measurement through the overlap $\langle\varphi_1(1)|\varphi_1(0)\rangle$ of the corresponding pointer states.

Finally the positions of the pointers are measured. The quantity of interest is the average of the product of the pointer positions, i.e., the expectation value $\langle\hat{x}_1 \otimes \hat{x}_2\rangle$. From Eq. (6.16), and using the facts that $\langle\varphi_2(1)|\hat{x}_2|\varphi_2(1)\rangle = 1$ and $\langle\varphi_2(0)|\hat{x}_2|\varphi_2(0)\rangle = 0$, we simply find that

$$
\langle\hat{x}_1 \otimes \hat{x}_2\rangle = \langle\Phi_1^{(1)}|\,\hat{x}_1\,|\Phi_1^{(1)}\rangle\,\langle\varphi_2(1)|\,\hat{x}_2\,|\varphi_2(1)\rangle + \langle\Phi_1^{(0)}|\,\hat{x}_1\,|\Phi_1^{(0)}\rangle\,\langle\varphi_2(0)|\,\hat{x}_2\,|\varphi_2(0)\rangle = \langle\Phi_1^{(1)}|\,\hat{x}_1\,|\Phi_1^{(1)}\rangle\,. \tag{6.18}
$$

Considering Gaussian pointers with widths σ_j for each measurement, and using the explicit forms of Eq. (6.17) for $|\Phi_1^{(1)}\rangle$, together with Eq. (6.3) for the first Gaussian pointer, one has

$$
\langle\hat{x}_1 \otimes \hat{x}_2\rangle = \frac{1}{16}(1 - 3\,\langle\varphi_1(1)|\hat{x}_1|\varphi_1(0)\rangle - 3\,\langle\varphi_1(0)|\hat{x}_1|\varphi_1(1)\rangle\,) = \frac{1}{16}\left(1 - 3e^{-\frac{1}{8\sigma_1^2}}\right). \tag{6.19}
$$

Notice that this quantity depends on σ_1 but not on σ_2: the strength of the second measurement has no effect here and can be made as strong as one wishes.

Since both observables being measured are projectors with spectra $\{0, 1\}$, one would naturally expect an average value within the range $[0, 1]$. Regardless of the strength of either measurement, each pointer, taken individually, indeed has an average position in $[0, 1]$. Specifically, one can calculate $\langle\hat{x}_1\rangle$ directly from the state after the first measurement, Eq. (6.14), as the second measurement does not interact with the first pointer at all. Observing that only the first term

from Eq. (6.14) contributes (because $\langle\varphi_1(0)|\,\hat{x}_1\,|\varphi_1(0)\rangle = 0$), we therefore have

$$\langle\hat{x}_1\rangle = \frac{1}{4}\,\langle\psi_1|\psi_1\rangle\,\langle\varphi_1(1)|\hat{x}_1|\varphi_1(1)\rangle\,\langle\varphi_2(0)|\varphi_2(0)\rangle = \frac{1}{4}\,. \tag{6.20}$$

For the second pointer, since $\langle\varphi_2(0)|\,\hat{x}_2\,|\varphi_2(0)\rangle = 0$, we find from Eqs. (6.16), (6.17) and (6.3) that

$$\langle\hat{x}_2\rangle = \langle\Phi_1^{(1)}|\Phi_1^{(1)}\rangle\,\langle\varphi_2(1)|\hat{x}_2|\varphi_2(1)\rangle = \langle\Phi_1^{(1)}|\Phi_1^{(1)}\rangle = \frac{1}{8}\left(5 - 3e^{-\frac{1}{8\sigma_1^2}}\right), \tag{6.21}$$

which ranges from $\langle\hat{x}_2\rangle \approx 1/4$ when the first measurement is weak ($\sigma_1 \gg 1$), to $\langle\hat{x}_2\rangle \approx 5/8$ when it is strong ($\sigma_1 \ll 1$). In the latter regime, Eq. (6.19) gives $\langle\hat{x}_1 \otimes \hat{x}_2\rangle \approx 1/16$, which is consistent with the above argument. However, if the first measurement is sufficiently weak, the average value can become negative; in the limit $\sigma_1 \to \infty$ we get

$$\langle\hat{x}_1 \otimes \hat{x}_2\rangle \approx -\frac{1}{8}. \tag{6.22}$$

This pointer reading is anomalous in that it gives an average value outside of the natural range of $[0, 1]$ that one would expect for the average of a product of two binary 0/1-valued measurements. As we will discuss in more detail below, this result can be linked to an anomalous weak value without post-selection, $(|\psi_2\rangle\langle\psi_2| \cdot |\psi_1\rangle\langle\psi_1|)_0^{\mathbb{1}} := \langle 0|\psi_2\rangle\langle\psi_2|\psi_1\rangle\langle\psi_1|0\rangle$ (see Eq. (6.30) below); specifically, we have here

$$\langle\hat{x}_1 \otimes \hat{x}_2\rangle \approx \mathrm{Re}\left(\,\langle 0|\psi_2\rangle\langle\psi_2|\psi_1\rangle\langle\psi_1|0\rangle\,\right) = -\frac{1}{8}. \tag{6.23}$$

We emphasise that this anomalous value is obtained despite the absence of post-selection. This effect can nevertheless be understood intuitively by considering that the second measurement acts as an effective post-selection on $|\psi_2\rangle$, as the corresponding pointer moves only in this case. This becomes apparent upon rewriting the above weak value as

$$\langle 0|\psi_2\rangle\langle\psi_2|\psi_1\rangle\langle\psi_1|0\rangle = |\langle\psi_2|0\rangle|^2\,\frac{\langle\psi_2|\psi_1\rangle\langle\psi_1|0\rangle}{\langle\psi_2|0\rangle}, \tag{6.24}$$

which differs from the standard weak value $(|\psi_1\rangle\langle\psi_1|)_0^{\psi_2}$ for a post-selection on $|\psi_2\rangle$ only by the factor $|\langle\psi_2|0\rangle|^2$, which is the probability that the projection of $|0\rangle$ onto $|\psi_2\rangle$ is successful. As it turns out, this factor ensures in particular that the anomalous weak value without post-selection cannot be arbitrary large, a fact that we prove further below. As we show in the published paper [36], a sequence of two projection observables $\hat{A}$ and $\hat{B}$ (with eigenvalues 0 and 1), the

above value of $-1/8$ for the real part is indeed the most anomalous value obtainable.

6.4 Analysis for arbitrary observables

In order to analyse more generally the phenomenon exhibited by the previous example, let us consider the general situation where a sequence of weak measurements is performed, each involving a coupling to a different pointer; these pointers can then be jointly measured following the sequence of interactions and, potentially, a post-selection on the system. Such sequential weak measurements were studied in Ref. [102, 103].

For the case of two sequential weak measurements, as in the example of the previous section, consider thus a system prepared in the pure state $|\psi\rangle$, which is subjected to a sequential weak measurement of the (generally noncommuting) observables $\hat{A}\,(=\sum_k a_k\,|a_k\rangle\langle a_k|)$ then $\hat{B}\,(=\sum_m b_m\,|b_m\rangle\langle b_m|)$, before being post-selected onto the state $|\phi\rangle$. The system-pointer interaction Hamiltonians are $\hat{H}_1=\gamma_1\hat{A}\hat{p}_1$ and $\hat{H}_2=\gamma_2\hat{B}\hat{p}_2$. We will choose again, for simplicity, the coupling constants and interaction times such that $\gamma_j\Delta t_j=1$, and take Gaussian pointers initially in the states $|\varphi_1(0)\rangle$ and $|\varphi_2(0)\rangle$ with widths σ_1 and σ_2, which dictate the measurement strengths.

Similarly to Eq. (6.4), the system-pointer state after the coupling is

$$e^{-i\hat{B}\hat{p}_2}e^{-i\hat{A}\hat{p}_1}\,|\psi\rangle\,|\varphi_1(0)\rangle\,|\varphi_2(0)\rangle=\sum_{k,l}\langle b_l|a_k\rangle\,\langle a_k|\psi\rangle\,|b_l\rangle\otimes|\varphi_1(a_k)\rangle\otimes|\varphi_2(b_l)\rangle \tag{6.25}$$

The (unnormalised) final state of the two pointers after the post-selection is given by

$$|\Psi\rangle=\sum_{k,l}\langle b_l|a_k\rangle\,\langle a_k|\psi\rangle\,\langle\phi|b_l\rangle\,|\varphi_1(a_k)\rangle\otimes|\varphi_2(b_l)\rangle\,. \tag{6.26}$$

With the expression for the Gaussian pointer of Eq. (6.3), and taking the weak regime approximation in which $\sigma_1\gg|a_k-a_{k'}|$, $\sigma_2\gg|b_l-b_{l'}|$, one finds, to the lowest order,

$$\langle\varphi_1(a_{k'})|\hat{x}_1|\varphi_1(a_k)\rangle=\frac{a_k+a_{k'}}{2}e^{-\frac{(a_k-a_{k'})^2}{8\sigma_1^2}}\approx\frac{a_k+a_{k'}}{2},\qquad\langle\varphi_1(a_{k'})|\varphi_1(a_k)\rangle=e^{-\frac{(a_k-a_{k'})^2}{8\sigma_1^2}}\approx 1, \tag{6.27}$$

and similarly for $\langle\varphi_2(b_{l'})|\hat{x}_2|\varphi_2(b_l)\rangle$ and $\langle\varphi_2(b_{l'})|\varphi_2(b_l)\rangle$. Therefore, the expectation value of the product of the pointer positions, given that the post-selection was successful, is

$$\begin{aligned}
\langle\hat{x}_1\otimes\hat{x}_2\rangle &= \frac{\mathrm{Tr}\left(\hat{x}_1\otimes\hat{x}_2\ |\Psi\rangle\langle\Psi|\right)}{\mathrm{Tr}\left(|\Psi\rangle\langle\Psi|\right)} \\
&\approx \frac{\sum_{klk'l'}\langle\phi|b_l\rangle\langle b_l|a_k\rangle\langle a_k|\psi\rangle\langle\psi|a_{k'}\rangle\langle a_{k'}|b_{l'}\rangle\langle b_{l'}|\phi\rangle)(a_k+a_{k'})(b_l+b_{l'})/4}{|\langle\phi|\psi\rangle|^2} \\
&= \frac{1}{4}\frac{\langle\phi|\hat{B}\hat{A}|\psi\rangle\langle\psi|\phi\rangle+\langle\phi|\hat{A}|\psi\rangle\langle\psi|\hat{B}|\phi\rangle+\langle\phi|\hat{B}|\psi\rangle\langle\psi|\hat{A}|\phi\rangle+\langle\phi|\psi\rangle\langle\psi|\hat{A}\hat{B}|\phi\rangle}{|\langle\phi|\psi\rangle|^2} \\
&= \frac{1}{2}\left[\mathrm{Re}\left(\frac{\langle\phi|\hat{B}\hat{A}|\psi\rangle}{\langle\psi|\phi\rangle}\right)+\mathrm{Re}\left(\frac{\langle\phi|\hat{A}|\psi\rangle}{\langle\phi|\psi\rangle}\frac{\langle\psi|\hat{B}|\phi\rangle}{\langle\psi|\phi\rangle}\right)\right] \\
&= \frac{1}{2}\left(\mathrm{Re}[(BA)^{\phi}_{\psi}]+\mathrm{Re}[A^{\phi}_{\psi}(B^{\phi}_{\psi})^*]\right).
\end{aligned} \tag{6.28}$$

Here, we introduced the *sequential weak value* $(BA)^{\phi}_{\psi}$, which, in analogy to Eq. (6.7), is defined[4] as [102]

$$(BA)^{\phi}_{\psi} := \frac{\langle\phi|\hat{B}\hat{A}|\psi\rangle}{\langle\phi|\psi\rangle}. \tag{6.29}$$

While the notion of an anomalous weak value for single (non-sequential) weak measurements is intimately linked to the pointer displacement (and even justified) by the relation $\langle\hat{x}\rangle = \mathrm{Re}(A^{\phi}_{\psi})$, we find (cf. also Refs. [102, 103]) that the relationship between the mean pointer positions and $(BA)^{\phi}_{\psi}$ is more subtle for sequential weak measurements. This cautions that some care must be taken when linking (possibly anomalous) pointer positions to weak values.

The sequential weak value of Eq. (6.11) can be generalised to the case without post-selection, by defining, in a similar way to before, the *sequential weak value with no post-selection* as

$$(BA)^{\mathbb{1}}_{\psi} := \langle\psi|\hat{B}\hat{A}|\psi\rangle\,. \tag{6.30}$$

Indeed, this quantity has previously been considered in the study of time asymmetry in sequential weak measurements [104, 105] and their quasiprobabilistic interpretation [106]. Connecting this to the pointer positions, contrary to Eq. (6.28) (which was obtained with post-selection), we recover here the direct relation[5]

$$\langle\hat{x}_1\otimes\hat{x}_2\rangle \approx \mathrm{Re}[(BA)^{\mathbb{1}}_{\psi}], \tag{6.31}$$

[4] While this terminology is standard (see, e.g., Refs. [102, 103]), note that this should be read as the weak value for measuring $\hat{A}$ then $\hat{B}$ and *not* the weak value of $\hat{B}\hat{A}$, which indeed is not a valid observable in general since it may not be Hermitian. As Eq. (6.29) shows, this quantity nonetheless behaves *as if* it were the weak value of $\hat{B}\hat{A}$.

[5] As in the single measurement case, the weak values $(BA)^{\mathbb{1}}_{\psi}$ (in the absence of post-selection) and $(BA)^{\psi}_{\psi}$ (in the case of "re-selection" [99]) coincide; cf. Footnote 3. Recall, however, that these correspond to different physical situations; crucially here, the mean pointer positions in Eqs. (6.31) and (6.28) differ in general between these scenarios.

as anticipated already in Eq. (6.23), which holds as long as the *first* measurement is sufficiently weak [106]. Indeed, the state of the pointers after tracing out the system in Eq. (6.25) is

$$\eta = \sum_{klk'} \left(\langle b_l|a_k\rangle\langle a_k|\psi\rangle\langle\psi|a_{k'}\rangle\langle a_{k'}|b_l\rangle \right) |\varphi_1(a_k)\rangle \langle\varphi_1(a_{k'})| \otimes |\varphi_2(b_l)\rangle \langle\varphi_2(b_l)| \,. \tag{6.32}$$

and we obtain

$$\begin{aligned} \langle \hat{x}_1 \otimes \hat{x}_2 \rangle = \frac{\mathrm{Tr}\,(\hat{x}_1 \otimes \hat{x}_2\, \eta)}{\mathrm{Tr}\,(\eta)} &\approx \sum_{klk'} \langle b_l|a_k\rangle\langle a_k|\psi\rangle\langle\psi|a_{k'}\rangle\langle a_{k'}|b_l\rangle\, (a_k + a_{k'})b_l/2 \\ &= \frac{1}{2}(\langle\psi|\hat{B}\hat{A}|\psi\rangle + \langle\psi|\hat{A}\hat{B}|\psi\rangle = \mathrm{Re}[(BA)^{\mathbb{1}}_{\psi}] \end{aligned} \tag{6.33}$$

This justifies that our earlier illustrative example could indeed be interpreted as yielding an anomalous weak value without post-selection. Crucially, although for a single measurement without post-selection $A^{\mathbb{1}}_{\psi}$ is simply the expectation value of $\hat{A}$, no such interpretation can be given to $(BA)^{\mathbb{1}}_{\psi}$: it is the weak value of a sequence of measurements and simply behaves *as if* it were the weak value of the operator $\hat{B}\hat{A}$ which, as already mentioned, is only Hermitian – and thus defines an observable – if $\hat{A}$ and $\hat{B}$ commute (see also footnote 4). In particular, this implies that unless $\hat{A}$ and $\hat{B}$ commute, $(BA)^{\mathbb{1}}_{\psi}$ need not be contained within the interval $[\Lambda_{\min}(\hat{A},\hat{B}), \Lambda_{\max}(\hat{A},\hat{B})]$, where $\Lambda_{\min(\max)}(\hat{A},\hat{B}) = \min(\max)_{k,\ell}\, \lambda_k(\hat{A})\lambda_\ell(\hat{B})$, as one would naturally expect for the product of outcomes for a measurement of $\hat{A}$ then $\hat{B}$ [107, 108].

Nevertheless, as we noted after Eq. (6.24), the value of $(BA)^{\mathbb{1}}_{\psi}$ cannot be amplified arbitrarily. It is possible to place a more quantitive bound on the values that it can in fact take. Using the Cauchy-Schwartz inequality, we indeed have

$$|(BA)^{\mathbb{1}}_{\psi}| = |\,\langle\psi|\hat{B}\hat{A}|\psi\rangle\,| \leq \sqrt{\langle\psi|\hat{A}^2|\psi\rangle\,\langle\psi|\hat{B}^2|\psi\rangle} \leq \|\hat{A}\|\,\|\hat{B}\|, \tag{6.34}$$

(where $\|\cdot\|$ is the spectral norm). Thus, although the mean pointer position can show anomalous weak values without post-selection, the *magnitude* of the mean pointer position cannot be pushed outside what one can obtain using strong measurements.

The bound above implies in particular that for observables with symmetric spectra (with respect to 0), the real part of the weak value – and therefore the mean product of pointer positions, see Eq. (6.31) – cannot be anomalous; anomalous pointer positions are only obtained for observables with asymmetric spectra, such as projection observables. Nevertheless, one can also obtain complex weak values for observables with symmetric spectra. Take, for example, a system initially prepared in the (+1)-eigenstate $|0\rangle$ of the Pauli matrix $\hat{\sigma}_{\mathrm{Z}}$, on which a sequential weak measurement of the Pauli observables $\hat{\sigma}_{\mathrm{Y}}$ and $\hat{\sigma}_{\mathrm{X}}$ is performed. One thus obtains $(\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}})^{\mathbb{1}}_{0} = i$. The imaginary part of the weak value here can be detected by measuring the pointer momenta [36, 103, 109]. Such complex anomalous weak values cannot be obtained without post-selection with only a single weak measurement or a sequence of weak measurements of

commuting observables, and can thus themselves be considered anomalous in this sense.

6.5 Discussion

In this chapter, we presented a result from Ref. [36], where we showed that anomalous sequential weak values can arise in scenarios without post-selection. We put the result into the context of this thesis by recalling the notion of pre- and post-selected quantum systems and their link with the process matrix formalism, and we recalled the concept of weak measurements and (anomalous) weak values. We then illustrated the main result via an example with projection observables, followed by a more general discussion of the effect.

The observation that anomalous weak values are possible without post-selection is already quite interesting in itself, since usually post-selection is considered to be a fundamental ingredient in the occurence of anomalous weak values, and it is surprising that this ingredient can be relaxed [110]. The phenomenon has been subject to experimental confirmation very recently [111].

In the publication [36], we furthermore consider the case with more than two measurements, and we derive bounds on how anomalous a weak value can be obtained in such a scenario. We also discuss some questions that arise in this context about the relation between weak values and pointer positions. In the case of sequential weak measurements, the link between the two turns out to be more subtle than for standard, single weak measurements, where the expectation value of the pointer is simply the real part of the weak value. While for two weak measurements without post-selection, the expectation value of the product of the two pointers is equal to the real part of the weak value, for more measurements (as well as for the case of two measurements with post-selection), this direct relationship does no longer hold.

Finally, we consider some implications of our results for the problem of distinguishing between different causal structures. We point out that the effect identified above does not arise in a weak measurement of a product observable on the parts of a joint system, and an anomalous weak value without post-selection could thus be used to discriminate between that case and the sequential one. An interesting question is also whether in scenarios with indefinite causal structure, anomalous weak values or pointer positions which lie further outside the expected range than possible in a well-defined causal structure can be obtained.

Chapter 7

Noncausal processes on time-delocalised subsystems

In this chapter some work in progress, conducted in collaboration with Ognyan Oreshkov and Cyril Branciard, is presented. The results of this chapter will be further developed and clarified in an upcoming manuscript.

One of the principal questions of this thesis is which process matrices—or, in the terminology of Chapter 3, quantum supermaps—have a physical interpretation, and how they can be realised in practice. In Chapter 3, we characterised some particular classes of quantum supermaps, and we showed that they can be implemented by inserting the input operations into certain circuits. In particular, we introduced the class of quantum circuits with quantum control (QC-QC)—a large class of quantum supermaps, containing the quantum switch as a prime example, which are causally nonseparable in general, and which can be realised through coherent control of which input operation is applied at any given time.

There is, however, some debate on whether these protocols can really be seen as "realisations" of the corresponding supermaps with indefinite causal order, or whether they just "simulate" these supermaps. In particular, an assertion that has often been called into question is that the input operations of the supermap are really applied once and only once in the experiments that are claimed to realise the quantum switch. This shows that it is necessary to clarify the exact sense in which these protocols with coherent control of the order of the operations "realise" the corresponding causally nonseparable process matrices, and, more generally, what one means by a "physical realisation" of a higher order transformation.

For the quantum switch and its implementation with controlled operations at different times, this issue was addressed in Ref. [38], where it was shown that the input and output systems in the mathematical description of the higher order transformation correspond to concrete, physical systems in the experiments, on which the operations are thus indeed applied once and only once. These systems are *time-delocalised subsystems*, that is, they are nontrivial subsystems of some

composite systems, consisting of elementary subsystems that are associated with different times. The arguments from Ref. [38] also apply to the quantum circuits with quantum control that we introduced in Chapter 3, justifying why they can be seen as realisations of the corresponding causally nonseparable process matrices.

As we will show here, one can even go beyond QC-QC, and find examples of supermaps outside of that class which have a realisation on time-delocalised subsystems. It was already shown in [38] that the argument generalises to all bipartite process matrices with a unitary extension—a particular class (cf. Sec. 7.3) which, although a priori more general than QC-QC, was in fact shown (in the bipartite case) to coincide with QC-QC [112, 113], cf. Sec. 7.4.

In this chapter, we now show that this proof can be generalised to the tripartite case. That is, we show that any tripartite process matrix with a unitary extension has a realisation on time delocalised subsystems, and we show how to construct the realisation and identify the subsystems explicitly. Remarkably, in the tripartite case, the class of tripartite processes with a unitary extension is strictly larger than that of QC-QC; in particular, it contains noncausal processes. The results of this chapter therefore show that noncausal processes with a realisation on time-delocalised subsystems exist.

First, we will introduce the concept of time-delocalised subsystems, and we will show how the considerations in Ref. [38] apply to the class of QC-QC supermaps that we studied in Chapter 3. Then, we recall the argument for unitary extensions of bipartite processes. We then discuss the general tripartite argument, and we discuss an example of a tripartite unitarily extensible noncausal process on time-delocalised subsystems.

7.1 Time-delocalised quantum subsystems

Let us first clarify what is understood by a time-delocalised quantum subsystem.

Consider a joint quantum system X, which is composed of several subsystems $X_1, \ldots, X_n$, associated with Hilbert spaces $\mathcal{H}^{X_1}, \ldots, \mathcal{H}^{X_n}$ (of dimensions $d_{X_1}, \ldots, d_{X_n}$). Its state space $\mathcal{H}^X$ is obtained by taking the tensor product $\mathcal{H}^X := \bigotimes_{i=1}^n \mathcal{H}^{X_i}$, and thus naturally carries a tensor product structure.

This tensor product structure is however not the only one that can be assigned to $\mathcal{H}^X$. That is, there are many other ways to view $\mathcal{H}^X$ as the state space of some multipartite quantum system. Equipping $\mathcal{H}^X$ with a tensor product structure means choosing an isomorphism (i.e., a unitary transformation) $\tau : \mathcal{H}^X \to \bigotimes_{i=1}^m \mathcal{H}^{Y_m}$, where $\mathcal{H}^{Y_1}, \ldots, \mathcal{H}^{Y_m}$ are Hilbert spaces of dimensions $d_{Y_1}, \ldots, d_{Y_m}$, with $\Pi_{i=1}^m d_{Y_i} = d_X$ [114]. Such a choice establishes a notion of locality on $\mathcal{H}^X$, and defines a decomposition of the system X into subsystems $Y_1, \ldots, Y_m$ (sometimes referred to as *virtual quantum subsystems* [115]), which are generally different from the "original" systems

$X_1, \ldots, X_n$[1]. For instance, the operators in $\mathcal{L}(\mathcal{H}^X)$ that are local on the subsystem Y_i are those of the form $\tau^{-1}(M^{Y_i} \otimes \mathbb{1}^{Y_1,\ldots,Y_{i-1}Y_{i+1}\ldots Y_m})\tau$ with $M^{Y_i} \in \mathcal{L}(\mathcal{H}^{Y_i})$.[2]

Most generally, a quantum subsystem is associated to a *tensor factor of a subspace* of some larger Hilbert space [117–119].

In a quantum circuit[3], there is some background time parameter according to which quantum operations are ordered. Each Hilbert space is thus associated with some definite time, and one usually takes tensor products of Hilbert spaces at the same time. One can however also consider a tensor product of Hilbert spaces that are associated with different times. When one then considers a tensor product structure on that joint Hilbert space which is different from the "original" one, the corresponding subsystems are in general not associated with a definite time. This is what is meant by a time-delocalised subsystem, and as we will see, it is on such time-delocalised subsystems that quantum circuits with quantum control, as well as the generalisations we present in this chapter, are realised.

7.2 Quantum circuits with quantum control of causal order on time-delocalised subsystems

For simplicity (and since it is the case of interest for the considerations in the later sections), we will focus on QC-QC with two input operations. The case of QC-QC with N operations is analogous.

The physical circuit we consider thus has the general form shown in Fig. 7.1.

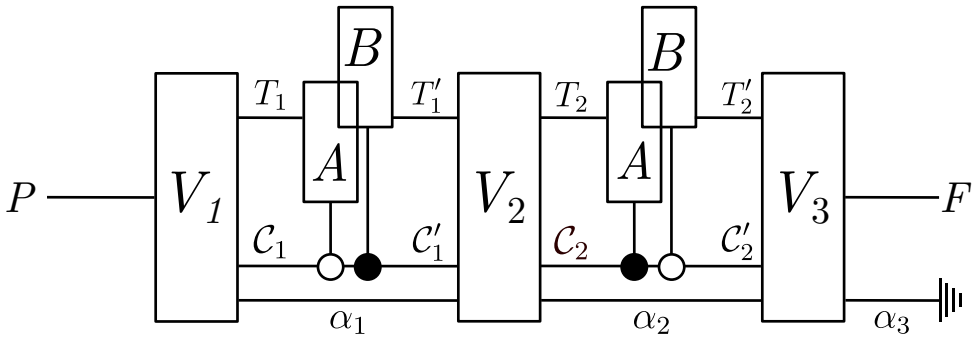

FIGURE 7.1: A QC-QC with two input operations.

The physical, time-local subsystems in this circuit are the initial and final systems P and F, the target systems T_1, T_1', T_2, T_2' (whose Hilbert spaces $\mathcal{H}^{T_1}$, $\mathcal{H}^{T_1'}$, $\mathcal{H}^{T_2}$ and $\mathcal{H}^{T_2'}$ are taken to be

[1]Except if τ is of the form $U_1^{X_1 \to Y_1} \otimes \cdots \otimes U_n^{X_n \to Y_n}$, with unitaries $U_1 : \mathcal{H}^{X_1} \to \mathcal{H}^{Y_1}$, etc. (and up to permutations).

[2]Equivalently, specifying the algebras of operators that are local on the different subsystems also defines the tensor product structure [116]. In Ref. [38], tensor product structures are defined in that way.

[3]That is, in a "standard" quantum circuit (QC-FO), and also in a QC-CC or QC-QC, when described in terms of a common "target" input and output system at each time slot on which the operations are applied (cf. Sec. 3.5.1).

of the same dimension d), the control qubits $\mathcal{C}_1, \mathcal{C}_1', \mathcal{C}_2, \mathcal{C}_2'$ and the ancillas $\alpha_1, \alpha_2, \alpha_3$. They are connected by circuit operations $\tilde{V}_1$, $\tilde{V}_2$ and $\tilde{V}_3$, which are isometries of the form (cf. (3.52)–(3.54))

$$
\begin{aligned}
|\tilde{V}_1\rangle\!\rangle &= |\tilde{V}^{\to A}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} |0\rangle^{\mathcal{C}_1} + |\tilde{V}^{\to B}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} |1\rangle^{\mathcal{C}_1} \\
|\tilde{V}_2\rangle\!\rangle &= |\tilde{V}^{\to B}_{\emptyset,A}\rangle\!\rangle^{T_1'\alpha_1 T_2\alpha_2} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_1'} + |\tilde{V}^{\to A}_{\emptyset,B}\rangle\!\rangle^{T_1'\alpha_1 T_2\alpha_2} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_1'} \\
|\tilde{V}_3\rangle\!\rangle &= |\tilde{V}^{\to F}_{\{A\},B}\rangle\!\rangle^{T_2'\alpha_2 F\alpha_3} |0\rangle^{\mathcal{C}_2'} + |\tilde{V}^{\to F}_{\{B\},A}\rangle\!\rangle^{T_2'\alpha_2 F\alpha_3} |1\rangle^{\mathcal{C}_2'} .
\end{aligned}
\tag{7.1}
$$

Since we restrict to the 2-operation case, we can simplify some of the notations that we used in Chapter 3. We choose the notation A and B (instead of A_1 and A_2) for the input operations of the circuit[4], with input (output) Hilbert spaces $\mathcal{H}^{A_I}$, $\mathcal{H}^{B_I}$ ($\mathcal{H}^{A_O}$, $\mathcal{H}^{B_O}$), and we denote the states of the control systems corresponding to "A first" (i.e., $|\emptyset, A\rangle$ and $|\{A\}, B\rangle$) by $|0\rangle$ (represented by a white dot in Fig. 7.1), and those corresponding to "B first" by $|1\rangle$ (represented by a black dot in Fig. 7.1).

Between $\tilde{V}_1$ and $\tilde{V}_2$, the controlled operation

$$
|\tilde{A}_1\rangle\!\rangle = |\tilde{A}\rangle\!\rangle^{T_1T_1'} |0\rangle^{\mathcal{C}_1} |0\rangle^{\mathcal{C}_1'} + |\tilde{B}\rangle\!\rangle^{T_1T_1'} |1\rangle^{\mathcal{C}_1} |1\rangle^{\mathcal{C}_1'} \tag{7.2}
$$

is applied. Between $\tilde{V}_2$ and $\tilde{V}_3$, we have

$$
|\tilde{A}_2\rangle\!\rangle = |\tilde{B}\rangle\!\rangle^{T_2T_2'} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_2'} + |\tilde{A}\rangle\!\rangle^{T_2T_2'} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_2'} . \tag{7.3}
$$

The global transformation from the past to the future that this circuit implements is $M^{PF} = \mathrm{Tr}_{\alpha_3} |V\rangle\!\rangle\langle\!\langle V|^{PF\alpha_3}$ with

$$
|V\rangle\!\rangle^{PF\alpha_3} = |\tilde{V}_1\rangle\!\rangle^{PT_1\mathcal{C}_1\alpha_1} * |\tilde{A}_1\rangle\!\rangle^{T_1\mathcal{C}_1T_1'\mathcal{C}_1'} * |\tilde{V}_2\rangle\!\rangle^{T_1'\mathcal{C}_1'\alpha_1T_2\mathcal{C}_2\alpha_2} * |\tilde{A}_2\rangle\!\rangle^{T_2\mathcal{C}_2T_2'\mathcal{C}_2'} * |\tilde{V}_3\rangle\!\rangle^{T_2'\mathcal{C}_2'\alpha_2F\alpha_3} . \tag{7.4}
$$

This can be simplified to

$$
\begin{aligned}
|V\rangle\!\rangle^{PF\alpha_3} =& \big(|0\rangle^{\mathcal{C}_1} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_1'} |0\rangle^{\mathcal{C}_2'} |\tilde{V}^{\to A}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} * |\tilde{V}^{\to B}_{\emptyset,A}\rangle\!\rangle^{T_1'\alpha_1T_2\alpha_2} * |\tilde{V}^{\to F}_{\{A\}B}\rangle\!\rangle^{T_2'\alpha_2F\alpha_3} \\
&+ |1\rangle^{\mathcal{C}_1} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_1'} |1\rangle^{\mathcal{C}_2'} |\tilde{V}^{\to B}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} * |\tilde{V}^{\to A}_{\emptyset,B}\rangle\!\rangle^{T_1'\alpha_1T_2\alpha_2} * |\tilde{V}^{\to F}_{\{B\},A}\rangle\!\rangle^{T_2'\alpha_2F\alpha_3} \big) \\
&* \big(|0\rangle^{\mathcal{C}_1} |0\rangle^{\mathcal{C}_1'} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_2'} |\tilde{A}\rangle\!\rangle^{T_1T_1'} |\tilde{B}\rangle\!\rangle^{T_2T_2'} + |1\rangle^{\mathcal{C}_1} |1\rangle^{\mathcal{C}_1'} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_2'} |\tilde{B}\rangle\!\rangle^{T_1T_1'} |\tilde{A}\rangle\!\rangle^{T_2T_2'} \big).
\end{aligned}
\tag{7.5}
$$

The process matrix corresponding to this QC-QC (cf. Def. 3.7 and Proposition 3.8) is $\mathrm{Tr}_{\alpha_3} |w\rangle\langle w|$, with the process vector

$$
\begin{aligned}
|w\rangle^{PA_{IO}B_{IO}F\alpha_3} = |V^{\to A}_{\emptyset,\emptyset}\rangle\!\rangle^{PA_I\alpha_1} * |V^{\to B}_{\emptyset,A}\rangle\!\rangle^{A_O\alpha_1B_I\alpha_2} * |V^{\to F}_{\{A\},B}\rangle\!\rangle^{B_O\alpha_2F\alpha_3} \\
+ |V^{\to B}_{\emptyset,\emptyset}\rangle\!\rangle^{PB_I\alpha_1} * |V^{\to A}_{\emptyset,B}\rangle\!\rangle^{B_O\alpha_1A_I\alpha_2} * |V^{\to F}_{\{B\},A}\rangle\!\rangle^{A_O\alpha_2F\alpha_3}
\end{aligned}
\tag{7.6}
$$

[4]For the operations themselves, as well as for their labels.

and when we insert the operations, we obtain

$$|V\rangle\!\rangle^{PF\alpha_3} = |w\rangle^{PA_{IO}B_{IO}F\alpha_3} * |A\rangle\!\rangle^{A_{IO}} * |B\rangle\!\rangle^{B_{IO}} \,. \tag{7.7}$$

The claim is now that for a suitable choice of tensor product structure, giving rise to time-delocalised subsystems, Eq. (7.5) takes the form of Eq. (7.7) (i.e., the circuit in Fig. 7.1 is indeed precisely described by Eq. (7.7)).

To see this, let us consider the Hilbert space $\mathcal{H}_{\text{in}} \coloneqq \mathcal{H}^{T_1T_2} \otimes \text{span}\{|00\rangle^{\mathcal{C}_1\mathcal{C}_2}, |11\rangle^{\mathcal{C}_1\mathcal{C}_2}\}$, which is a subspace of $\mathcal{H}^{T_1T_2\mathcal{C}_1\mathcal{C}_2}$, and assign it a tensor product structure $\tau_{\text{in}} : \mathcal{H}_{\text{in}} \to \mathcal{H}^{A_I} \otimes \mathcal{H}^{B_I} \otimes \mathcal{H}^R$ (with $\dim \mathcal{H}^{A_I} = \dim \mathcal{H}^{B_I} = d$ and $\dim \mathcal{H}^R = 2$), given by the isomorphism

$$\tau_{\text{in}} = |0\rangle^R \langle 00|^{\mathcal{C}_1\mathcal{C}_2} \otimes \mathbb{1}^{T_1\to A_I} \otimes \mathbb{1}^{T_2\to B_I} + |1\rangle^R \langle 11|^{\mathcal{C}_1\mathcal{C}_2} \otimes \mathbb{1}^{T_1\to B_I} \otimes \mathbb{1}^{T_2\to A_I} \tag{7.8}$$

(with $\mathbb{1}^{X\to Y}$ denoting the identity from $\mathcal{H}^X$ to $\mathcal{H}^Y$). Similarly, we consider the Hilbert space $\mathcal{H}_{\text{out}} \coloneqq \mathcal{H}^{T_1'T_2'} \otimes \text{span}\{|00\rangle^{\mathcal{C}_1'\mathcal{C}_2'}, |11\rangle^{\mathcal{C}_1'\mathcal{C}_2'}\}$ (which is a subspace of $\mathcal{H}^{T_1'T_2'\mathcal{C}_1'\mathcal{C}_2'}$) and assign it a tensor product structure $\tau_{\text{out}} : \mathcal{H}_{\text{out}} \to \mathcal{H}^{A_O} \otimes \mathcal{H}^{B_O} \otimes \mathcal{H}^{R'}$, given by the isomorphism

$$\tau_{\text{out}} = |0\rangle^{R'} \langle 00|^{\mathcal{C}_1'\mathcal{C}_2'} \otimes \mathbb{1}^{T_1'\to A_O} \otimes \mathbb{1}^{T_2'\to B_O} + |1\rangle^{R'} \langle 11|^{\mathcal{C}_1'\mathcal{C}_2'} \otimes \mathbb{1}^{T_1'\to B_O} \otimes \mathbb{1}^{T_2'\to A_O}. \tag{7.9}$$

With respect to these choices of tensor product structures, we have

$$\begin{aligned}
&|0\rangle^{\mathcal{C}_1} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_1'} |0\rangle^{\mathcal{C}_2'} |\tilde{V}^{\to A}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} * |\tilde{V}^{\to B}_{\emptyset,A}\rangle\!\rangle^{T_1'\alpha_1T_2\alpha_2} * |\tilde{V}^{\to F}_{\{A\}B}\rangle\!\rangle^{T_2'\alpha_2F\alpha_3} \\
&\qquad + |1\rangle^{\mathcal{C}_1} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_1'} |1\rangle^{\mathcal{C}_2'} |\tilde{V}^{\to B}_{\emptyset,\emptyset}\rangle\!\rangle^{PT_1\alpha_1} * |\tilde{V}^{\to A}_{\emptyset,B}\rangle\!\rangle^{T_1'\alpha_1T_2\alpha_2} * |\tilde{V}^{\to F}_{\{B\},A}\rangle\!\rangle^{T_2'\alpha_2F\alpha_3} \\
\equiv\ &|0\rangle^R |0\rangle^{R'} |V^{\to A}_{\emptyset,\emptyset}\rangle\!\rangle^{PA_I\alpha_1} * |V^{\to B}_{\emptyset,A}\rangle\!\rangle^{A_O\alpha_1B_I\alpha_2} * |V^{\to F}_{\{A\},B}\rangle\!\rangle^{B_O\alpha_2F\alpha_3} \\
&\qquad + |1\rangle^R |1\rangle^{R'} |V^{\to B}_{\emptyset,\emptyset}\rangle\!\rangle^{PB_I\alpha_1} * |V^{\to A}_{\emptyset,B}\rangle\!\rangle^{B_O\alpha_1A_I\alpha_2} * |V^{\to F}_{\{B\},A}\rangle\!\rangle^{A_O\alpha_2F\alpha_3}
\end{aligned} \tag{7.10}$$

and

$$\begin{aligned}
&|0\rangle^{\mathcal{C}_1} |0\rangle^{\mathcal{C}_1'} |0\rangle^{\mathcal{C}_2} |0\rangle^{\mathcal{C}_2'} |\tilde{A}\rangle\!\rangle^{T_1T_1'} |\tilde{B}\rangle\!\rangle^{T_2T_2'} + |1\rangle^{\mathcal{C}_1} |1\rangle^{\mathcal{C}_1'} |1\rangle^{\mathcal{C}_2} |1\rangle^{\mathcal{C}_2'} |\tilde{B}\rangle\!\rangle^{T_1T_1'} |\tilde{A}\rangle\!\rangle^{T_2T_2'} \\
&\qquad \equiv |\mathbb{1}\rangle\!\rangle^{RR'} |A\rangle\!\rangle^{A_IA_O} |B\rangle\!\rangle^{B_IB_O}
\end{aligned} \tag{7.11}$$

and when taking the link product of the right-hand sides in Eq. (7.10) and Eq. (7.11), we indeed obtain Eq. (7.7). We thus have indeed identified the concrete, physical (time-delocalised) subsystems with respect to which the circuit in Fig. 7.1 is precisely described by the process matrix (7.6). The input systems A_I and B_I are subsystems of $T_1T_2\mathcal{C}_1\mathcal{C}_2$ (the linear operators in $\mathcal{L}(\mathcal{H}^{T_1T_2\mathcal{C}_1\mathcal{C}_2})$ that are local to these subsystems are given by $\tau_{\text{in}}^{-1}(M^{A_I} \otimes \mathbb{1}^{B_IR})\tau_{\text{in}}$ and $\tau_{\text{in}}^{-1}(M^{B_I} \otimes \mathbb{1}^{A_IR})\tau_{\text{in}}$, respectively) and the output systems A_O and B_O are subsystems of $T_1'T_2'\mathcal{C}_1'\mathcal{C}_2'$ (the linear operators in $\mathcal{L}(\mathcal{H}^{T_1'T_2'\mathcal{C}_1'\mathcal{C}_2'})$ that are local to these subsystems are given by $\tau_{\text{out}}^{-1}(M^{A_O} \otimes \mathbb{1}^{B_OR'})\tau_{\text{out}}$ and $\tau_{\text{out}}^{-1}(M^{B_O} \otimes \mathbb{1}^{A_OR})\tau_{\text{out}}$, respectively).

The fact that Alice's and Bob's operations are applied on these input and output systems can in principle experimentally tested by tomography, i.e., by preparing suitable states on the input systems, and suitable measurements on the output systems [38].

7.3 Bipartite unitarily extensible process matrices on time-delocalised subsystems

We will now move on to a particular class of unitary quantum supermaps with two input operations, which were first introduced in Ref. [60]. Such supermaps are defined as those which output a unitary transformation whenever all input operations are unitaries, and they are characterised by rank-one process matrices of the form $W_U = |U\rangle\!\rangle\langle\!\langle U|$, where U is a unitary $U : \mathcal{H}^{PA_OB_O} \to \mathcal{H}^{FA_IB_I}$.

In Ref. [60], the question was raised whether any process matrix can be recovered from a larger process matrix of that form, by preparing some fixed state in the global past space, and tracing out the global future. It was found that this is not the case, and that there are process matrices which do not have such a unitary extension (for instance, W^{ocb} from Sec. 1.2.2).

It was then shown in Ref. [38] that any unitary bipartite process matrix has a realisation on time-delocalised subsystems. We will briefly sketch the proof here. For a more detailed discussion, see Ref. [38].

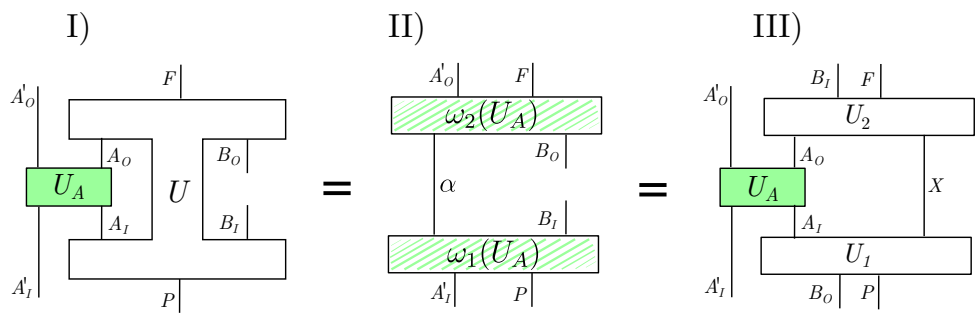

FIGURE 7.2: For a unitary process with two input operations, we have that $|U\rangle\!\rangle^{PA_{IO}B_{IO}F} * |U_A\rangle\!\rangle^{A_{IO}A'_{IO}} = |\omega_1(U_A)\rangle\!\rangle^{PA'_IB_I\alpha} * |\omega_2(U_A)\rangle\!\rangle^{B_O\alpha A'_OF} = |U_1\rangle\!\rangle^{B_OPA_IX} * |U_2\rangle\!\rangle^{A_OXB_IF} * |U_A\rangle\!\rangle^{A_{IO}A'_{IO}}$.

The argument is based on the equivalence shown graphically in Fig. 7.2. One considers some bipartite unitary supermap, described by a process matrix $|U\rangle\!\rangle\langle\!\langle U|$. When one plugs a given operation into the slot of Alice (I) (which we can assume to be unitary without loss of generality, since one can always use additional ancillary spaces $\mathcal{H}^{A'_I}$ and $\mathcal{H}^{A'_O}$ to purify it), one obtains a reduced supermap with one input operation (that of Bob). This reduced supermap is a quantum circuit with fixed causal order (trivially, as it has only one input operation). It therefore has a realisation as in (II) (with circuit operations $\omega_1(U_A)$ and $\omega_2(U_A)$ that depend on Alice's operation, and that can be taken to be unitaries, as the reduced one-operation supermap is still unitary).

The claim is now that this circuit (II) is a realisation of the process $|U\rangle\!\rangle\langle\!\langle U|$ on time delocalised subsystems, as one can again establish a precise correpondence between the input and output systems in the process matrix, and some concrete, physical input and output systems in the circuit.

For Bob's operation, this is obvious. It is applied at a fixed time, on the time-local input and output systems B_I and B_O. The situation is not so clear for Alice's operation, which is applied somehow "within the circuit operations" $\omega_1(U_A)$ and $\omega_2(U_A)$ that constitute the circuit (II). It turns out that one can make this statement precise, and identify the concrete, physical time-delocalised subsystems on which Alice's operation takes place in (II).

Alice's time-delocalised input system is a subsystem of PB_O, and Alice's time-delocalised output system is a subsystem of B_IF. To see this, and to identify these subsystems explicitly, one considers the right-hand side (III) of Fig. 7.2. The equivalence between (I) and (III) follows from the fact that we can decompose the unitary $|U\rangle\!\rangle$ as

$$|U\rangle\!\rangle = |U_1\rangle\!\rangle^{B_OPA_IX} * |U_2\rangle\!\rangle^{A_OXB_IF} \tag{7.12}$$

with unitaries $U_1 : \mathcal{H}^{PB_O} \to \mathcal{H}^{A_IX}$, and $U_2 : \mathcal{H}^{A_OX} \to \mathcal{H}^{B_IF}$. Such a decomposition always exists, and can be constructed explicitly. This follows from the validity conditions satisfied by $|U\rangle\!\rangle\langle\!\langle U|$, which imply that ${}_{[1-A_O]B_IF}\,|U\rangle\!\rangle\langle\!\langle U| = 0$, as well as ${}_{[1-PB_O]A_{IO}B_IF}\,|U\rangle\!\rangle\langle\!\langle U| = 0$[5]. These are precisely the conditions that mathematically characterise a quantum comb with global past space $\mathcal{H}^{PB_O}$, global future space $\mathcal{H}^{B_IF}$ and one input operation for Alice (cf. Eq. (2.12)).

It follows that Alice's input space $\mathcal{H}^{A_I}$ can be identified with a tensor factor of the Hilbert space $\mathcal{H}^{PB_O}$ [6], and Alice's output space $\mathcal{H}^{A_O}$ can be identified with a tensor factor of $\mathcal{H}^{B_IF}$. Again, the claim that Alice acts precisely on the systems we identified can in principle be tested experimentally by tomography [38].

7.4 Generalisation to tripartite unitarily extensible processes

In Ref. [38], the question was raised whether the bipartite construction can be generalised to more parties. Indeed, the equivalence between (I) and (III) in Fig. 7.2 generalises straightforwardly to

[5] To see this, consider the validity conditions (cf. Eq. (2.11)) of a process matrix with global past P and global future F, and two parties in between

$$\begin{aligned} &(i) \quad {}_{[1-B_O]A_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| = 0 \qquad (ii) \quad {}_{[1-A_O]B_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| = 0 \\ &(iii) \quad {}_{[1-A_O][1-B_O]F}\,|U\rangle\!\rangle\langle\!\langle U| = 0 \qquad (iv) \quad {}_{[1-P]A_{IO}B_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| = 0. \end{aligned} \tag{7.13}$$

Adding (ii) to (iii) with B_I traced out and replaced yields ${}_{[1-A_O]B_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| + {}_{[1-A_O][1-B_O]B_IF}\,|U\rangle\!\rangle\langle\!\langle U| = {}_{[1-A_O]B_IF}\,|U\rangle\!\rangle\langle\!\langle U| = 0$. Adding (i) with B_I traced out and replaced to (iv) yields ${}_{[1-B_O]B_IA_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| + {}_{[1-P]A_{IO}B_{IO}F}\,|U\rangle\!\rangle\langle\!\langle U| = {}_{[1-PB_O]A_{IO}B_IF}\,|U\rangle\!\rangle\langle\!\langle U| = 0$.

[6] More precisely, the argument in Ref. [38] is that there is a tensor product structure $\mathcal{H}^{PB_O} \cong \mathcal{H}^{\tilde{A}_I} \otimes \mathcal{H}^{\tilde{X}}$, such that U_1 maps identically $\mathcal{H}^{\tilde{A}_I}$ to $\mathcal{H}^{A_I}$, and $\mathcal{H}^{\tilde{X}}$ to $\mathcal{H}^{X}$. It is then argued that the identities can be suppressed, so that U_2 can directly be taken to relate the different subsystem decompositions [38]; similarly for Alice's output system.

the multipartite case. One can, for instance, decompose any tripartite unitarily entended process as on the right-hand side in Fig. 7.3, which follows analogously from the validity conditions.

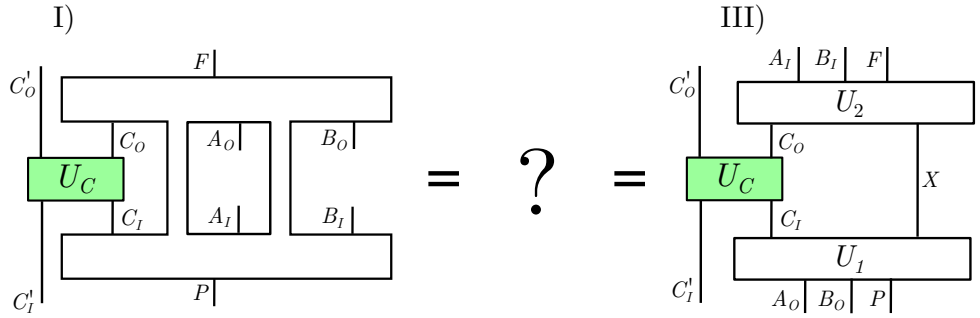

FIGURE 7.3: For three parties, we still have the equivalence between I) and III), i.e., $|U\rangle\!\rangle^{PA_{IO}B_{IO}C_{IO}F} * |U_C\rangle\!\rangle^{C_{IO}C'_{IO}} = |U_1\rangle\!\rangle^{A_OB_OPC_IX} * |U_1\rangle\!\rangle^{C_OXA_IB_IF} * |U_C\rangle\!\rangle^{C_{IO}C'_{IO}}$, but it is not so clear how step II) from the bipartite proof generalises.

We can therefore identify, for instance, Charlie's input system C_I with a subsystem of A_OB_OP, and his output system C_O with a subsystem of A_IB_IF (a fact that was also pointed out in Ref. [120]).

It is however not so clear how the step (II) from the above bipartite proof generalises to the multipartite case, i.e., how one can actually construct a circuit in which all operations are applied on concrete, (possibly time-delocalised) physical systems. The bipartite proof relies crucially on the fact that Bob's operation, and his input and output systems can be "fixed" in time, such that the subsystems of Alice can then be identified with respect to the fixed, time-localised ones. Here, one would need to "fix" both Alice and Bob in order to identify concrete, physical input and output systems also for the remaining party Charlie.

It now turns out that, by combining our considerations about the characterisation of QC-QC from Chapter 3 with the idea of the above Section 7.3, such a construction is actually possible. Suppose that we have some tripartite unitary process matrix $|U\rangle\!\rangle\langle\!\langle U| \in \mathcal{L}(\mathcal{H}^{PA_{IO}B_{IO}C_{IO}F})$, which has the property that whenever we insert a unitary operation $|U_C\rangle\!\rangle^{C_{IO}C'_{IO}}$ into the slot of Charlie, the reduced supermap with two input operations (for Alice and Bob) satisfies our characterisation of quantum circuits with quantum control. For a process of that kind, we can complete Fig. 7.3 with the circuit shown in Fig. 7.4, i.e., a QC-QC (as in Fig. 7.1), with circuit operations that depend on Charlie's operation U_C.

In such a circuit, Alice's and Bob's input and output systems are time-delocalised through quantum control, precisely in the way described in Section 7.2, and can thus be identified in the same way as described above; they are time-delocalised subsystems of $T_1T_2\mathcal{C}_1\mathcal{C}_2$ and $T_1'T_2'\mathcal{C}_1'\mathcal{C}_2'$, respectively. Charlie's systems can then be identified as subsystems of the (now known, fixed) subsystems A_OB_OP and A_IB_IF, respectively, through the decomposition (III) on the right hand side of Fig. 7.3, analogously to the bipartite case.

The question is now, do such examples of tripartite unitary process matrices, with the reduced bipartite process matrix being that of a QC-QC for any operation performed by Charlie, actually

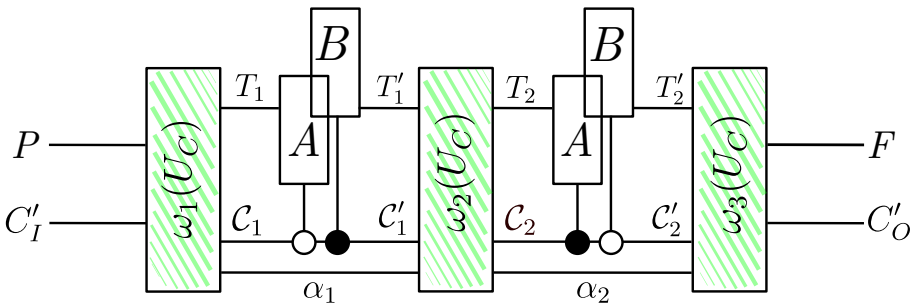

FIGURE 7.4: Any tripartite unitary process matrix has a realisation on time-delocalised subsystem as shown here, i.e., as a QC-QC $|\omega_1(U_C)\rangle\!\rangle^{PC_I'T_1\mathcal{C}_1\alpha_1} * |\omega_2(U_C)\rangle\!\rangle^{T_1'\mathcal{C}_1'\alpha_1T_2\mathcal{C}_2\alpha_2} * |\omega_3(U_C)\rangle\!\rangle^{T_2'\mathcal{C}_2'\alpha_2FC_O'}$ with circuit operations depending on Charlie's operation.

exist? In fact, all tripartite unitary processes satisfy this property. This follows from a result shown recently by Barrett, Lorenz and Oreshkov, who proved that all bipartite process matrices (with trivial P and F) that have a unitary extension are causally separable [112, 113]. This implies that all unitary extensions of bipartite process matrices are in the QC-QC class,[7] and therefore shows that all tripartite unitary processes have the desired property (since inserting a unitary operation into the slot of Charlie leads to a bipartite unitary process for Alice and Bob).

This established the main result of this chapter: All unitarily extensible tripartite processes have a realisation on time-delocalised subsystems. In the next section, we will discuss a concrete example.

7.5 Example: A noncausal process on time-delocalised subsystems

In Ref. [67], it was shown that, for three and more parties, there exist process matrices that violate causal inequalities and that can be interpreted as "classical" process matrices, since they are diagonal in the computational basis. An example, introduced in Ref. [67], is the process matrix

$$W_{\text{bw}} = \sum_{abc} |a,b,c\rangle\langle a,b,c|^{A_OB_OC_O} \otimes |\neg b\wedge c, \neg c\wedge a, \neg a\wedge b\rangle\langle \neg b\wedge c, \neg c\wedge a, \neg a\wedge b|^{A_IB_IC_I}$$

where $a,b,c\in\{0,1\}$.

[7] Indeed, assume a bipartite unitary process matrix $|U\rangle\!\rangle\langle\!\langle U| \in \mathcal{L}(\mathcal{H}^{PA_{IO}B_{IO}F})$ which is not in the QC-QC class. $\text{Tr}_F\, |U\rangle\!\rangle\langle\!\langle U|$ would then be causally nonseparable, and $W_{\text{ext}} = |U\rangle\!\rangle\langle\!\langle U| \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|^{P'A_I'} \otimes |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|^{A_O'F'}$ would be a unitary extension of the causally nonseparable bipartite (with trivial P and F) process matrix $\text{Tr}_{FF'}\, W_{\text{ext}} * |\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|^{PP'} \in \mathcal{L}(\mathcal{H}^{A_{IO}A_{IO}'B_{IO}})$, in contradiction with the result of Barrett, Lorenz and Oreshkov.

In Ref. [60], W_{bw} was then shown to have a unitary extension $|U_{\text{bw}}\rangle\rangle\langle\langle U_{\text{bw}}|$, with

$$|U_{\text{bw}}\rangle\rangle = \sum_{\substack{abc\\ijk}} |a,b,c\rangle^{F} |a,b,c\rangle^{A_O B_O C_O} \otimes |i \oplus \neg b \wedge c, j \oplus \neg c \wedge a, k \oplus \neg a \wedge b\rangle^{A_I B_I C_I} |i,j,k\rangle^{P} \tag{7.14}$$

(with $i, j, k \in \{0, 1\}$, i.e., P and F consisting of three qubits each).

How does a realisation of this process matrix on time-delocalised subsystems, according to what we outlined in the last section, look like? An example of such a realisation is given by the circuit shown in Fig. 7.5 (a similar circuit corresponding to this process has also been found in Ref. [120]).

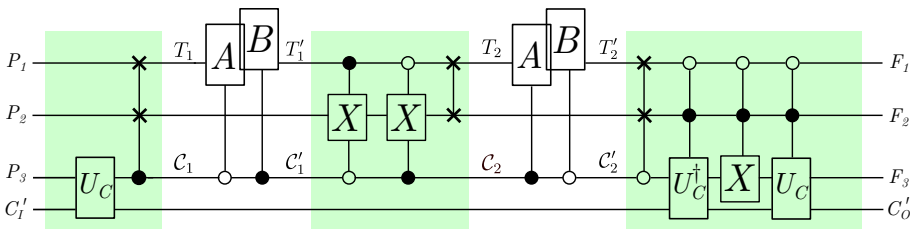

FIGURE 7.5: Realisation on time-delocalised subsystems of $|U_{\text{bw}}\rangle\rangle\langle\langle U_{\text{bw}}|$.

Note that this circuit is, in fact, classical, and its implementation is straightforward.

Intuitively speaking, in this circuit, Charlie determines the state of the control systems, which control the order of Alice and Bob. After they have acted, the operation of Charlie can be "reversed and applied again", and whether this happens or not is determined jointly by Alice and Bob. Importantly, however, from the perspective of time-delocalised subsystems, the operation of Charlie is not applied multiple times. Instead, each operation is applied once and only once, on some time-delocalised input and output systems, with respect to which the circuit in Fig. 7.5 is precisely described by the process matrix (7.14).

Alice and Bob are time-delocalised via coherent control, and their input and output systems A_I and B_I are time-delocalised subsystems of $T_1 T_2 \mathcal{C}_1 \mathcal{C}_2$ (as specified in Eq. (7.8)), and their output systems A_O and B_O are time-delocalised subsystems of $T_1' T_2' \mathcal{C}_1' \mathcal{C}_2'$ (as specified in Eq. (7.9)). Charlie's systems can then be identified as subsystems of $A_O B_O P$ and $A_I B_I F$, respectively, with the different subsystem decompositions related by the unitaries on the right hand side (III) in Fig. 7.3. For this particular process matrix, such a decomposition is given by $|U\rangle\rangle = |U_1\rangle\rangle * |U_2\rangle\rangle$, with the unitaries $U_1 : \mathcal{H}^{P A_O B_O} \to \mathcal{H}^{C_I X}$ and $U_2 : \mathcal{H}^{C_O X} \to \mathcal{H}^{A_I B_I F}$

$$U_1 = \sum_{ijkab} |k \oplus \neg\, a \wedge b, i, j, a, b\rangle^{C_I X} \langle i, j, k, a, b|^{P A_O B_O} \tag{7.15}$$

$$U_2 = \sum_{pqcxy} |p \oplus \neg\, y \wedge c, q \oplus \neg\, c \wedge x, x, y, c\rangle^{A_I B_I F} \langle c, p, q, x, y|^{C_O X} \tag{7.16}$$

7.6 Discussion

In this chapter, we continued investigating the question of which quantum processes with indefinite causal order have a physical realisation in quantum theory. As we pointed out, in order to shed more light on this question, an important step is to actually clarify more precisely what it means to say that a given process matrix has a quantum realisation. In Ref. [38], this issue was adressed, in particular with the aim of putting on a firm ground the claim that certain experimental procedures are indeed physical realisations of the quantum switch. It was shown that these experiments indeed realise the quantum switch transformation, in a precise sense, with the operations acting on time-delocalised subsystems, and it was shown that this argument generalises to all unitary extensions of bipartite process matrices. In this chapter, we further generalised that argument to unitary extensions of tripartite process matrices, a class which is strictly larger than the process matrices obtainable through quantum control that we characterised in Chapter 3, and which in particular contains processes violating causal inequalities. These processes can thus be said to exist in quantum theory in that same sense.

One may however question whether the violation of causal inequalities is really meaningful as a device-independent concept to qualify these particular realisations as incompatible with a definite causal order. Indeed, what makes causal inequality violations "surprising" is the idea of "closed laboratories" in which the parties act. It is not so clear whether such a closed laboratory assumption is satisfied in the realisations of noncausal processes we identified here.

Remember that the circuit we found for the example W_{bw} in Sec. 7.5 is classical. It can therefore be studied in the framework of classical causal models, a formalism where one considers classical, random variables that are connected in a graph, by arrows that indicate the causal relations between them. Recently, it was pointed out by Oreshkov [112] that in this framework, an argument can in fact be made to capture the non-causality of such a circuit in a device-independent way, by formulating an analogy of the closed laboratory assumption. In this sense, the circuit in Fig. 7.5 can be understood as a realisation of a cyclic causal model, with *time-delocalised variables*, giving rise to correlations that cannot be explained in the presence of a definite causal order. It will be interesting to study this further.

Another open question is whether the proof can be generalised to the case of $N \geq 4$ parties. In the tripartite case, we have a particular, standard form for the reduced supermap that we obtain when plugging in an operation for one of the parties, and that allows us to identify the systems; it is not clear whether and how this extends to the general multipartite case, and it would be interesting to determine whether such a standard form can also be found in the multipartite case, potentially allowing for a generalisation of the proof. Nevertheless, we may be able to characterise particular classes of multipartite supermaps for which the construction is possible. In particular, the argument we presented here works for all supermaps such that the reduced supermap, when plugging in an operation for one of the parties, is a QC-QC.

Conclusion

One principal aim of this work was to gain a deeper understanding of multipartite quantum causal relations, by extending the mathematical and conceptual tools required for their analysis to the multipartite case. Another main goal was to obtain more insight into the question of which multipartite processes with indefinite causal order have a physical interpretation, and to find new examples of physically realisable processes with an indefinite causal order. A third objective was to investigate the potential usefulness of these examples for quantum information processing beyond "standard", causally ordered quantum circuits.

We started our exploration into these questions by studying the concept of multipartite causal (non)separability, which qualifies multipartite scenarios that are incompatible with a definite causal order. We identified and resolved an inconsistency in previous definitions of this concept, and established the previously introduced notion of "extensible causal separability" as the most natural generalisation of the bipartite concept to the multipartite case. We then provided criteria allowing to certify the absence of a definite causal order in practice, by constructing multipartite witnesses of causal nonseparability. The main open question that remains is that about the complete characterisation of causally separable process matrices for the case of $N \geq 4$ parties, where we found distinct sufficient and necessary conditions, and we do not know whether they coincide, or whether one is both necessary and sufficient. This problem is particularly intriguing in the light of Chapter 3, where we equipped the process matrices satisfying the sufficient condition with an operational interpretation, by showing that they correspond precisely to *quantum circuits with classical control of causal order*. It would be important to clarify whether these represent all causally separable multipartite processes, or whether there exist causally separable processes that do not fall into this category. Another important next step, naturally extending the work presented in this thesis, is to formally develop and characterise a notion of "genuinely multipartite causal nonseparability", analogously to the concept of "genuinely multipartite noncausal correlations" studied in Chapter 5.

In the third chapter, we then went further beyond the above-mentioned quantum circuits with classical control. We introduced *quantum circuits with quantum control of causal order*, a class of generally causally nonseparable multipartite processes that generalise the quantum switch. We identified the subset of multipartite process matrices that corresponds to these circuits, and we also characterised it in terms of simple conditions. On the one hand, this will allow us to identify and study interesting new examples of causally nonseparable processes, with features different

from that of the quantum switch, that can be practically realised and studied in laboratory experiments.

Studying the properties of these processes for which we have an operational interpretation will also allow us to investigate in a comprehensive way what new computational protocols and tasks these processes allow to realise, which will enable a more systematic and general understanding of causal nonseparability as a computational resource. A particularly interesting perspective, opened up by the characterisations given in Chapter 3, is the systematic search for advantages of these processes through semidefinite optimisation. We already gave a few examples of new advantages that we could identify in this way, and we think that there are many more results to be found in this direction.

By defining the class of quantum circuits with quantum control of causal order, and characterising the subset of process matrices they define, we have considerably reduced the gap between the causally nonseparable processes that we know to have a physical realisation, and the set of all quantum processes, obtained through a "top-down" approach from very general considerations. However, there is still a gap that remains, and not all process matrices can be realised through quantum control of causal order. In particular, quantum circuits with quantum control of causal order do not violate causal inequalitites. It will be important to understand precisely how this gap comes about, and how the process matrices outside of the class of QC-QC are to be interpreted. Our results of Chapter 7, where we show that certain circuits can be understood as realisations of some of these process matrices on time-delocalised subsystems, are a promising first step into this direction, which deserves to be further explored.

Another possible direction is to study what possibilities there are beyond the process matrix framework, for instance by building on the suggestions put forward in Chapter 4.

Résumé en Français

Contexte et considérations préliminaires

Le concept de la causalité est essentiel en science, et pour notre perception du monde en général. L'idée habituelle est la suivante : les événements suivent un ordre causal bien défini, avec des événements dans le passé qui peuvent influencer des événements dans le futur, et non l'inverse. On peut se demander si cette notion est vraiment fondamentale, ou bien si une situation qui n'est pas compatible avec un tel ordre causal est concevable. Cette question se pose notamment dans le domaine de la théorie quantique, où il y a de nombreux concepts et phénomènes étranges qui contredisent notre intuition. Nous pouvons alors nous demander si les relations causales entre les événements peuvent elles-mêmes être soumises à des effets quantiques. En considérant deux événements A et B, peut-on par exemple avoir une superposition quantique de « A est la cause de B » et « B est la cause de A » ? Ces questions sont très importantes d'un point de vue fondamental, ainsi que d'un point de vue plus appliqué dans le contexte de l'informatique quantique. De telles relations causales pourraient permettre d'aller au-delà du paradigme habituel qui est utilisé dans ce domaine : celui des circuits quantiques, qui présuppose un ordre causal fixe. Ces dernières années, une grande attention a donc été portée à l'étude des relations causales en théorie quantique, avec plusieurs approches qui explorent ce type de questions sous différents angles [4–10].

Une approche particulière pour étudier des relations causales entre des événements quantiques est le formalisme des matrices de processus (*process matrix formalism*), introduit dans [4]. Dans cette approche, on considère des opérations réalisées par des parties, qui suivent les lois de la théorie quantique « localement ». Cela signifie que les parties reçoivent un système quantique, choisissent et effectuent une opération qui donne lieu à un résultat de mesure, et renvoient un système en sortie. Par contre, on ne présuppose pas de structure causale « globale » dans laquelle les parties sont intégrées a priori. Par cette approche, on retrouve d'une part des scénarios quantiques standards, comme des états quantiques partagés, ou des circuits quantiques qui les connectent d'une manière causalement ordonnée, d'autre part, on trouve aussi des scénarios qui ne sont pas compatibles avec un ordre causal bien défini entre les parties.

L'objet central de ce formalisme, qui lie les parties entre elles et qui encode leurs relations causales, est la matrice de processus (*process matrix*) [4]. Elle peut être comprise comme la généralisation

d'une matrice densité qui permet aussi la description des scénarios avec de la communication entre les parties. Les matrices de processus qui décrivent des scénarios incompatibles avec un ordre causal bien défini sont appelées causalement non-séparables (*causally nonseparable*) [4]. Ce choix de terminologie suggère qu'il y a une analogie avec le concept de l'intrication (à l'opposé de la séparabilité) des états quantiques. Effectivement, cette analogie existe, et elle s'étend plus loin. Certaines matrices de processus causalement non-séparables permettent aux parties d'établir des *corrélations non-causales*, et de violer des inégalités causales [4, 12], qui sont des contraintes similaires à des inégalités de Bell. De plus, la non-séparabilité causale peut être certifiée par des témoins de non-séparabilité causale [11, 13], similaires à des témoins d'intrication.

Ces concepts et outils mathématiques pour étudier les relations causales indéfinies ont été développés principalement pour le cas bipartite. L'un des objectifs de cette thèse est d'étendre ces résultats au cas multipartite.

Une autre question ayant une importance majeure est de trouver des réalisations physiques pour des processus avec un ordre causal indéfini. Nous ignorons si toutes les matrices de processus ont une réalisation ou une interprétation physique, et à ce jour, il n'y a qu'un seul exemple d'un processus causalement non-séparable qui a été étudié en profondeur d'un point de vue plus pratique. Il s'agit du *quantum switch* [3], un protocole dans lequel l'ordre de deux opérations quantiques est contrôlé par un qubit, qui est dans un état de superposition. Cela mène à une « superposition d'ordres », et on ne peut pas dire laquelle des deux opérations a été appliquée en premier. Des protocoles de ce type ont été réalisés expérimentalement [16, 17, 19] et leur non-séparabilité causale a été vérifiée dans ces expériences. De plus, il a été démontré que le *quantum switch* permet des avantages dans certaines tâches de traitement de l'information quantique, comparé à des circuits quantiques avec un ordre causal défini [23–29]. Un deuxième but de cette thèse est de trouver de nouveaux exemples de processus causalement non-séparables avec une interprétation physique.

Enfin, un troisième objectif de cette thèse est d'identifier de nouvelles applications en information quantique pour des processus causalement non-séparables qui sont réalisables en pratique.

La non-séparabilité causale dans des scénarios multipartites

La première étape vers des analyses multipartites est d'étendre la définition de la (non)-séparabilité causale à des situations multipartites. Cela est un problème complexe, notamment en raison des *ordres dynamiques* qui doivent être pris en compte en présence de plus de deux parties. Cela signifie que l'ordre causal des parties dans le futur peut changer en fonction des événements précédents. De plus, dans le formalisme des matrices de processus, les parties peuvent partager des états intriqués supplémentaires, ce qui implique des complexités additionnelles. Dans des travaux précédents, plusieurs définitions ont été proposées. Araújo *et al.* ont proposé une définition pour le cas restreint du *quantum switch*, afin d'étudier sa non-séparabilité causale [11]. Ensuite, Oreshkov et Giarmatzi ont considéré le cas multipartite général, notamment en prenant

en compte les ordres dynamiques, et ont proposé deux généralisations possibles [15]. Dans le chapitre 2 (qui se base sur la publication [32]), nous comparons ces différentes définitions dans le scénario tripartite particulier du *quantum switch*, et nous trouvons qu'elles ne sont pas toutes équivalentes, un point qui nécessite d'être clarifié. Nous proposons et justifions ensuite notre propre définition pour le cas multipartite général, et nous trouvons que notre définition est équivalente à celle de la « séparabilité causale extensible » (*extensible causal separability*) d'Oreshkov et Giarmatzi. Cette dernière évite l'« activation » de la non-séparabilité causale en fournissant des états intriqués additionnels aux parties. Nous établissons donc ce concept comme la généralisation naturelle de la séparabilité causale bipartite.

Dans la deuxième partie du chapitre 2, nous nous intéressons à comment caractériser l'ensemble des matrices de processus multipartites causalement séparables en termes de conditions simples. Nous déduisons une conditions nécessaire, ainsi qu'une condition suffisante pour la séparabilité causale multipartite. Dans les cas bipartite et tripartite, les deux conditions (nécessaire et suffisante) coïncident. Dans le cas de quatre parties et plus, la question reste ouverte. Ces conditions N-partites permettent toutefois de déterminer pour toute matrice de processus en dehors de cette zone grise potentielle, si elle est causalement séparable ou non en généralisant la technique des témoins de la non-séparabilité causale.

Circuits quantiques avec un ordre causal contrôlé de manière classique et quantique

Dans le chapitre 3, nous adoptons une approche constructive pour trouver de nouveaux exemples de réalisations concrètes de certains processus quantiques. Tout d'abord, nous considérons des circuits quantiques avec un ordre causal fixe, où les opérations sont appliquées l'une après l'autre, avec des transformations intermédiaires qui envoient l'état de sortie d'une opération vers l'entrée de la prochaine (tout en agissant sur des systèmes auxiliaires). Ces circuits ont été étudiés dans des travaux antérieurs, où ils ont été appelés « peignes quantiques » (*quantum combs*) [45, 56]. Les matrices de processus correspondantes sont compatibles avec un ordre fixe. Inversement, toute matrice de processus compatible avec un ordre fixe peut être réalisée en tant que peigne quantique, et le circuit peut être construit explicitement à partir de la matrice de processus [44, 45]. Ensuite, nous généralisons ces résultats en introduisant deux nouvelles classes de circuits. Premièrement, nous considérons des « circuits quantiques avec un ordre contrôlé de manière classique », dans lesquels les transformations intermédiaires sont des instruments quantiques avec plusieurs résultats classiques, qui déterminent la prochaine opération qui sera appliquée. Ces circuits restent compatibles avec un ordre causal bien défini (qui n'est cependant pas fixe a priori, mais établi graduellement pendant le déroulement du processus). Ces circuits correspondent aux matrices des processus qui satisfont la condition suffisante pour la séparabilité causale multipartite trouvée dans le chapitre 2. Sur cette base, nous introduisons ensuite des « circuits quantiques avec un ordre causal contrôlé de manière quantique » (QC-QC), dans lesquels l'ordre causal est contrôlé de manière cohérente. Cette classe contient le *quantum switch*,

mais aussi de nouveaux exemples de processus causalement non-séparables avec des nouvelles propriétés, qui sont réalisables en pratique. Par la suite, nous étendons notre analyse à des circuits « post-sélectionnés », qui peuvent eux-mêmes générer plusieurs résultats de mesure. De tels circuit définissent des ensembles de matrices de processus « probabilistes », appellés des « superinstruments quantiques » (*quantum superinstruments*). Nous trouvons que nos caractérisations des circuits, ainsi que des superinstruments correspondants, sont très utiles pour trouver des applications potentiels des processus causalement non-séparables avec une réalisation physique. En effet, elles permettent de comparer les différentes classes de circuits pour des tâches pratiques en utilisant des algorithmes d'optimisation semi-définie. Nous illustrons cela par un exemple.

Communication à travers des canaux quantiques contrôlés de manière cohérente

Dans le chapitre 4 (qui est basé sur la publication [34]), nous revenons sur un effet étrange qui a été montré dans un travail récent [29] : de l'information peut être transmise à travers certains canaux quantiques bruyants quand ils sont connectés dans le *quantum switch*, d'une manière qui n'est pas possible quand ils le sont dans une configuration causalement ordonnée. Plus particulièrement, dans un *quantum switch* avec deux canaux complètement dépolarisants, de l'information sur l'état initial du système sur lequel les canaux agissent peut être transmise à travers le switch. Nous montrons que cet effet est présent même dans une situation plus simple, dans un « demi-switch » où l'on applique un canal ou l'autre, en fonction d'un qubit de contrôle en superposition. Nous en concluons que ce phénomène est lié à l'utilisation d'un contrôle quantique, plutôt qu'à un ordre causal indéfini. L'analyse de cet exemple nous conduit à une analyse plus générale du concept d'une « opération quantique contrôlée de manière cohérente », et nous trouvons que cette notion est mal définie en général. Plus précisément, nous trouvons que, dans le scénario du demi-switch (dont la description n'entre pas dans le cadre des matrices de processus puisque les opérations ne sont pas appliquées une et une seule fois), l'état final dépend de la manière exacte dont les deux opérations sont effectuées. Cela signifie que la description habituelle d'une opération quantique en termes d'une application complètement positive préservant la trace n'est pas suffisante pour une description complète de la situation. Nous formalisons l'information supplémentaire qui est nécessaire pour compléter la description en introduisant des « matrices de transformation » (*transformation matrices*).

Des corrélations non-causales « véritablement multipartites »

Une autre question qui se pose dans les scénarios multipartites est la suivante : un phénomène donné est-il « véritablement multipartite » (*genuinely multipartite*) ? Autrement dit, nécessite-t-il vraiment la coopération d'un certain nombre de parties, ou peut-il être réduit à un phénomène

qui se produit entre les parties d'un sous-ensemble ? Dans le chapitre 5 (qui est basé sur la publication [35]), nous définissions les corrélations véritablement multipartites comme celles qui ne sont pas « 2-causales », c'est-à-dire pour lesquelles il n'est pas possible d'identifier un ordre causal « à une plus grande échelle » même en considérant des groupements des parties en deux sous-ensembles (ou des mélanges probabilistes de tels groupements). Ensuite, nous caractérisons l'ensemble des corrélations 2-causales, et nous dérivons des inégalités 2-causales dans le scénario le plus simple où le concept a du sens.

Des valeurs faibles anormales en absence de post-sélection

Dans le chapitre 6 (qui est basé sur la publication [36]), le résultat présenté a été obtenu en étudiant le lien entre le formalisme des matrices de processus et un autre formalisme qui étudie le rôle du temps dans la théorie quantique, celui des « états pré- et post- sélectionnés » (*pre-and post-selected quantum states*) ou « états quantiques à deux temps » (*two-time states*) [5–7]. Dans le cadre de ces systèmes pré-et post-sélectionnés, on étudie souvent des « mesures faibles ». Par cela, on entend des mesures qui ne perturbent que très faiblement le système mesuré, mais qui dévoilent très peu d'information sur celui-ci. Un concept lié à celui de la mesure faible est la « valeur faible » d'une observable, qui est, dans un certain sens, une quantité analogue à la valeur moyenne. Il est connu que la valeur faible peut être en dehors du spectre de l'observable, ce qui est appelé une « valeur faible anormale » (*anomalous weak value*). En général, cet effet est associé à la présence d'une post-sélection non triviale. Dans le chapitre 6, nous montrons que ce n'est pas forcément le cas : dans des scénarios avec des mesures faibles séquentielles, on peut obtenir des valeurs faibles anormales sans post-sélection.

Des processus non-causaux sur des sous-systèmes temporellement délocalisés

Dans le chapitre 7 un autre travail en cours est présenté. Les protocoles expérimentaux relatifs au *quantum switch* ont été accompagnés par un débat portant sur leur interprétation comme véritables « réalisations » d'un processus avec un ordre causal indéfini, ou alors plutôt comme des « simulations » d'un tel processus [37–39]. En particulier, la question qui a été soulevée est : les opérations sont-elles appliquées une et une seule fois dans ces implémentations, et sur quels systèmes agissent-t-elles ? Dans [38], il a été montré que ces expériences peuvent effectivement être vues comme des réalisations, en introduisant le concept des « sous-systèmes temporellement délocalisés » (*time-delocalised subsystems*)—des sous-systèmes quantiques qui ne sont pas associés avec un temps bien défini.
Dans le chapitre 7, nous montrons dans un premier temps que les arguments présentés dans [38] s'appliquent aussi aux circuits quantiques avec un contrôle d'ordre quantique (chapitre 3), justifiant ainsi leur interprétation en tant que « réalisations ». Dans [38], il a ensuite été montré

qu'une classe particulière de matrices de processus bipartites, celles qui sont « unitairement extensibles », ont également une réalisation sur des sous-systèmes temporellement délocalisés. Bien que cette classe soit a priori plus grande que celle des QC-QC, il a en fait été montré que dans le cas de deux parties, les deux classes coïncident. Sur la base des résultats du chapitre 3, nous généralisons la preuve de [38] aux processus unitairement extensibles avec trois parties, une classe qui est strictement plus grande que celle des QC-QC avec trois opérations, et qui, en particulier, contient des processus capables de violer des inégalités causales.

Appendix A

Quantum circuits with quantum control of causal order cannot violate causal inequalities

In this Appendix we prove Proposition 3.10, that QC-QCs (and *a fortiori*, QC-CCs or QC-FOs) can only generate causal correlations.

As we consider probabilities rather than induced maps here, we will take the global past and future Hilbert spaces P and F to be trivial[1]. For reference, let us explicitly give the characterisation for that case:

Proposition A.1. *The process matrix $W \in \mathcal{L}(\mathcal{H}^{A^{IO}_{\mathcal{N}}})$ of a quantum circuit with quantum control of causal order is such that there exist positive semidefinite matrices $W_{(\mathcal{K},\ell)} \in \mathcal{L}(\mathcal{H}^{A^{IO}_{\mathcal{K}} A^{I}_{\ell}})$, for all strict subsets $\mathcal{K}$ of $\mathcal{N}$ and all $\ell \in \mathcal{N}\backslash\mathcal{K}$, satisfying*

$$\begin{aligned} &W = \sum_{k\in\mathcal{N}} W_{(\mathcal{N}\backslash\{k\},k)} \otimes \mathbb{1}^{A^{O}_{k}}, \\ &\forall\, \emptyset \subsetneq \mathcal{K} \subsetneq \mathcal{N},\ \sum_{\ell\in\mathcal{N}\backslash\mathcal{K}} \mathrm{Tr}_{A^{I}_{\ell}} W_{(\mathcal{K},\ell)} = \sum_{k\in\mathcal{K}} W_{(\mathcal{K}\backslash\{k\},k)} \otimes \mathbb{1}^{A^{O}_{k}}, \\ &\text{and} \quad \sum_{\ell\in\mathcal{N}} \mathrm{Tr}\, W_{(\emptyset,\ell)} = 1. \end{aligned} \tag{A.1}$$

Conversely, any Hermitian matrix $W \in \mathcal{L}(\mathcal{H}^{A^{IO}_{\mathcal{N}}})$ such that there exist positive semidefinite matrices $W_{(\mathcal{K},\ell)} \in \mathcal{L}(\mathcal{H}^{A^{IO}_{\mathcal{K}} A^{I}_{\ell}})$ for all $\mathcal{K} \subsetneq \mathcal{N}$ and $\ell \in \mathcal{N}\backslash\mathcal{K}$ satisfying Eq. (A.1) is the process matrix of a quantum circuit with quantum control of causal order.

[1] Without loss of generality of course—any process matrix with non-trivial global past and future spaces can equivalently be seen as one with trivial global past and future spaces and two additional parties.

Here, consider again N "parties" that apply quantum instruments; the party A_k's instruments, corresponding to some setting ("input") x_k are denoted (in their CJ representations) $\{M_{a_k|x_k}\}_{a_k}$, where each a_k corresponds to one possible outcome ("output") of the instrument.

For any subset $\mathcal{K} = \{k_1, \ldots, k_n\}$ of $\mathcal{N}$, we shall denote by $\vec{x}_\mathcal{K} := (x_{k_1}, \ldots, x_{k_n})$ and $\vec{a}_\mathcal{K} := (a_{k_1}, \ldots, a_{k_n})$ the list of inputs and outputs for the parties in $\mathcal{K}$, and by $M_{\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}} := \bigotimes_{k\in\mathcal{K}} M_{a_k|x_k} \in \mathcal{L}(\mathcal{H}^{A_\mathcal{K}^{IO}})$ the corresponding joint operations (in their Choi representation). With these notations, the correlations $P(\vec{a}_\mathcal{N}|\vec{x}_\mathcal{N})$ obtained from a QC-QC are given, as in Eq. (1.31) (and for all $\vec{x}_\mathcal{N}, \vec{a}_\mathcal{N}$), by

$$P(\vec{a}_\mathcal{N}|\vec{x}_\mathcal{N}) = \mathrm{Tr}\left[W\, M^T_{\vec{a}_\mathcal{N}|\vec{x}_\mathcal{N}}\right], \tag{A.2}$$

with W satisfying the constraints of Proposition A.1, i.e., such that there exist positive semidefinite matrices $W_{(\mathcal{K},\ell)}$ satisfying Eq. (A.1).[2]

Let us introduce for convenience the following notations, for any strict subset $\mathcal{K}$ of $\mathcal{N}$ and any $\ell \in \mathcal{N}\backslash\mathcal{K}$:

$$\begin{aligned} r_{(\mathcal{K},\ell)}(\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}) &:= \mathrm{Tr}\left[\left(\mathrm{Tr}_{A_\ell^I} W_{(\mathcal{K},\ell)}\right) M^T_{\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}}\right], \\ s_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K},\ell}|\vec{x}_{\mathcal{K},\ell}) &:= \mathrm{Tr}\left[(W_{(\mathcal{K},\ell)} \otimes \mathbb{1}^{A_\ell^O}) M^T_{\vec{a}_{\mathcal{K},\ell}|\vec{x}_{\mathcal{K},\ell}}\right], \end{aligned} \tag{A.3}$$

with the first definition simplifying to $r_{(\emptyset,\ell)} := r_{(\emptyset,\ell)}(\vec{a}_\emptyset|\vec{x}_\emptyset) := \mathrm{Tr}[W_{(\emptyset,\ell)}]$ for $\mathcal{K} = \emptyset$, and with $\vec{x}_{\mathcal{K},\ell} := \vec{x}_{\mathcal{K}\cup\{\ell\}}$ and $\vec{a}_{\mathcal{K},\ell} := \vec{a}_{\mathcal{K}\cup\{\ell\}}$ in the second definition. We note that $r_{(\mathcal{K},\ell)}(\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K})$ and $s_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K},\ell}|\vec{x}_{\mathcal{K},\ell})$ are nonnegative functions of the inputs and outputs of the parties in $\mathcal{K}$ and in $\mathcal{K} \cup \{\ell\}$, respectively, but are not valid probability distributions in general, since neither $\mathrm{Tr}_{A_\ell^I} W_{(\mathcal{K},\ell)}$ nor $W_{(\mathcal{K},\ell)} \otimes \mathbb{1}^{A_\ell^O}$ are valid process matrices in general. Nevertheless, using the fact that the sum $\sum_{a_\ell} M_{a_\ell|x_\ell}$ is a CPTP map, i.e., that $\mathrm{Tr}_{A_\ell^O} \sum_{a_\ell} M_{a_\ell|x_\ell} = \mathbb{1}^{A_\ell^I}$, one finds that

$$\begin{aligned} &\sum_{a_\ell} s_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K},\ell}|\vec{x}_{\mathcal{K},\ell}) \\ &\quad = \mathrm{Tr}\left[(W_{(\mathcal{K},\ell)} \otimes \mathbb{1}^{A_\ell^O})(M^T_{\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}} \otimes \sum_{a_\ell} M_{a_\ell|x_\ell})\right] \\ &\quad = \mathrm{Tr}\left[W_{(\mathcal{K},\ell)}\left(M^T_{\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}} \otimes \mathbb{1}^{A_\ell^I}\right)\right] \\ &\quad = \mathrm{Tr}\left[\left(\mathrm{Tr}_{A_\ell^I} W_{(\mathcal{K},\ell)}\right) M^T_{\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}}\right] = r_{(\mathcal{K},\ell)}(\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}). \end{aligned} \tag{A.4}$$

It follows that one can define (for all $\emptyset \subseteq \mathcal{K} \subsetneq \mathcal{N}$, all $\ell \in \mathcal{N}\backslash\mathcal{K}$ and all $\vec{x}_\mathcal{K}, \vec{a}_\mathcal{K}$) a valid conditional probability distribution $P^{(\ell)}_{\mathcal{K},\vec{x}_\mathcal{K},\vec{a}_\mathcal{K}}(a_\ell|x_\ell)$ for party A_ℓ—which, as indicated by the subscript, depends on $\mathcal{K}$, $\vec{x}_\mathcal{K}$ and $\vec{a}_\mathcal{K}$—such that

$$s_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K},\ell}|\vec{x}_{\mathcal{K},\ell}) = P^{(\ell)}_{\mathcal{K},\vec{x}_\mathcal{K},\vec{a}_\mathcal{K}}(a_\ell|x_\ell)\; r_{(\mathcal{K},\ell)}(\vec{a}_\mathcal{K}|\vec{x}_\mathcal{K}). \tag{A.5}$$

[2] As is straightforward to see, the set of QC-QC process matrices is closed under extension with shared ancillary systems. We can thus consider Eq. (A.2) without additional ancillas without loss of generality.

One can further easily see that the functions $r_{(\mathcal{K},\ell)}$ and $s_{(\mathcal{K},\ell)}$ inherit the following properties from Eq. (A.1):

$$\begin{aligned}
&P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}}) = \sum_{k\in\mathcal{N}} s_{(\mathcal{N}\setminus\{k\},k)}(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}}),\\
&\forall\, \emptyset\subsetneq\mathcal{K}\subsetneq\mathcal{N},\ \sum_{\ell\in\mathcal{N}\setminus\mathcal{K}} r_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}}) = \sum_{k\in\mathcal{K}} s_{(\mathcal{K}\setminus\{k\},k)}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}}),\\
&\text{and}\quad \sum_{\ell\in\mathcal{N}} r_{(\emptyset,\ell)} = 1.
\end{aligned} \tag{A.6}$$

With this in place, we now proceed by recursively proving the following relation for the correlations $P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}})$:

Proposition A.2. *For any n with $0 \le n \le N-1$, one can write the correlations (A.2) as*

$$P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}}) = \sum_{\substack{\mathcal{K}\subsetneq\mathcal{N}:\\ |\mathcal{K}|=n}} P^{c}_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}}) \sum_{\ell\in\mathcal{N}\setminus\mathcal{K}} r_{(\mathcal{K},\ell)}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}}), \tag{A.7}$$

where the $P^{c}_{\mathcal{K},\vec{a}_{\mathcal{K}},\vec{x}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}})$'s are causal correlations for the parties in $\mathcal{N}\setminus\mathcal{K}$.

Proof. For $n = N-1$, the first line of Eq. (A.6), combined with Eq. (A.5), gives

$$\begin{aligned}
P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}}) = \sum_{k\in\mathcal{N}} & P^{(k)}_{\mathcal{N}\setminus\{k\},\vec{x}_{\mathcal{N}\setminus\{k\}},\vec{a}_{\mathcal{N}\setminus\{k\}}}(a_k|x_k)\\
& r_{(\mathcal{N}\setminus\{k\},k)}(\vec{a}_{\mathcal{N}\setminus\{k\}}|\vec{x}_{\mathcal{N}\setminus\{k\}}),
\end{aligned} \tag{A.8}$$

which is clearly of the form of Eq. (A.7), with $\mathcal{K} = \mathcal{N}\setminus\{k\}$, $\ell = k$, and with $P^{c}_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}} = P^{(k)}_{\mathcal{N}\setminus\{k\},\vec{x}_{\mathcal{N}\setminus\{k\}},\vec{a}_{\mathcal{N}\setminus\{k\}}}$, which, as a single-partite probability distribution, is trivially causal.

Suppose now that Eq. (A.7) holds for some given $n \le N-1$, with some causal correlations $P^{c}_{\mathcal{K},\vec{a}_{\mathcal{K}},\vec{x}_{\mathcal{K}}}$. Then if $n > 0$, by using the second line of Eq. (A.6), by replacing the double sum $\sum_{\mathcal{K}\subsetneq\mathcal{N},|\mathcal{K}|=n}\sum_{k\in\mathcal{K}}$ by the equivalent double sum $\sum_{\mathcal{K}'\subsetneq\mathcal{N},|\mathcal{K}'|=n-1}\sum_{k\in\mathcal{N}\setminus\mathcal{K}'}$ (with $\mathcal{K}' = \mathcal{K}\setminus\{k\}$), and by then using Eq. (A.5), we obtain

$$\begin{aligned}
&P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}})\\
&= \sum_{\substack{\mathcal{K}\subsetneq\mathcal{N}:\\ |\mathcal{K}|=n}} P^{c}_{\mathcal{K},\vec{x}_{\mathcal{K}},\vec{a}_{\mathcal{K}}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}}) \sum_{k\in\mathcal{K}} s_{(\mathcal{K}\setminus\{k\},k)}(\vec{a}_{\mathcal{K}}|\vec{x}_{\mathcal{K}})\\
&= \sum_{\substack{\mathcal{K}'\subsetneq\mathcal{N}:\\ |\mathcal{K}'|=n-1}} \sum_{k\in\mathcal{N}\setminus\mathcal{K}'} P^{c}_{\mathcal{K}'\cup\{k\},\vec{x}_{\mathcal{K}',k},\vec{a}_{\mathcal{K}',k}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}})\\
&\qquad\qquad s_{(\mathcal{K}',k)}(\vec{a}_{\mathcal{K}',k}|\vec{x}_{\mathcal{K}',k})\\
&= \sum_{\substack{\mathcal{K}'\subsetneq\mathcal{N}:\\ |\mathcal{K}'|=n-1}} \sum_{k\in\mathcal{N}\setminus\mathcal{K}'} P^{c}_{\mathcal{K}'\cup\{k\},\vec{x}_{\mathcal{K}',k},\vec{a}_{\mathcal{K}',k}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}})\\
&\qquad\qquad P^{(k)}_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(a_k|x_k)\ r_{(\mathcal{K}',k)}(\vec{a}_{\mathcal{K}'}|\vec{x}_{\mathcal{K}'}).
\end{aligned} \tag{A.9}$$

Let us now define, for each $\mathcal{K}', \vec{x}_{\mathcal{K}'}, \vec{a}_{\mathcal{K}'}$, the weights

$$q_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(k) := \frac{r_{(\mathcal{K}',k)}(\vec{a}_{\mathcal{K}'}|\vec{x}_{\mathcal{K}'})}{\sum_{\ell'\in\mathcal{N}\setminus\mathcal{K}'} r_{(\mathcal{K}',\ell')}(\vec{a}_{\mathcal{K}'}|\vec{x}_{\mathcal{K}'})} \quad \forall\, k\in\mathcal{N}\setminus\mathcal{K}' \tag{A.10}$$

(such that $q_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(k) \geq 0$, $\sum_k q_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(k) = 1$; when the denominator above is zero, we can take arbitrary weights) and the correlations

$$\begin{aligned} &P^c_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}'}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}'}) \\ &:= \sum_{k\in\mathcal{N}\setminus\mathcal{K}'} q_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(k)\; P^{(k)}_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(a_k|x_k) \\ &\qquad\qquad P^c_{\mathcal{K}'\cup\{k\},\vec{x}_{\mathcal{K}',k},\vec{a}_{\mathcal{K}',k}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}'\setminus\{k\}}) \end{aligned} \tag{A.11}$$

for the parties in $\mathcal{N}\setminus\mathcal{K}'$. By construction, the correlations $P^c_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}$ are convex mixtures (with weights $q_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(k)$) of correlations compatible with a given party $k \in \mathcal{N}\setminus\mathcal{K}'$ acting first (with a response function $P^{(k)}_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(a_k|x_k)$ which does not depend on the inputs of the other parties in $\mathcal{N}\setminus\mathcal{K}'$) and such that whatever that first party does, the conditional correlations shared by the other parties are causal. In other words: this means that the correlations $P^c_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}$ are causal [14, 15]. In terms of these, Eq. (A.9) can be written as

$$P(\vec{a}_{\mathcal{N}}|\vec{x}_{\mathcal{N}}) = \sum_{\substack{\mathcal{K}'\subsetneq\mathcal{N}:\\ |\mathcal{K}'|=n-1}} P^c_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}(\vec{a}_{\mathcal{N}\setminus\mathcal{K}'}|\vec{x}_{\mathcal{N}\setminus\mathcal{K}'}) \sum_{\ell'\in\mathcal{N}\setminus\mathcal{K}'} r_{(\mathcal{K}',\ell')}(\vec{a}_{\mathcal{K}'}|\vec{x}_{\mathcal{K}'}), \tag{A.12}$$

which is of the form of Eq. (A.7). Hence, Eq. (A.7) also holds with $|\mathcal{K}'| = n-1$ instead of $|\mathcal{K}| = n$ (and with some causal correlations $P^c_{\mathcal{K}',\vec{x}_{\mathcal{K}'},\vec{a}_{\mathcal{K}'}}$), which by recursion proves that it indeed holds for all n such that $0 \leq n \leq N-1$, as claimed in Proposition A.2. □

To conclude the proof of Proposition 3.10, it then suffices to notice (using the fact that $\sum_{\ell\in\mathcal{N}} r_{(\emptyset,\ell)} = 1$) that the latter simply corresponds to the case $n = 0$ of Proposition A.2 above.

Bibliography

[1] M. Zych, F. Costa, I. Pikovski, and Č. Brukner. Bell's theorem for temporal order. *Nat. Commun.*, 10:3772, 2019. URL https://doi.org/10.1038/s41467-019-11579-x.

[2] L. Hardy. Probability theories with dynamic causal structure: a new framework for quantum gravity. *arXiv:0509120 [gr-qc]*, 2005. URL https://arxiv.org/abs/gr-qc/0509120.

[3] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron. Quantum computations without definite causal structure. *Phys. Rev. A*, 88:022318, 2013. URL http://dx.doi.org/10.1103/PhysRevA.88.022318.

[4] O. Oreshkov, F. Costa, and Č. Brukner. Quantum correlations with no causal order. *Nat. Commun.*, 3:1092, 2012. URL http://dx.doi.org/10.1038/ncomms2076.

[5] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman. Multiple-time states and multiple-time measurements in quantum mechanics. *Phys. Rev. A*, 79:052110, 2009. URL https://link.aps.org/doi/10.1103/PhysRevA.79.052110.

[6] R. Silva, Y. Guryanova, N. Brunner, N. Linden, A. J. Short, and S. Popescu. Pre- and postselected quantum states: Density matrices, tomography, and kraus operators. *Phys. Rev. A*, 89:012121, 2014. URL https://link.aps.org/doi/10.1103/PhysRevA.89.012121.

[7] R. Silva, Y. Guryanova, A. J. Short, P. Skrzypczyk, N. Brunner, and S. Popescu. Connecting processes with indefinite causal order and multi-time quantum states. *New J. Phys.*, 19 (10):103022, 2017. URL https://doi.org/10.1088%2F1367-2630%2Faa84fe.

[8] O. Oreshkov and N. J. Cerf. Operational quantum theory without predefined time. *New J. Phys.*, 18(7):073037, 2016. URL https://doi.org/10.1088%2F1367-2630%2F18%2F7%2F073037.

[9] J. Barrett, R. Lorenz, and O. Oreshkov. Quantum causal models. *arXiv:1906.10726 [quant-ph]*, 2019. URL https://arxiv.org/abs/1906.10726.

[10] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens. Quantum common causes and quantum causal models. *Phys. Rev. X*, 7:031021, 2017. URL https://doi.org/10.1103/PhysRevX.7.031021.

[11] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner. Witnessing causal nonseparability. *New J. Phys.*, 17:102001, 2015. URL http://dx.doi.org/10.1088/1367-2630/17/10/102001.

[12] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner. The simplest causal inequalities and their violation. *New J. Phys.*, 18:013008, 2016. URL http://dx.doi.org/10.1088/1367-2630/18/1/013008.

[13] C. Branciard. Witnesses of causal nonseparability: an introduction and a few case studies. *Sci. Rep.*, 6:26018, 2016. URL http://dx.doi.org/10.1038/srep26018.

[14] A. A. Abbott, C. Giarmatzi, F. Costa, and C. Branciard. Multipartite causal correlations: Polytopes and inequalities. *Phys. Rev. A*, 94:032131, 2016. URL http://dx.doi.org/10.1103/PhysRevA.94.032131.

[15] O. Oreshkov and C. Giarmatzi. Causal and causally separable processes. *New J. Phys.*, 18 (9):093020, 2016. URL http://dx.doi.org/10.1088/1367-2630/18/9/093020.

[16] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther. Experimental superposition of orders of quantum gates. *Nat. Commun.*, 6:7913, 2015. URL http://dx.doi.org/10.1038/ncomms8913.

[17] 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. *Sci. Adv.*, 3(3): e1602589, 2017. URL http://dx.doi.org/10.1126/sciadv.1602589.

[18] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther. Experimental entanglement of temporal orders. *arXiv:1712.06884 [quant-ph]*, 2017. URL https://arxiv.org/abs/1712.06884.

[19] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White. Indefinite causal order in a quantum switch. *Phys. Rev. Lett.*, 121:090503, 2018. URL http://dx.doi.org/10.1103/PhysRevLett.121.090503.

[20] K. Wei, N. Tischler, S. Zhao, Y. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J. Pan. Experimental quantum switching for exponentially superior quantum communication complexity. *Phys. Rev. Lett.*, 122:120504, 2019. URL http://dx.doi.org/10.1103/PhysRevLett.122.120504.

[21] Y. Guo, X. Hu, Z. Hou, H. Cao, J. Cui, B. Liu, Y. Huang, C. Li, and G. Guo. Experimental investigating communication in a superposition of causal orders. *arXiv:1811.07526 [quant-ph]*, 2018. URL https://arxiv.org/abs/1811.07526.

[22] K. Goswami, Y. Cao, G.A. Paz-Silva, J. Romero, and A. G. White. Communicating via ignorance. *arXiv:1807.07383 [quant-ph]*, 2018. URL https://arxiv.org/abs/1807.07383.

[23] G. Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Phys. Rev. A*, 86:040301, 2012. URL http://dx.doi.org/10.1103/PhysRevA.86.040301.

[24] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti. Quantum computation with programmable connections between gates. *Phys. Lett. A*, 376(45):2940–2943, 2012. URL http://dx.doi.org/10.1016/j.physleta.2012.08.028.

[25] M. Araújo, F. Costa, and Č. Brukner. Computational Advantage from Quantum-Controlled Ordering of Gates. *Phys. Rev. Lett.*, 113(25):250402, 2014. URL http://dx.doi.org/10.1103/PhysRevLett.113.250402.

[26] S. Facchini and S. Perdrix. Quantum circuits for the unitary permutation problem. In Rahul Jain, Sanjay Jain, and Frank Stephan, editors, *Theory and Applications of Models of Computation*, pages 324–331, Cham, 2015. Springer International Publishing. ISBN 978-3-319-17142-5. URL https://doi.org/10.1007/978-3-319-17142-5_28.

[27] A. Feix, M. Araújo, and Č. Brukner. Quantum superposition of the order of parties as a communication resource. *Phys. Rev. A*, 92:052326, 2015. URL http://dx.doi.org/10.1103/PhysRevA.92.052326.

[28] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Phys. Rev. Lett.*, 117:100502, 2016. URL http://dx.doi.org/10.1103/PhysRevLett.117.100502.

[29] D. Ebler, S. Salek, and G. Chiribella. Enhanced communication with the assistance of indefinite causal order. *Phys. Rev. Lett.*, 120:120502, 2018. URL http://dx.doi.org/10.1103/PhysRevLett.120.120502.

[30] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, NY, USA, 2011. ISBN 1107002176, 9781107002173. URL https://doi.org/10.1017/CBO9780511976667.

[31] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013. URL http://dx.doi.org/10.1017/CBO9781139525343.

[32] J. Wechs, A. A. Abbott, and C. Branciard. On the definition and characterisation of multipartite causal (non)separability. *New J. Phys.*, 21:013027, 2019. URL http://dx.doi.org/10.1088/1367-2630/aaf352.

[33] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard. Quantum circuits with classical versus quantum control of causal orders. in preparation, 2020.

[34] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard. Communication through coherent control of quantum channels. *arXiv:1810.09826 [quant-ph]*, 2018. URL https://arxiv.org/abs/1810.09826.

[35] A. A. Abbott, J. Wechs, F. Costa, and C. Branciard. Genuinely Multipartite Noncausality. *Quantum*, 1:39, 2017. URL https://doi.org/10.22331/q-2017-12-14-39.

[36] A. A. Abbott, R. Silva, J. Wechs, N. Brunner, and C. Branciard. Anomalous Weak Values Without Post-Selection. *Quantum*, 3:194, 2019. URL https://doi.org/10.22331/q-2019-10-14-194.

[37] J-P. W. MacLean, K. Ried, R. W. Spekkens, and K. J. Resch. Quantum-coherent mixtures of causal relations. *Nat. Commun.*, 8(1):15149, 2017. URL https://doi.org/10.1038/ncomms15149.

[38] O. Oreshkov. Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics. *Quantum*, 3:206, December 2019. URL https://doi.org/10.22331/q-2019-12-02-206.

[39] N. Paunkovic and M. Vojinovic. Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders. *arXiv:1905.09682 [quant-ph]*, 2019. URL https://arxiv.org/abs/1905.09682.

[40] F. Giacomini, E. Castro-Ruiz, and Č. Brukner. Indefinite causal structures for continuous-variable systems. *New J. Phys.*, 18(11):113026, 2016. URL http://stacks.iop.org/1367-2630/18/i=11/a=113026.

[41] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Commun. Math. Phys.*, 17(3):239–260, 1970. URL http://dx.doi.org/10.1007/BF01647093.

[42] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3:275–278, 1972. URL http://dx.doi.org/10.1016/0034-4877(72)90011-0.

[43] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra Appl.*, 10:285–290, 1975. URL http://dx.doi.org/10.1016/0024-3795(75)90075-0.

[44] G. Gutoski and J. Watrous. Toward a general theory of quantum games. In *Proceedings of 39th ACM STOC*, pages 565–574, 2006. URL https://doi.org/10.1145/1250790.1250873.

[45] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Theoretical framework for quantum networks. *Phys. Rev. A*, 80(2):022339, 2009. URL http://dx.doi.org/10.1103/PhysRevA.80.022339.

[46] D. Kretschmann and R. F. Werner. Quantum channels with memory. *Phys. Rev. A*, 72:062323, 2005. URL https://link.aps.org/doi/10.1103/PhysRevA.72.062323.

[47] Ä. Baumeler and S. Wolf. Perfect signaling among three parties violating predefined causal order. In *2014 IEEE Int. Symp. Inf. Theory (ISIT)*, pages 526–530, June 2014. doi: 10.1109/ISIT.2014.6874888. URL http://dx.doi.org/10.1109/ISIT.2014.6874888.

[48] S. Salek, D. Ebler, and G. Chiribella. Quantum communication in a superposition of causal orders. *arXiv:1809.06655 [quant-ph]*, 2018. URL https://arxiv.org/abs/1809.06655.

[49] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar. Indefinite causal order enables perfect quantum communication with zero capacity channels. *arXiv:1810.10457 [quant-ph]*, 2018. URL https://arxiv.org/abs/1810.10457.

[50] Ognyan Oreshkov. private communication, 2016.

[51] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[52] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera. Classification of mixed three-qubit states. *Phys. Rev. Lett.*, 87:040401, 2001. URL http://dx.doi.org/10.1103/PhysRevLett.87.040401.

[53] G. Svetlichny. Distinguishing three-body from two-body nonseparability by a Bell-type inequality. *Phys. Rev. D*, 35:3066, 1987. URL http://dx.doi.org/10.1103/PhysRevD.35.3066.

[54] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués. Operational framework for nonlocality. *Phys. Rev. Lett.*, 109:070401, 2012. URL http://dx.doi.org/10.1103/PhysRevLett.109.070401.

[55] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio. Definitions of multipartite nonlocality. *Phys. Rev. A*, 88:014102, 2013. URL http://dx.doi.org/10.1103/PhysRevA.88.014102.

[56] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Quantum circuit architecture. *Phys. Rev. Lett.*, 101:060401, 2008. URL https://doi.org/10.1103/PhysRevLett.101.060401.

[57] G. Chiribella and D. Ebler. Optimal quantum networks and one-shot entropies. *New J. Phys.*, 18(9):093053, 2016. URL http://stacks.iop.org/1367-2630/18/i=9/a=093053.

[58] A. Shimbo, A. Soeda, and M. Murao. Equivalence determination of unitary operations. *arXiv:1803.11414 [quant-ph]*, 2018. URL https://arxiv.org/abs/1803.11414.

[59] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Transforming quantum operations: Quantum supermaps. *EPL*, 83(3):30004, 2008. URL https://doi.org/10.1209/0295-5075/83/30004.

[60] M. Araújo, A. Feix, M. Navascués, and Č. Brukner. A purification postulate for quantum mechanics with indefinite causal order. *Quantum*, 1:10, 2017. URL https://doi.org/10.22331/q-2017-04-26-10.

[61] A. Feix, M. Araújo, and Č. Brukner. Causally nonseparable processes admitting a causal model. *New. J. Phys.*, 18(8):083040, 2016. URL http://stacks.iop.org/1367-2630/18/i=8/a=083040.

[62] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao. Probabilistic exact universal quantum circuits for transforming unitary operations. *Phys. Rev. A*, 100:062339, 2019. URL https://doi.org/10.1103/PhysRevA.100.062339.

[63] N. Johnston. QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9. http://qetlab.com, 2016.

[64] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao. Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations. *Phys. Rev. Lett.*, 123:210502, 2019. URL https://doi.org/10.1103/PhysRevLett.123.210502.

[65] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti. Optimal quantum learning of a unitary transformation. *Phys. Rev. A*, 81:032324, 2010. URL https://doi.org/10.1103/PhysRevA.81.032324.

[66] M. Sedlák, A. Bisio, and M. Ziman. Optimal probabilistic storage and retrieval of unitary channels. *Phys. Rev. Lett.*, 122:170502, 2019. URL https://doi.org/10.1103/PhysRevLett.122.170502.

[67] Ä. Baumeler and S. Wolf. The space of logically consistent classical processes without causal order. *New. J. Phys.*, 18(1):013036, 2016. URL https://doi.org/10.1088/1367-2630/18/1/013036.

[68] K. Kraus. *States, Effects, and Operations: Fundamental Notions of Quantum Theory*. Springer-Verlag, Berlin Heidelberg, 1983. URL https://doi.org/10.1007/3-540-12732-1.

[69] L. M. Procopio, F. Delgado, M. Enriquez, N. Belabas, and J. A. Levenson. Communication enhancement through quantum coherent control of N channels in an indefinite causal-order scenario. *Entropy*, 21(10):1012, 2019. URL https://doi.org/10.3390/e21101012.

[70] M. Araújo, A. Feix, F. Costa, and Č. Brukner. Quantum circuits cannot control unknown operations. *New J. Phys.*, 16:093026, 2014. URL http://dx.doi.org/10.1088/1367-2630/16/9/093026.

[71] N. Friis, V. Dunjko, W. Dür, and H. J. Briegel. Implementing quantum control for unknown subroutines. *Phys. Rev. A*, 89:030303(R), 2014. URL https://doi.org/10.1103/PhysRevA.89.030303.

[72] T. M. Rambo, J. B. Altepeter, P. Kumar, and G. M. D'Ariano. Functional quantum computing: An optical approach. *Phys. Rev. A*, 93:052321, 2016. URL https://doi.org/10.1103/PhysRevA.93.052321.

[73] J. Thompson, K. Modi, V. Vedral, and M. Gu. Quantum plug n' play: modular computation in the quantum regime. *New J. Phys.*, 20:013004, 2018. URL https://doi.org/10.1088/1367-2630/aa99b3.

[74] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White. Simplifying quantum logic using higher-dimensional Hilbert spaces. *Nat. Phys.*, 5:134–140, 2009. URL https://doi.org/10.1038/nphys1150.

[75] X. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien. Adding control to arbitrary unknown quantum operations. *Nat. Commun.*, 2:413, 08 2011. URL http://dx.doi.org/10.1038/ncomms1392.

[76] X. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien. Calculating unknown eigenvalues with a quantum algorithm. *Nat. Photonics*, 7:223, 02 2013. URL http://dx.doi.org/10.1038/nphoton.2012.360.

[77] N. Friis, A. A. Melnikov, G. Kirchmair, and H. J. Briegel. Coherent controlization using superconducting qubits. *Sci. Rep.*, 5:18036, 12 2015. URL http://dx.doi.org/10.1038/srep18036.

[78] V. Dunjko, N. Friis, and H. J. Briegel. Quantum-enhanced deliberation of learning agents using trapped ions. *New J. Phys.*, 17:023006, 2015. URL http://dx.doi.org/10.1088/1367-2630/17/2/023006.

[79] P. A. Guérin, G. Rubino, and Č. Brukner. Communication through quantum-controlled noise. *Phys. Rev. A*, 99:062317, 2019. URL https://doi.org/10.1103/PhysRevA.99.062317.

[80] A. Bisio, M. Dall'Arno, and P. Perinotti. Quantum conditional operations. *Phys. Rev. A*, 94:022340, 2016. URL https://doi.org/10.1103/PhysRevA.94.022340.

[81] G. Chiribella and H. Kristjánsson. Quantum shannon theory with superpositions of trajectories. *Proc. R. Soc. A*, 475:20180903, 2019. URL http://dx.doi.org/10.1098/rspa.2018.0903.

[82] W. F. Stinespring. Positive functions on c^*-algebras. *Proc. Amer. Math. Soc.*, 6:211–216, 1955. URL http://dx.doi.org/10.1090/S0002-9939-1955-0069403-4.

[83] J. Åberg. Subspace preservation, subspace locality, and gluing of completely positive maps. *Ann. Phys.*, 313(2):326 – 367, 2004. URL https://doi.org/10.1016/j.aop.2004.04.013.

[84] Daniel K. L. Oi. Interference of quantum channels. *Phys. Rev. Lett.*, 91:067902, 2003. URL http://dx.doi.org/10.1103/PhysRevLett.91.067902.

[85] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, Cambridge, 2018. URL https://doi.org/10.1017/9781316848142.

[86] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131, 1997. URL https://doi.org/10.1103/PhysRevA.56.131.

[87] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44(1):269, 1998. URL https://doi.org/10.1109/18.651037.

[88] N. Gisin, N. Linden, S. Massar, and S. Popescu. Error filtration and entanglement purification for quantum communication. *Phys. Rev. A*, 72:012338, 2005. URL http://dx.doi.org/10.1103/PhysRevA.72.012338.

[89] S. J. Devitt, W. J. Munro, and K. Nemoto. Quantum error correction for beginners. *Rep. Prog. Phys.*, 76(7):076001, 2013. URL http://dx.doi.org/10.1088/0034-4885/76/7/076001.

[90] A. Ambainis, M. Mosca, A. Tapp, and R. De Wolf. Private quantum channels. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, page 547. IEEE, 2000. URL https://doi.org/10.1109/SFCS.2000.892142.

[91] K. Fukuda. CDD, v0.94g, 2012. https://www.inf.ethz.ch/personal/fukudak/cdd_home/.

[92] A. A. Abbott, J. Wechs, F. Costa, and C. Branciard. Supplementary material for "genuinely multipartite noncausality", 2017. https://arxiv.org/src/1708.07663v2/anc/Supplementary_Material.cdf.

[93] N. Miklin, A. A. Abbott, C. Branciard, R. Chaves, and C. Budroni. The entropic approach to causal correlations. *New J. Phys.*, 19(11):113041, 2017. URL http://dx.doi.org/10.1088/1367-2630/aa8f9f.

[94] Ä. Baumeler, A. Feix, and S. Wolf. Maximal incompatibility of locally classical behavior and global causal order in multi-party scenarios. *Phys. Rev. A*, 90:042106, 2014. URL http://dx.doi.org/10.1103/PhysRevA.90.042106.

[95] J. von Neumann and R.T. Beyer. *Mathematical Foundations of Quantum Mechanics*. Goldstine Printed Materials. Princeton University Press, 1955. URL https://doi.org/10.1063/1.3061789.

[96] Y. Aharonov, D. Z. Albert, and L. Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, Apr 1988. URL https://doi.org/10.1103/PhysRevLett.60.1351.

[97] H. M. Wiseman. Weak values, quantum trajectories, and the cavity-qed experiment on wave-particle correlation. *Phys. Rev. A*, 65:032111, 2002. URL http://dx.doi.org/10.1103/PhysRevA.65.032111.

[98] N. Brunner, A. Acín, D. Collins, N. Gisin, and V. Scarani. Optical telecom networks as weak quantum measurements with postselection. *Phys. Rev. Lett.*, 91:180402, Oct 2003. URL https://doi.org/10.1103/PhysRevLett.91.180402.

[99] L. Diósi. Structural features of sequential weak measurements. *Phys. Rev. A*, 94:010103(R), 2016. URL https://doi.org/10.1103/PhysRevA.94.010103.

[100] Y. Aharonov and L. Vaidman. Properties of a quantum system during the time interval between two measurements. *Phys. Rev. A*, 41:11–20, 1990. URL https://doi.org/10.1103/PhysRevA.41.11.

[101] Y. Aharonov and L. Vaidman. *The Two-State Vector Formalism: An Updated Review*, pages 399–447. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. URL https://doi.org/10.1007/978-3-540-73473-4_13.

[102] G. Mitchison, R. Jozsa, and S. Popescu. Sequential weak measurement. *Phys. Rev. A*, 76:062105, 2007. URL https://dx.doi.org/10.1103/PhysRevA.76.062105.

[103] K. J. Resch and A. M. Steinberg. Extracting joint weak values with local, single-particle measurements. *Phys. Rev. Lett.*, 92:130402, 2004. URL https://dx.doi.org/10.1103/PhysRevLett.92.130402.

[104] A. Bednorz, K. Franke, and W. Belzig. Noninvasiveness and time symmetry of weak measurements. *New J. Phys.*, 15:023043, 2013. URL https://doi.org/10.1088/1367-2630/15/2/023043.

[105] D. Curic, M. C. Richardson, G. S. Thekkadath, J. Flórez, L. Giner, and J. S. Lundeen. Experimental investigation of measurement-induced disturbance and time symmetry in quantum physics. *Phys. Rev. A*, 97:042128, 2018. URL https://doi.org/10.1103/PhysRevA.97.042128.

[106] A. Bednorz and W. Belzig. Quasiprobabilistic interpretation of weak measurements in mesoscopic junctions. *Phys. Rev. Lett.*, 105:106803, 2010. URL https://doi.org/10.1103/PhysRevLett.105.106803.

[107] J. S. Lundeen and C. Bamber. Procedure for direct measurement of general quantum states using weak measurement. *Phys. Rev. Lett.*, 108:070402, 2012. URL https://doi.org/10.1103/PhysRevLett.108.070402.

[108] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen. Direct measurement of the density matrix of a quantum system. *Phys. Rev. Lett.*, 117:120401, 2016. URL https://doi.org/10.1103/PhysRevLett.117.120401.

[109] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. Colloquium: Understanding quantum weak values: Basics and applications. *Rev. Mod. Phys.*, 86:307–316, 2014. URL https://doi.org/10.1103/RevModPhys.86.307.

[110] E. Cohen. Quantum measurements - yet another surprise. *Quantum Views*, 3:27, 2019. URL https://doi.org/10.22331/qv-2019-11-15-27.

[111] M. Yang, Q. Li, Z. Liu, Z. Hao, C. Ren, J. Xu, C. Li, and G. Guo. Experimental observation of anomalous weak value without post-selection. *arXiv:1911.10958 [quant-ph]*, 2019. URL https://arxiv.org/abs/1911.10958.

[112] O. Oreshkov. Quantum processes on time-delocalized systems. Causality in the quantum world, Anacapri, 2019. URL https://www.quantumlab.it/wp-content/uploads/2019/09/Oreshkov_Quantum-processes-on-time-delocalized-systems.pdf.

[113] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. private communication, 2019.

[114] J. S. Cotler, G. R. Penington, and D. H. Ranard. Locality from the spectrum. *Commun. Math. Phys.*, 368(3):1267–1296, 2019. URL https://doi.org/10.1007/s00220-019-03376-w.

[115] P. Zanardi. Virtual quantum subsystems. *Phys. Rev. Lett.*, 87:077901, 2001. URL https://doi.org/10.1103/PhysRevLett.87.077901.

[116] P Zanardi, D. A. Lidar, and S. Lloyd. Quantum tensor product structures are observable induced. *Phys. Rev. Lett.*, 92:060402, 2004. URL https://doi.org/10.1103/PhysRevLett.92.060402.

[117] L. Viola, E. Knill, and R. Laflamme. Constructing qubits in physical systems. *Journal of Physics A: Mathematical and General*, 34(35):7067–7079, 2001. URL https://doi.org/10.1088%2F0305-4470%2F34%2F35%2F331.

[118] E. Knill. Protected realizations of quantum information. *Phys. Rev. A*, 74:042301, 2006. URL https://doi.org/10.1103/PhysRevA.74.042301.

[119] D. W. Kribs and R. W. Spekkens. Quantum error-correcting subsystems are unitarily recoverable subsystems. *Phys. Rev. A*, 74:042329, 2006. URL https://doi.org/10.1103/PhysRevA.74.042329.

[120] P. A. Guérin and Č. Brukner. Observer-dependent locality of quantum events. *New J. Phys.*, 20(10):103031, 2018. URL https://doi.org/10.1088%2F1367-2630%2Faae742.