

SIMON MILZ

ON THE OPERATIONAL THEORY OF GENERAL (QUANTUM)
STOCHASTIC PROCESSES



MONASH
University

ON THE OPERATIONAL THEORY
OF GENERAL (QUANTUM)
STOCHASTIC PROCESSES

SIMON MILZ
School of Physics and Astronomy
Monash University

A thesis submitted for the degree of
Doctor of Philosophy
January 2019

To Friedl,
who would have read this thesis cover to cover.

ABSTRACT

Stochastic processes are ubiquitous in nature. They are used to model the stock market, predict the weather, describe transport processes in cells and understand the random motion of particles suspended in a fluid, amongst other fields of application. In the realm of classical physics, stochastic processes are well-defined and their properties are conceptually clear. The situation presents itself much less obvious in quantum mechanics, where ‘natural’ classical properties do not hold anymore, and existing approaches to their description break down when memory effects play a non-negligible role.

Here, using the language of higher-order quantum maps, we develop a comprehensive framework for describing quantum stochastic processes, that constitutes a natural extension of its classical counterpart and coincides with it in the correct limit. With this framework at hand, we recover fundamental mathematical properties of quantum stochastic processes, and prove their equivalence with the theory of quantum causal modeling.

Furthermore, the methods we develop allow for an unambiguous characterization of memory effects in quantum processes that we use to derive the connection between Markovianity, i.e., memorylessness, of a process, and previously used witnesses thereof. By tailoring it to relevant scenarios of restricted control and limited resources, we make this framework amenable to experimental implementation, classifying the maximal amount of information that can be inferred about a quantum stochastic process with limited experimental means. As a special case, this result enables the delineation of classical and quantum processes in the non-Markovian regime.

Finally, we connect the theory of quantum stochastic processes to the closely related field of causally indefinite processes, and provide a constructive probabilistic simulation protocol that exceeds previous ones in terms of its success probability. The corresponding results tie indefinite causal order to the non-local properties of the resources used for their implementation, thus offering a direct link between entanglement, non-locality and causal indefiniteness.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Melbourne, January 2019

Simon Milz

PUBLICATIONS

Most of the main results in this thesis have been presented in the following publications:

- [1] *CP divisibility does not mean Markovianity*
S. Milz, M. S. Kim, F. A. Pollock, and K. Modi
arXiv:1901.05223 (2019). Accepted for publication in Phys. Rev. Lett.
- [2] *Reconstructing open quantum system dynamics with limited control*
S. Milz, F. A. Pollock, and K. Modi
Phys. Rev. A, **98**, 012108 (2018).
- [3] *Entanglement, non-Markovianity, and causal non-separability*
S. Milz, F. A. Pollock, T. P. Le, G. Chiribella, and K. Modi
New J. Phys. **20**, 033033 (2018).
- [4] *Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories*
S. Milz, F. Sakuldee, F. A. Pollock, and K. Modi
arXiv:1712.02589 (2017).

Some ideas and figures have also appeared previously in:

- [5] *The Structure of Quantum Stochastic Processes with Finite Markov Order*
P. Taranto, S. Milz, F. A. Pollock, and K. Modi
Phys. Rev. A **99**, 042108 (2019).
- [6] *Quantum Markov Order*
P. Taranto, F. A. Pollock, S. Milz, M. Tomamichel, and K. Modi
Phys. Rev. Lett. **122**, 140401 (2019).
- [7] *Non-Markovian quantum control as coherent stochastic trajectories*
F. Sakuldee, S. Milz, F. A. Pollock, and K. Modi
J. Phys. A **51**, 414014 (2018).
- [8] *An Introduction to Operational Quantum Dynamics*
S. Milz, F. A. Pollock, and K. Modi
Open Sys. Info. Dyn. **24**, 1740016 (2017).

One of the main aims of this thesis is to provide a comprehensive overview of the theory of general stochastic processes. In order to emphasize the interconnectedness of the above results, they are embedded into a broader research context in this work. While properly referenced throughout, for easier delineation between the existing literature and the original results of the above papers, the following list provides their explicit occurrence by chapter:

Chapter 2 is an overview of the theory of higher order quantum maps and does not contain original results.

The original results of **Chapter 3** are based on Refs. [4]. Specifically, Secs. 3.5 – 3.9, the discussion of the relation between causal modeling and stochastic processes (both in classical physics and quantum mechanics), as well as the derivation of the generalized extension theorem can be found in [4]. The generative models for quantum stochastic processes, presented in Sec. 3.10, are original, unpublished work.

The original results of **Chapter 4** are based on Refs. [1, 5, 6]. In detail, the discussion of CP divisibility and its relation to Markovianity (Secs. 4.8 – 4.8.4) is taken from [1]. The brief discussion of quantum Markov order (Sec. 4.9) can be found in [5, 6].

The original results of **Chapter 5** are based on Ref. [2]. The discussion of restricted process tensors (Secs. 5.1 – 5.5) is entirely original work, taken from [2]. The ensuing results on non-Markovian classical processes (Sec. 5.6.2 – 5.6.4) are original, unpublished work.

The original results of **Chapter 6** are based on Ref. [3]. Specifically, the simulation procedure, as well as the subsequent discussion of simulation probabilities, and the analysis of simulation resources (Secs. 6.3 – 6.5) can be found in [3].

The findings of Refs. [7, 8] are referenced throughout, but do not constitute any of the main results of the thesis.

*A skyscraper skyline along the horizon
As mariners sail in towards Melbourne town,
The city that's full of delightful surprises,
For this is a city of well-earned renown.*

— Dacre Smyth [9]

ACKNOWLEDGMENTS

There are a multitude of people whose help, advice and friendship have been paramount for the completion of my PhD. First and foremost, I am deeply indebted to my supervisors, Kavan Modi and Felix A. Pollock. Their contagious enthusiasm for research, as well as their constant endeavour to create a sociable working environment, has been truly inspiring. Both of them went above and beyond to support me, and I am grateful for having had the privilege to profit from their guidance and tremendous insights over the past three years.

It has been a true pleasure working in the Monash Quantum Information Science group, and I extend my sincere gratitude for all the support, creative space, and insightful conversations its members offered.

I would also like to express my great appreciation to the whole staff of the School of Physics and Astronomy. Without fail, every single one of them provided me with the utmost support in both administrative and academic matters. The accommodating and amicable atmosphere they created made me feel like being part of a great work family. My special thanks goes to Jean Pettigrew, whose last-minute help with respect to bureaucratic procedure was invaluable in the submission process.

Financially, I could not have undertaken this degree without the assistance of the Monash International Postgraduate Research Scholarship and the J. L. William Postgraduate Scholarship, whose support is gratefully acknowledged.

Without people in my life to constantly remind me of the lightness of being, this thesis might never have come into existence. I am thankful to Cody Duncan, Josh Morris, and Francesco Campaioli for their friendship; Dacre Smyth, for always finding the means to raise morale; Corey Jacobs for asking the right questions; and Top Notoh for all the insightful discussions. Many of the ideas this thesis is based upon have been honed in countless conversations with Phil Taranto, and I owe him a great deal of gratitude for always steering me in the right direction.

I am deeply grateful to Agnieszka Górecka, Johanna Schäfer and Phil Taranto for proofreading the manuscript of this thesis, and providing me with invaluable comments and advice to improve upon my early drafts.

I would neither be who, nor where, I am today without the loving and unconditional support from my parents and family, and I am grateful for the undying confidence they have in me. My special thanks extend to Lee and Mikki, who have been my family away from home and made my time in Melbourne an extraordinary one.

Finally, I would like to thank Ida and Martha for bringing color to my life, and Judith for being my never-failing compass. Last, but surely not least, for her unwavering support, I owe particular gratitude to Johanna, the greatest partner in crime.

CONTENTS

1	MOTIVATION AND OUTLINE	1
1.1	Motivation	1
1.2	Outline	2
2	HIGHER-ORDER QUANTUM MAPS	5
2.1	States, Effects and Channels	6
2.2	Graphical Representation of Quantum Networks – Basic Examples	9
2.3	Graphical Representation of Quantum Networks – The General Case	11
2.4	Quantum Combs and Directed Acyclic Graphs	16
2.5	Unified Representation: Choi-Jamiołkowski Isomorphism	17
2.6	Choi-Jamiołkowski Isomorphism: Examples	20
2.7	Trace Preservation, Causal Ordering, and Deterministic Implementation	22
2.8	Concatenating Building Blocks – The Link Product	24
2.9	Deterministic Quantum Networks and Causal Order	26
2.10	Deterministic Combs and Causal Ordering	30
2.11	Higher-Order Quantum Maps – Summary	32
3	CAUSAL MODELING AND GENERAL STOCHASTIC PROCESSES	35
3.1	Classical Stochastic Processes – Basic Concepts	38
3.2	σ -algebras, POVMs and Instruments	41
3.3	Classical Causal Modeling	43
3.4	The Kolmogorov Extension Theorem (KET)	45
3.5	The KET and Classical Causal Modeling	47
3.6	Quantum Mechanics and Compatibility Conditions	49
3.7	Quantum Stochastic Processes	51
3.8	Generalized Extension Theorem (GET)	53
3.8.1	Quantum Stochastic Processes and Quantum Causal Modeling	57
3.8.2	GET \Rightarrow KET	58
3.9	Relation to previous Work	58
3.10	Quantum Stochastic Processes – Outlook	59
3.11	General Stochastic Processes – Summary	62
4	OPEN QUANTUM SYSTEM DYNAMICS	65
4.1	Open Quantum System Dynamics – The Traditional Approach	67
4.2	Initial Correlations and Superchannels	71
4.3	Superchannels, Linearity, Complete Positivity, and Trace Preservation	73
4.4	Multi-Step Processes and Process Tensors	76
4.5	Purifications, Dilations, and Choi States	79
4.6	Memory in Classical Processes	82
4.7	Memory in Quantum Processes	83
4.8	CP Divisibility and Markovianity	88

4.8.1	iCP and oCP Divisibility	90
4.8.2	iCP Divisibility \neq oCP Divisibility	92
4.8.3	Markovianity and oCP Divisibility	93
4.8.4	Temporal Correlations in oCP-Divisible Processes	95
4.9	Quantum Markov Order	98
4.10	Open Quantum Processes – Summary	102
5	PROCESS TENSORS AND LIMITED RESOURCES	105
5.1	Restricted Tomographic Reconstruction	106
5.2	Projective Measurements and Unitary Control	108
5.2.1	Unitary Operations	109
5.2.2	Projective Operations	109
5.3	Properties of Restricted Process Tensors	110
5.3.1	Partial Containment Properties	111
5.3.2	Complete Positivity and Trace Preservation	113
5.4	Witnesses for System-Environment Correlations from Limited Control	115
5.4.1	Witnesses for Non-Markovianity	117
5.5	Examples and Applications of Restricted Process Tensors	119
5.5.1	Reconstruction of Qubit Dynamics	119
5.5.2	Restricted Process Tensors and Quantum Control	122
5.6	Classical Resources	125
5.6.1	Classical Markovian Dynamics	126
5.6.2	Maximally Incoherent Combs (MIC)	128
5.6.3	Non-Discord-Generating-and-Detecting Dynamics and Classical Processes	133
5.6.4	Temporal Correlations in Classical Processes	135
5.7	Limited Resources – Summary	137
6	CAUSALLY INDEFINITE PROCESSES	141
6.1	Causally Ordered Processes – Recap	144
6.2	Local Causal Order	146
6.3	Conditional Simulation of Causally Indefinite Processes	150
6.3.1	Conditional Simulation of Arbitrary Causally non-Separable Processes	150
6.3.2	Conditional Circuits and Valid Processes	154
6.3.3	Probability of Success	156
6.4	Resources for Causally non-Separable Process Matrices	158
6.4.1	The Parallel Case	158
6.4.2	The Serial Case	161
6.5	Causally Indefinite Processes – Summary	162
7	CONCLUSIONS	165
I	APPENDIX	
A	CAUSAL MODELING AND GENERAL STOCHASTIC PROCESSES	169
A.1	Definition of Stochastic Processes	169
A.2	Proof of the Generalized Extension Theorem (GET)	170
A.3	Proof GET \Rightarrow KET	172

A.4	Properties of $1 \rightarrow 3$ Adapters	173
B	OPEN QUANTUM SYSTEM DYNAMICS	177
B.1	Dual Matrices	177
B.2	Proof that oCP Divisibility \neq iCP Divisibility	177
B.3	Conditional non-Signaling Processes that are not oCP-Divisible	178
B.4	Structural Properties of Processes with Finite Markov Order	179
C	PROCESS TENSORS AND LIMITED RESOURCES	181
C.1	Basis of Unitary Maps Acting on a Qubit	181
C.2	Basis of Projective Maps	182
C.3	Derivation of intermediate Restricted Process Tensors	183
C.4	NDGD and Classicality	184
D	CAUSALLY INDEFINITE PROCESSES	187
D.1	Allowed Terms in the Process Matrix $W^{B_2 B_1 A_2 A_1}$	187
	BIBLIOGRAPHY	189

LIST OF FIGURES

Figure 2.1	Graphical illustration of states, transformations, and effects	10
Figure 2.2	Graphical representation of Eq. (2.5)	10
Figure 2.3	Graphical representation of the teleportation protocol	11
Figure 2.4	Graphical representation of Eq. (2.6)	11
Figure 2.5	Composition of quantum maps as a network	12
Figure 2.6	Temporal ordering of quantum networks	13
Figure 2.7	Ordered quantum network	14
Figure 2.8	Depiction of an ordered quantum network as a comb	14
Figure 2.9	Teleportation as the contraction of a comb with preparations and effects	14
Figure 2.10	Contraction of a two-slot comb with a preparation, a transformation and a measurement	15
Figure 2.11	Contraction of a comb with another comb	15
Figure 2.12	Directed acyclic graph (DAG)	16
Figure 2.13	Quantum network and comb corresponding to the DAG in Fig. 2.12. .	17
Figure 2.14	Graphical representation of the trace preservation property	21
Figure 2.15	Graphical representation of the unit trace of a quantum state as a trace preservation condition.	21
Figure 2.16	Difference between the identity map, a measurement outcome that corresponds to the effect Φ^+ , and a preparation of Φ^+	22
Figure 2.17	Different networks corresponding to the same link product $L \star K$	25
Figure 2.18	Deterministic quantum network with final output for $N = 3$	27
Figure 2.19	Graphical representation of the hierarchy of trace conditions for a de- terministic comb $C^{(3)}$	28
Figure 2.20	Structure of a deterministic comb with trivial final output space	29
Figure 2.21	Structure of a comb with trivial initial input space	29
Figure 2.22	Causality constraints for combs	32
Figure 2.23	General causality constraints for combs	32
Figure 3.1	Example trajectory of a particle undergoing Brownian motion	38
Figure 3.2	Consistency condition for classical stochastic processes	41
Figure 3.3	(Quantum) Causal network	44
Figure 3.4	Classical process with and without intervention	48
Figure 3.5	Concatenated Stern-Gerlach experiment	50
Figure 3.6	Concatenated Stern-Gerlach experiment (two-step)	50
Figure 3.7	Graphical representation of a four step process	53
Figure 3.8	Consistency condition for combs	54
Figure 3.9	Graphical representation of a $1 \rightarrow 3$ adapter $A_{1 \rightarrow 3}$	60
Figure 3.10	Action of an $1 \rightarrow 3$ adapter on a one-slot comb $C^{(1)}$	60

Figure 3.11	Consistency condition for adapters	61
Figure 3.12	Requirement for universal $1 \rightarrow 3$ adapters	61
Figure 3.13	Construction of a family of consistent combs	62
Figure 4.1	Quantum channel resulting from open dynamics	69
Figure 4.2	Resulting superchannel from an initially correlated system-environment state and open quantum system dynamics	72
Figure 4.3	Different ways of considering $\mathcal{T}_{2:1}$	74
Figure 4.4	Complete positivity of $\mathcal{T}_{2:1}$	75
Figure 4.5	Illustration of a multi-time experiment	76
Figure 4.6	Example of a correlated control operation over two times	78
Figure 4.7	Experimental preparation of the Choi state of a channel \mathcal{L}	80
Figure 4.8	Experimental preparation of $\mathbb{T}_{N+1:1}$	81
Figure 4.9	Action of a causal break $\mathcal{A}_{x_\alpha y_\alpha}$ on a system-environment state ρ_{se}	84
Figure 4.10	Process tensors and causal breaks	86
Figure 4.11	Structure of a Markovian process	87
Figure 4.12	Possible dilation of a Markovian process	88
Figure 4.13	Experimental scenarios for the reconstruction of the maps $\mathcal{L}_{y x}$	91
Figure 4.14	Non-Markovian oCP-divisible process	94
Figure 4.15	Descriptor $\mathcal{T}_{w:v:u}$ of a process defined on three times u , v , and w	95
Figure 4.16	Hierarchy of processes	96
Figure 4.17	Non-signaling requirement for oCP divisible dynamics	97
Figure 4.18	Graphical representation of $L_{v u}$	97
Figure 4.19	Process with finite Markov order	101
Figure 5.1	Containment property of process tensors	111
Figure 5.2	Partial containment property	112
Figure 5.3	Different scenarios in the reconstruction of $\mathbb{T}_{2:1}$ and/or $\mathbb{T}_{2:1}^R$	115
Figure 5.4	Preparation of a product state	116
Figure 5.5	Markovian process tensor	118
Figure 5.6	Goal of dynamical decoupling	123
Figure 5.7	Markovian process that leads to classical statistics	126
Figure 5.8	Definition of the maps $L_{\beta \alpha}$	127
Figure 5.9	NCGD dynamics	127
Figure 5.10	Maximally incoherent comb	129
Figure 5.11	Non-classical MIC	130
Figure 5.12	Concise dilation of an open process	133
Figure 5.13	NDGD	134
Figure 6.1	Possible global temporal orders for Alice's and Bob's respective experiments	144
Figure 6.2	Causally ordered process $A \prec B$ without an open output wire	145
Figure 6.3	Process where Alice and Bob are independent	146
Figure 6.4	Sketch of causal disorder	147
Figure 6.5	Local causality of process matrices	148
Figure 6.6	Causally ordered circuit with an additional measurement on the environment	151

LIST OF FIGURES

Figure 6.7 Circuit that can simulate any process matrix 152

Figure 6.8 Structural consequences of an initial product state in a parallel circuit . . 159

Figure 6.9 Structural consequences of a unitary of product form 160

Figure 6.10 Structural consequences of product state and product unitary (serial case)162

Figure A.1 Graphical representation of a $1 \rightarrow 3$ adapter $A_{1 \rightarrow 3}$ 173

Figure A.2 Possible deterministic quantum network corresponding to a $1 \rightarrow 3$ adapter174

Figure A.3 Graphical representation of Eq. (A.17) 174

Figure C.1 Non-NDGD dynamics that leads to classical statistics 185

MOTIVATION AND OUTLINE

1.1 MOTIVATION

No physical system is ever fully isolated from its surroundings. In any real-world scenario, a system that is controlled and manipulated for storing or processing information will inevitably interact with an unknown environment. This fact leads to non-trivial dynamics, where memory effects play a non-negligible role. With increasing miniaturization and read-out frequencies in modern-day technologies, the role of such effects will become even more prominent in the future.

Consequently, there have been intense efforts – theoretical, experimental, and engineering – to analyze, model and characterize the dynamics of open systems, and the frameworks developed for their description find application in every branch of the quantitative sciences, including outside the realm of physics.

Despite these efforts, there are still many conceptual and practical open questions, pertaining to the modeling of complex open processes, that hinder both theoretical and technological progress. This is particularly true in quantum mechanics. There, the mere act of observation can disturb the system, a fact which has proven an obstacle to the proper mathematical definition of open quantum processes, and the characterization of their memory effects.

Going forward, a unifying framework for open processes – or, more generally, stochastic processes – is of paramount importance, not only for our theoretical understanding of classical physics and quantum mechanics, but also from an applied perspective; any technological development that aims to harness quantum and memory effects when they are advantageous, or suppress them when they are detrimental, has to start from a clear and experimentally accessible understanding of the underlying concepts.

The overarching purpose of this thesis is hence twofold: On the one hand, it supplies such an understanding, and develops the ‘umbrella theory’, that unifies existing approaches to stochastic processes (classical, quantum, and beyond), simultaneously providing the axiomatic underpinnings for many different frameworks in physics and other disciplines. On the other hand, leveraging this generalized vantage point, the second goal of this thesis is to render the resulting theoretical framework amenable to experimental investigation and implementation, and uncover the boundary between classical and quantum processes.

These results are achieved by assuming a distinctly operational viewpoint, which models processes only in terms of entities that are experimentally accessible in principle, and does away with many of the conceptual pitfalls that earlier approaches are plagued by. Additionally,

the operational approach allows one to investigate the minimal fundamental assumptions that have to go into the description of stochastic processes.

One of the seemingly irrevocable axioms that can be relaxed is that of causal order, and one of the aims of this work is to provide a clear connection between processes with indefinite causal order, and the resources that are necessary for their simulation, thus connecting the well-established formulation of stochastic processes that have causal order with the more speculative field of those that do not.

Overall, this thesis is intended to provide a comprehensive picture of the underlying operational structure of open quantum processes that demonstrates the interconnectedness of seemingly disparate fields of research. The operational approach taken throughout then enables the resolution of apparent roadblocks, and paves the way to an unambiguous investigation of the structural properties of general stochastic processes.

1.2 OUTLINE

Since the research areas covered in this work are intimately linked but contend with a rather wide selection of concepts, much emphasis has been put on a clear presentation of their mutual relation. Consequently, the original research results contained in this thesis are interwoven with the existing literature in their respective fields. Each chapter contains its own introduction and summary, to clarify its position in the corresponding wider context, and to point out potential future research avenues.

This thesis begins in Ch. 2, with a review of the theory of higher-order quantum maps and general quantum networks. The corresponding discussion provides the terminology, as well as the mathematical framework we will employ throughout this thesis. In particular, we encounter the Choi-Jamiołkowski isomorphism, and the link product, the necessary tools to succinctly describe higher-order quantum maps. Additionally, the graphical representation of quantum networks, that will be used frequently, is introduced. The chapter ends with a discussion of the constraints that causal ordering imposes on higher-order quantum maps.

In Ch. 3, we investigate classical stochastic processes and their generalization to the quantum regime in detail. The chapter starts with a discussion of the properties that classical processes satisfy and provide the Kolmogorov extension theorem, the fundamental theorem of the theory of classical stochastic processes. Subsequently, the breakdown of the Kolmogorov extension theorem in quantum mechanics, and the accompanying conceptual problems are highlighted. Their resolution is provided by means of a description of quantum processes in terms of higher-order quantum maps, which enables the derivation of a generalized extension theorem – one of the main results of this thesis – defining the concept of a quantum stochastic process. Additionally, by proving the generalized extension theorem, we demonstrate the equivalence of the fields of quantum stochastic processes and quantum causal modeling and provide an axiomatic underpinning for all dynamical probabilistic theories. Finally, this extension theorem is leveraged to propose abstract generative models of quantum stochastic processes.

Ch. 4 introduces the concept of open quantum system dynamics. After a brief review of traditional approaches to its description, their breakdown in the presence of non-negligible

memory effects is discussed. We present the operational way out of this problem in detail and discuss the mathematical properties of the process tensor, the higher-order quantum map that is the proper descriptor of multi-time open quantum dynamics. With this operational approach at hand, the presence of memory effects can be unambiguously characterized. The main result of the chapter lies in the subsequent alternative definition of CP divisibility and the comprehensive analysis of the temporal correlations, this witness of non-Markovianity is blind to. We round off the chapter with a discussion of the concept of Markov order in the quantum case.

In order to make the process tensor framework more amenable to experimental investigation, in Ch. 5, we tailor it to the case of limited experimental control. The structural properties of the resulting restricted process tensors are investigated, and their experimental reconstruction and ability to detect memory effects is discussed. We provide specific examples for the case, where experimental control is constrained to unitary operations and projective measurements, respectively. Subsequently, classical processes are defined as a special case of limited experimental control, and the direct connection between vanishing quantum discord and the classicality of dynamics is shown. The chapter ends with a comprehensive algebraic characterization of the set of classical processes.

In Ch. 6, we investigate the simulation of causally indefinite processes. After a brief review of the corresponding research area, highlighting the structural properties of processes with indefinite causal processes, we provide a constructive procedure for their simulation by means of a temporally ordered process with an additional conditioning. Following this discussion, the success probability of simulation is analyzed. Finally, we show that the simulation of causally disordered processes requires multipartite entanglement and non-locality as a resource, thus revealing a direct connection between quantum correlations and the absence of a clear causal order.

The thesis concludes in Ch. 7 with final remarks and an outlook on possible future research directions.

HIGHER-ORDER QUANTUM MAPS

The aim of a dynamical description of quantum processes is to quantify how measurement statistics of different observables can change from one moment to the next, even when the system in question may be interacting with its wider environment, which is typically large, uncontrollable and experimentally inaccessible. Quantum mechanics – and more generally any operational probabilistic theory (OPT)¹ – can be considered to consist of three main components: preparations/states, transformations, and measurements/effects. *Any* physically realizable and experimentally probable process can always be represented as a (possibly highly complex) concatenation of these three fundamental building blocks [11]. In order to be able to depict and manipulate general quantum processes in an intuitive manner, in this chapter we will introduce *higher-order quantum maps* alongside their graphical representation in terms of quantum networks, the natural mathematical and graphical formalism to describe quantum dynamics. Put a little bit less prosaically, we will provide both the vocabulary and the syntax, that will be employed to phrase concepts and results throughout this thesis.

Quite naturally, given the huge body of work that exists on higher-order quantum maps and their graphical representation, this chapter cannot give a comprehensive or fully rigorous account of their theory, but, in anticipation of the main subject of this thesis – the description of open quantum system dynamics and the simulation of causal disorder – will rather focus on the key aspects, that are pertinent to the discussion of these subjects. More extensive discussions of graphical calculus for the description of quantum dynamics (and beyond), can, for example, be found in [11, 12] (and with a slightly different motivation in [13, 14]), while the theory of higher-order quantum maps is fully developed, for example, in [15–17]. The latter three articles also form the basis for this chapter’s discussion.

While in subsequent chapters we will explicitly discuss the relation of higher-order quantum maps to open quantum system dynamics in detail, here, we will be agnostic about their physical realization and assume an axiomatic point of view that establishes higher-order quantum maps as arising as general networks of the aforementioned building blocks. The thusly developed framework will then be equipped with a clear underlying physical picture in Ch. 4.

As it turns out, the set of higher-order quantum maps obtained through this constructive approach is equivalent to the set of *temporally ordered processes*, *i.e.*, the set of maps that are compatible with a fixed underlying causal order of events [17, 18]. The discussion in Ch. 4 will show that, additionally, they can be experimentally implemented and reconstructed, and

¹ The natural framework to formalize OPTs is *category theory*. While we will occasionally employ the language of category theory to describe quantum dynamics, we will not make this relation explicit. The interested reader is referred to, *e.g.*, [10] for an introduction to this topic.

are thus faithful to what is observed in nature. In Ch. 6 we will relax the requirement of an underlying global causal order, and discuss the simulation of causally indefinite processes. Here, again, higher-order quantum maps will provide the overarching framework to discuss causally disordered structures and their implementation.

The aim of this chapter is twofold. On the one hand, it provides a pedestrian introduction to the theory of higher-order quantum maps and the powerful graphical calculus it can be equipped with. On the other hand, it establishes the notation, mathematical framework and key concepts, like linearity, complete positivity, trace preservation and causality, that will be used throughout this thesis. Before we start the discussion of quantum networks, though, a disclaimer is in order: We will establish a notation in this chapter that is tailored to higher-order quantum maps. However, this thesis spans topics from seemingly distinct fields of physics, with traditionally differing notations, and in different chapters, we will aim to emphasize different aspects of the theory. It would thus be futile – and quite often obfuscating – to stubbornly adhere to one and the same notational convention. Nonetheless, all future notation will be based upon the one we will now introduce, and any change in notation will be made obvious. With this in mind, higher-order quantum maps provide the versatile framework, that allows us to phrase the results in the subsequent chapters in a concise way, and a thorough introduction of the key concepts will avoid later conceptual detours. Additionally, whenever possible, we shall equip our results with their graphical representation, thus making them more tangible, and alleviating potential notational pitfalls.

2.1 STATES, EFFECTS AND CHANNELS

As we briefly mentioned above, the fundamental building blocks of quantum mechanics are states/preparations, transformations and measurements/effects. After introducing them, and describing their properties in this section, we will see below, how to concatenate them to form complex quantum networks.

The **state** of a quantum system is given by a density operator $\rho \in \mathcal{B}(\mathcal{H})$ that is a trace class (*i.e.*, its trace is well-defined) element of the space of bounded operators on the Hilbert space \mathcal{H} . Unless stated otherwise, we will always consider the dimension $d = \dim(\mathcal{H})$, corresponding to the number of perfectly distinguishable states the system can assume, to be finite.² Consequently, all operators have a matrix representation and density operators will henceforth be called density *matrices*. Intuitively, a density matrix contains all the information that can be deduced about the state of the system, *i.e.*, it contains all probabilities for all possible outcomes of all measurements that could possibly be performed on the system [19]. Accordingly, we have

$$\rho = \rho^\dagger \text{ (Hermiticity)}, \quad \rho \geq 0 \text{ (Positivity)} \quad \text{and} \quad \text{tr}(\rho) = 1 \text{ (Unit Trace)}, \quad (2.1)$$

to ensure that probabilities are real, positive, and sum up to one, respectively.

The second building block, **effects** (or measurements), are given by positive linear functionals \mathcal{E}_α , that map states to probabilities, $\rho \mapsto \mathbb{P}_\alpha = \mathcal{E}_\alpha[\rho]$ [20, 21]. In detail, for any instrument \mathcal{J} ,

² For the finite dimensional case, the space $\mathcal{B}(\mathcal{H})$ coincides with the space of linear operators on \mathcal{H} and all operators in $\mathcal{B}(\mathcal{H})$ are trace class.

that is used to interrogate the state of the system, every measurement outcome α corresponds to an effect \mathcal{E}_α . For the moment, we will consider an instrument to simply be a rule that relates an outcome for a given measurement procedure to its corresponding effect. For example, if a qubit – a quantum mechanical two-level system – was measured in the X -direction, the two possible outcomes could be labeled with 0 and 1, and the corresponding effects would act as $\{\mathcal{E}_0[\rho] = \langle +|\rho|+\rangle, \mathcal{E}_1[\rho] = \langle -|\rho|-\rangle\}$, where $|\pm\rangle$ are the eigenstates of the Pauli σ_x operator. The concept of instruments will be generalized and discussed in more detail in Ch. 3. As probabilities add up to unity, the effects corresponding to an instrument are normalized: $\sum_{\alpha=1}^D \mathcal{E}_\alpha = \text{tr}$, where tr is the trace operator, and, importantly, the number D of effects does not have to coincide with d .

It is convenient, using the Riesz representation theorem, to express effects in terms of positive operator valued measure (**POVM**) elements $E_\alpha^\text{T} \in \mathcal{B}(\mathcal{H})$, where the transpose \cdot^T is added for consistency with later notation (see Sec.2.5). We will employ the convention that effects are denoted by calligraphic letters, and the corresponding POVM element by the same, upright sans-serif letter. Positivity of the effects translates to positivity of the corresponding POVM elements, *i.e.*, $E_\alpha \geq 0$, and their normalization to $\sum_{\alpha=1}^D E_\alpha = \mathbb{1}$, where $\mathbb{1}$ is the identity matrix on \mathcal{H} . With this, the probability $\mathbb{P}(\alpha|\mathcal{J})$ for a measurement outcome α given that the instrument \mathcal{J} was used to probe the system can be computed via Born’s rule

$$\mathbb{P}(\alpha|\mathcal{J}) = \text{tr}(\rho E_\alpha^\text{T}). \quad (2.2)$$

We note that in the literature, quite often the role of states and effects is reversed, *i.e.*, effects are introduced as positive matrices, and states as linear functionals on the effects [20, 21]. In finite dimensions, both points of view are equivalent, and the choice we made corresponds to choosing a Schrödinger picture of dynamics. It is worth noting, that, starting with the notion of a POVM and demanding a probabilistic structure, and the premise that the probabilities corresponding to each of the POVM elements are positive and sum up to one, suffices to prove both the existence of a density matrix that with the properties (2.1) as well as the computation of probabilities via Born’s rule (2.2) [22–24]. We will return to this point when we discuss general stochastic processes in Ch. 3.

Finally, **transformations** of quantum states are maps $\mathcal{L} : \mathcal{B}(\mathcal{H}^i) \rightarrow \mathcal{B}(\mathcal{H}^o)$ that map quantum states $\rho \in \mathcal{B}(\mathcal{H}^i)$ defined on an ‘input’ space (denoted by i) to states $\mathcal{L}[\rho] = \rho' \in \mathcal{B}(\mathcal{H}^o)$ defined on an ‘output’ space (denoted by o).³ In general, \mathcal{H}^i and \mathcal{H}^o can differ (*e.g.*, \mathcal{L} could create or destroy particles and/or add or trace out degrees of freedom). For easier ‘bookkeeping’, we will clearly distinguish between the input and output space, even in cases where both spaces are isomorphic (denoted by $\mathcal{H}^i \cong \mathcal{H}^o$). To represent a deterministic physical process, a quantum map \mathcal{L} must preserve the basic properties of the density operator ρ , *i.e.*, it has to preserve trace and positivity. Trace preservation (**TP**) ensures that the probabilities obtained from a state $\rho' = \mathcal{L}[\rho]$ still add up to one, while positivity preservation ensures that each of these probabilities is still positive. We will call a map \mathcal{L} that is positivity preserving

³ As was the case for effects, we will use the Schrödinger picture to describe transformations, and, in general higher-order maps. For a discussion in the Heisenberg picture, see, for example, Ref. [18].

‘positive’, and denote it by $\mathcal{L} \geq 0$. More generally, we could imagine a situation, where \mathcal{L} only acts on a subset of the degrees of freedom of a quantum state $\eta \in \mathcal{B}(\mathcal{H}^i \otimes \mathcal{H}_a^i)$, *i.e.*,

$$\eta' = (\mathcal{L} \otimes \mathcal{I}_a) [\eta], \quad (2.3)$$

where \mathcal{I}_a is the identity map on the additional degrees of freedom that \mathcal{L} does not act on, *i.e.*, $\mathcal{I}_a[\rho] = \rho$ for all $\rho \in \mathcal{H}_a^i$. In order for \mathcal{L} to represent a valid physical evolution, η' has to be positive, for any size of additional Hilbert space \mathcal{H}_a^i , *i.e.*, $\mathcal{L} \otimes \mathcal{I}_a \geq 0$ for all $d_a = \dim(\mathcal{H}_a^i)$. Maps that satisfy this requirement are called *completely positive (CP)*. While there has been some debate about whether quantum maps have to be CP in general [25–27], giving up complete positivity would mean giving up the Holevo quantity [28], data processing inequality [29], and entropy production inequality [30], and is therefore not desirable. It has been shown that the problems that led researchers to question the requirement of complete positivity in the first place can be alleviated by properly accounting for system-environment correlations, and there is no need to consider giving it up [31, 32]. We will discuss this resolution in detail in Ch. 3. A thorough discussion of the arguments against CP maps can, *e.g.*, be found in [8].

Lastly, a quantum map has to act linearly on the set of quantum states:

$$\mathcal{L} \left[\sum p_k \rho_k \right] = \sum p_k \mathcal{L}[\rho_k] = \sum p_k \rho'_k, \quad (2.4)$$

where $\{p_k\}$ is a probability distribution. It is worth noting that this requirement does *not* follow from the fact that quantum mechanics – in the sense of quantum state vectors formed from linear superpositions of a basis set – is a linear theory (in fact, \mathcal{L} is not generally linear in this sense). Instead, the linearity of the quantum map is analogous to the linearity of mixing in a statistical theory.

To better appreciate this, consider a communication channel \mathcal{L} from Alice to Bob, where Alice randomly prepares a system in either state ρ_1 (with probability p) or ρ_2 (with probability $1 - p$); she then sends the system to Bob (via \mathcal{L}), who performs measurements on the states that he receives. Having no idea whether Alice sent ρ_1 or ρ_2 in each run, from his measurement outcomes, Bob would conclude that the state he receives is $\rho' = \mathcal{L}[\rho]$, where $\rho = p\rho_1 + (1 - p)\rho_2$. That is, we can interpret Alice’s preparation to be the average state. Now suppose that, at some later point, Alice reveals which state was sent in which run; Bob can now go back to his logbook and conclude that he received the state ρ'_1 (ρ'_2) whenever Alice sent him ρ_1 (ρ_2). Conversely, averaging over that data would amount to Bob receiving ρ' . Thus we must have $\rho' = p\rho'_1 + (1 - p)\rho'_2$. This simple thought experiment demands that the action of quantum channels must be linear.⁴ Somewhat interestingly, this argument does not employ quantum mechanics at any point; linearity of mixing is a general concept that applies to all probabilistic theories. With this, we will henceforth consider quantum maps to be linear, completely positive and trace-preserving (CPTP) maps. Such maps are also called *channels* or *quantum maps*.

Intuitively, these three building blocks describe the basic constituents of information flow in quantum dynamics. Information can be created by preparing a state, it can be transformed and

⁴ Strictly speaking, this is only an argument for linearity on convex combinations of states. Linearity in the complete sense is somewhat hard to argue for in general [33], however, maps that are linear on convex combinations of states can always be extended to linear maps [17] (or rather, they can be linearized [34]), and we will hence assume all quantum operations to be linear from now on.

changed by undergoing a dynamics, and it can be read-out via a measurement. There is no need to introduce further elements, and the remainder of this chapter will be concerned with the question of how to concatenate these building blocks, and how to compute the higher-order quantum maps that describe the resulting quantum networks.

2.2 GRAPHICAL REPRESENTATION OF QUANTUM NETWORKS – BASIC EXAMPLES

Every quantum mechanical process can be thought of as a network of preparations, transformations and measurements in some kind of spatio-temporal arrangement.⁵ For example, an experiment could consist of the initial preparation of a quantum state $\rho \in \mathcal{B}(\mathcal{H}_a^i \otimes \mathcal{H}_b^i)$, followed by a transformation $\mathcal{L} : \mathcal{B}(\mathcal{H}_a^i) \rightarrow \mathcal{B}(\mathcal{H}_c^o)$, that only acts non-trivially on the Hilbert space \mathcal{H}_a^i , and finally a measurement on the whole state with POVM element $E_\alpha^T \in \mathcal{B}(\mathcal{H}_b^i \otimes \mathcal{H}_c^o)$. Using Eq. (2.2), the probability for the outcome α that E_α^T corresponds to can be computed via

$$\mathbb{P}(\alpha|\mathcal{J}) = \text{tr} \left[E_\alpha^T (\mathcal{L} \otimes \mathcal{I}_b) [\rho] \right] \quad (2.5)$$

It proves helpful to introduce a graphical notation in the spirit of quantum circuits notation [19] to represent complex networks of elementary building blocks. To this end, we will denote transformations (preparations/states and effects) as boxes (triangles) with a number of wires coming out of them. Each wire corresponds to the space of bounded operators on a Hilbert space, but for conciseness, we will simply label them by the corresponding Hilbert space (or the label thereof, see below). In this notation, a state is depicted as a triangle with outgoing wires (see Fig. 2.1), a transformation is depicted as a box with incoming and outgoing wires, and an effect is depicted as a triangle with only incoming wires. By convention, in the following diagrams, we will consider time to flow from left to right. Neither the grouping, nor the positioning of the wires is unique and can be chosen freely (see Fig. 2.1 and the examples below). With this, we can easily represent Eq. (2.5) graphically with Fig. 2.2. Importantly, the type of the result of a concatenation can be directly read off from the remaining open wires. For example, the network corresponding to Eq. (2.5) has no open wires, and as such represents a number (more precisely, a probability). Accordingly, networks with only open outgoing wires, *i.e.*, wires pointing to the right, correspond to states, networks with only open incoming wires correspond to effects, and networks with both open incoming *and* outgoing wires correspond to transformations.

In order to build some intuition, it is insightful to provide two simple examples of a quantum network for the well-known quantum teleportation protocol, and a communication scenario over noisy channels.

EXAMPLE 2.1 (STATE TELEPORTATION): *Consider the well-known state teleportation protocol [19, 35, 36] (the representation of this example is inspired by the similar graphical representation in [14]). In its simplest version, Bob would like to teleport a qubit state $\Psi := |\Psi\rangle\langle\Psi| \in \mathcal{B}(\mathcal{H}_{B_2}^i)$ to Alice, by using classical communication and a shared maximally entangled two-qubit state $\Phi^+ := |\Phi^+\rangle\langle\Phi^+| \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_{B_1}^i)$, where $|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i_A\rangle |i_{B_1}\rangle$, and $\{|i_X\rangle\}$ is the*

⁵ Most of the following graphical arguments do not rely on the particular properties of quantum mechanics, but would also hold in OPTs (see, for example, Ref. [11]).

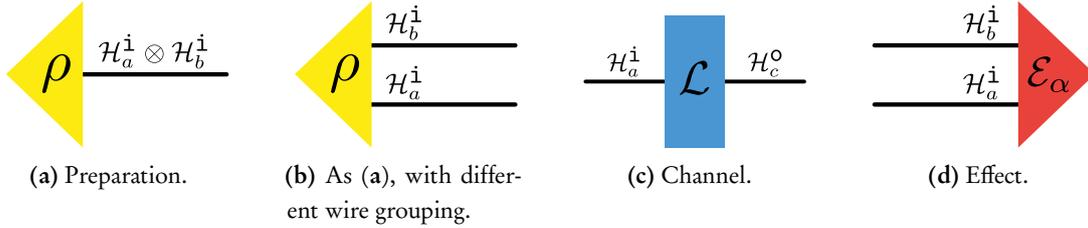


Figure 2.1: Graphical illustration of states ((a) and (b)), transformations (c) and effects (d). Time runs from left to right, *i.e.*, states (effects) only have outputs (inputs). The difference in representation of (a) and (b) demonstrates the non-uniqueness of the wire grouping.

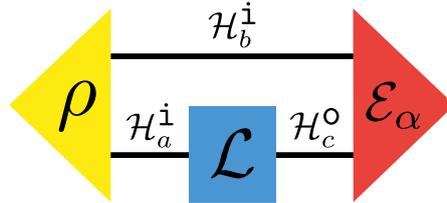


Figure 2.2: Graphical representation of Eq. (2.5). The resulting network has no open wires and corresponds to a number (the probability for the outcome α corresponding to \mathcal{E}_α).

computational basis of \mathcal{H}_X . In order to teleport the state, Bob performs a Bell measurement on his half of the maximally entangled state and the state he aims to teleport. This measurement yields one of four possible outcomes corresponding to the four Bell states $\Phi_j := (\sigma_j \otimes \mathbb{1})\Phi^+(\sigma_j \otimes \mathbb{1})$, where $\{\sigma_j\}_{j=0}^3 = \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ are the Pauli matrices. Communicating his measurement outcome j to Alice, she can then perform a correction unitary σ_j^* , where $*$ denotes complex conjugation in the computational basis. The corresponding network is depicted in Fig. 2.3. As expected, it has one outgoing wire, and it can be shown straightforwardly that the resulting state coincides with Ψ .⁶

EXAMPLE 2.2: As a second example, consider quantum communication from Alice and Bob over a noisy channel \mathcal{L} . In the simplest scenario, Alice can make multiple use of the channel, and each use is independent of the preceding ones [41] (for a discussion of this scenario that considers memory effects, see for example [18]). In order to send her message X , Alice can encode it in a quantum state ρ_X , and make multiple use of \mathcal{L} to transmit the state to Bob, who performs a measurement of the state he receives. Assuming that outcome x is interpreted by Bob as ‘Alice sent message X ’, the goal of this communication is to maximize the probability $\mathbb{P}(x|\mathcal{J}, X)$ to measure x using instrument \mathcal{J} , given that the message X was encoded. We have

$$\mathbb{P}(x|\mathcal{J}, X) = \text{tr} \left[\mathbf{E}_x^\top (\mathcal{L} \otimes \cdots \otimes \mathcal{L}) [\rho_X] \right]. \quad (2.6)$$

The corresponding diagrammatic representation is shown in 2.4. Here, we do not intend to discuss the optimization of this communication task, but merely use it as an example for the depiction of quantum processes in our circuit formalism. Additionally, we want to point out explicitly, that

⁶ Besides the algebraic proof, it can also be shown graphically by means of tensor network calculus [13, 14, 37–40] that the network of Fig. 2.3 indeed corresponds to teleportation of Ψ . As we will not make use of this graphical calculus throughout this thesis, the reader is referred to the aforementioned sources for an in-depth discussion of these techniques.

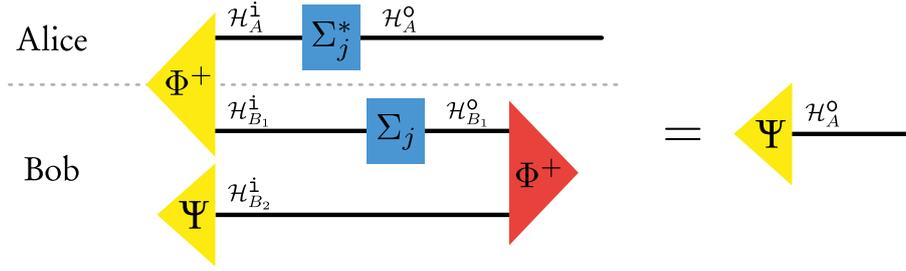


Figure 2.3: Graphical representation of the teleportation protocol. Initially, Alice and Bob share a maximally entangled state (for better orientation, the separation between their respective labs is represented by the dotted gray line). The map Σ_j is the unitary transformation $\Sigma_j[\rho] = \sigma_j \rho \sigma_j$, analogously Σ_j^* acts as $\Sigma_j^*[\rho] = \sigma_j^* \rho \sigma_j^*$. The resulting network has one outgoing wire, *i.e.*, it corresponds to a state. Knowing Bob's measurement outcome j – corresponding to a projection on the Bell state $(\Sigma_j \otimes \mathcal{I})(\Phi^+)$ – Alice performs the proper correcting transformation such that this state is equal to Ψ . In anticipation of later discussions (see Sec. 2.8), all input and output spaces have different labels, even if they are isomorphic.

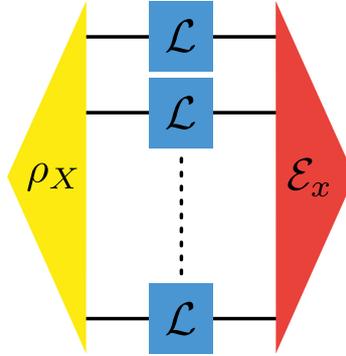


Figure 2.4: Graphical representation of Eq. (2.5). The message X is encoded in ρ_X by Alice, who sends it to Bob via multiple use of \mathcal{L} . Bob performs a measurement to decode the message. The diagram has no open wires, and as such represents a probability, namely $\mathbb{P}(x|\mathcal{J}, X)$.

objects that act on different Hilbert spaces are drawn in parallel (for example the different instances of \mathcal{L} in Fig. 2.4), while outputs and inputs that correspond to the same Hilbert spaces are joined.

2.3 GRAPHICAL REPRESENTATION OF QUANTUM NETWORKS – THE GENERAL CASE

With this intuition for small networks in place, we can now move on to more general quantum networks. First, let us consider the case of a network that consists of transformations $\{\mathcal{L}_n\}_{n=0}^{N-1}$ that are performed in succession, *i.e.*, \mathcal{L}_i comes before \mathcal{L}_j if $i < j$. Each of the maps has an input and an output space, and the input (output) space of a map \mathcal{L}_j corresponds to the time t_j (t_{j+1}). In general, the input space of any map \mathcal{L}_j has a non-trivial overlap with the output spaces of all maps \mathcal{L}_i with $i < j$. This concatenation of maps can be written as

$$\mathcal{C} = \mathcal{L}_{N-1} \circ \cdots \circ \mathcal{L}_1 \circ \mathcal{L}_0, \quad (2.7)$$

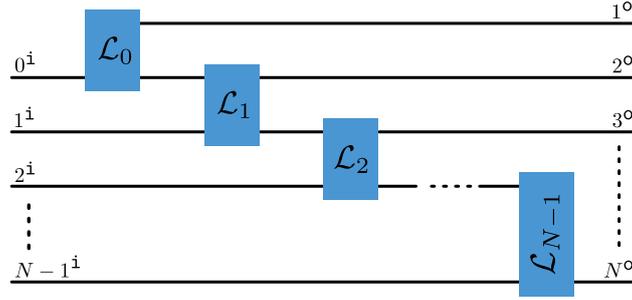


Figure 2.5: Composition of quantum maps as a network. The wires are labeled according to the space they represent. For example, a wire labeled 2^i corresponds to the space $\mathcal{B}(\mathcal{H}_2^i)$, etc. The resulting object is a Channel $\mathcal{C} : \mathcal{B}(\mathcal{H}_0^i \otimes \cdots \otimes \mathcal{H}_{N-1}^i) \rightarrow \mathcal{B}(\mathcal{H}_1^o \otimes \cdots \otimes \mathcal{H}_N^o)$ (for conciseness, only the Hilbert spaces of the open wires are labeled). Importantly, each label denotes the time it corresponds to. For example, both 2^i and 2^o correspond to the time t_2 .

where \circ denotes the composition of maps, and its corresponding quantum network is depicted in Fig. 2.5. Denoting the remaining open input spaces by $\{\mathcal{H}_0^i, \dots, \mathcal{H}_{N-1}^i\}$ and the remaining open output spaces by $\{\mathcal{H}_1^o, \dots, \mathcal{H}_N^o\}$, the resulting map \mathcal{C} can be considered a CPTP map

$$\mathcal{C} : \mathcal{B}(\mathcal{H}_0^i \otimes \cdots \otimes \mathcal{H}_{N-1}^i) \rightarrow \mathcal{B}(\mathcal{H}_1^o \otimes \cdots \otimes \mathcal{H}_N^o). \quad (2.8)$$

See Fig. 2.5 for a graphical representation. However, the order in which the maps $\{\mathcal{L}_i\}$ act provides us with more detailed structural properties corresponding to the temporal ordering, encapsulated in \mathcal{C} , that would not necessarily be reflected by an arbitrary CPTP map from the correct input to the correct output space. A general CPTP map $\mathcal{C} : \mathcal{B}(\mathcal{H}_0^i \otimes \cdots \otimes \mathcal{H}_{N-1}^i) \rightarrow \mathcal{B}(\mathcal{H}_1^o \otimes \cdots \otimes \mathcal{H}_N^o)$ could describe an experiment that happens at *one* point in time, in contrast to the network in Fig. 2.5, that describes a succession of events at *different* times.⁷ For example, consider an experimental situation, where the degrees of freedom in Alice’s laboratory are described by $\mathcal{B}(\mathcal{H}_A) = \mathcal{B}(\mathcal{H}_0^i \otimes \cdots \otimes \mathcal{H}_{N-1}^i)$ and she implements a transformation \mathcal{C} that maps some initial state ρ she prepares to an output state $\rho' = \mathcal{C}[\rho] \in \mathcal{B}(\mathcal{H}_1^o \otimes \cdots \otimes \mathcal{H}_N^o)$. While the involved input and output spaces are the same as in the network depicted in Fig. 2.5, both experimental situations differ fundamentally. For example, in the single-time experiment, changing the input state on any of the input spaces on Alice’s side can potentially influence the output state on *any* of the output spaces on Bob’s side. This is not the case in the multi-time case 2.5, where, for instance, a preparation of a state at t_3 has no influence on the measurement statistics at an earlier time t_2 . On the other hand, in the single time experiment, no output can have an influence on any of the inputs, while later inputs can be conditioned on earlier outputs in the network 2.5.

In detail, given the network that corresponds to \mathcal{C} from Eq. (2.7), an experimenter⁸ could prepare states $\{\rho_{y_j} \in \mathcal{B}(\mathcal{H}_j^i)\}_{j=0}^{N-1}$ and feed them into the network. Analogously, they can make measurements with outcomes corresponding to Effects $\{\mathcal{E}_{x_{j+1}}\}_{j=0}^{N-1}$. Now, the experimenter

⁷ We are somewhat lax in our use of the term ‘event’ throughout this thesis. Here, for all intents and purposes, we will mean preparations and measurements associated with a Hilbert space when we use the word event. We will make this terminology more precise in Ch. 6. For a more thorough discussion of the concept of an event, see for example [42, 43]

⁸ Here, and in what follows, we will often use the idea of an experimenter to explain the concepts at hand. This anthropomorphism is not a necessity, but merely a helpful and intuitively accessible visualization.

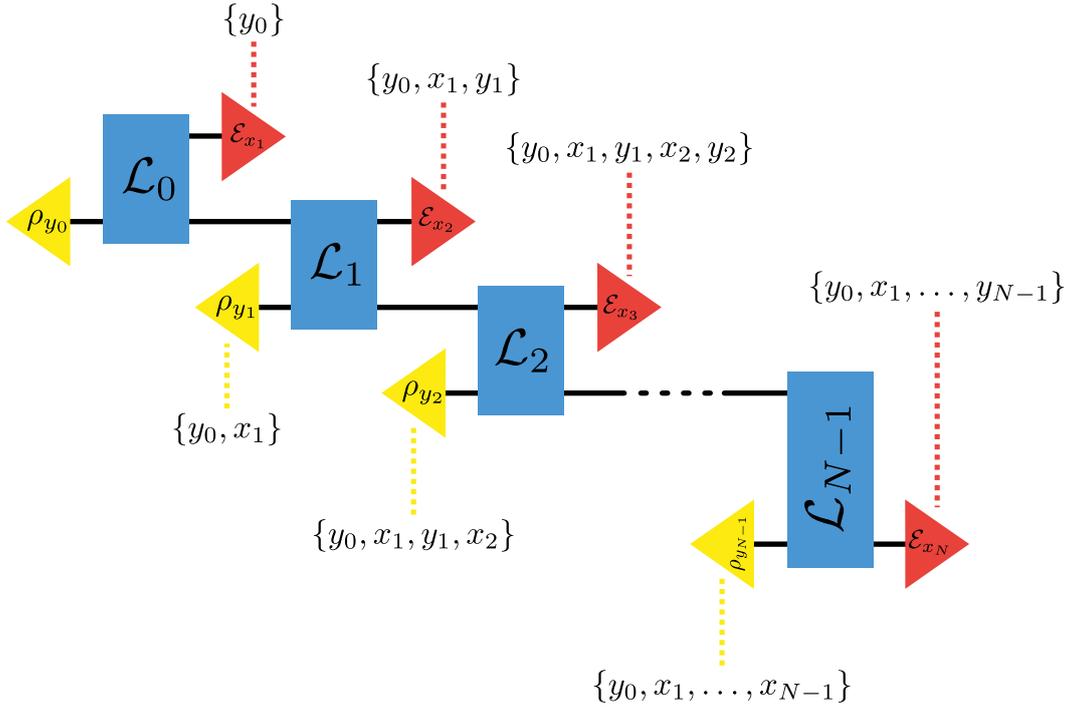


Figure 2.6: Temporal ordering of quantum networks. Statistics of measurements can be influenced by earlier state preparations and earlier measurement outcomes. Analogously, an experimenter can condition their preparations on earlier preparations, as well as earlier measurement outcomes. To emphasize this point, each effect and preparation is additionally labeled with the outcomes (denoted by x_j) and preparations (denoted by y_j) they can be influenced by/conditioned on.

could condition what state they prepare at a time t_j on all the states they prepared and all the measurement outcomes they obtained at times $t_i < t_j$. Quite naturally, the choice of preparation at t_j cannot depend on any *later* time $t_k > t_j$. Similarly, the probability to obtain an outcome x_j at t_j corresponding to the Effect E_{x_j} can be influenced both by preparations of states and measurement outcomes at earlier times $t_i < t_j$, but not by anything that happens at a time $t_k > t_j$ (see Fig. 2.6 for a graphical representation)⁹. Simply put, the structure of the network dictates, how information can flow, and what events can have an influence on others. More technically, the structure of the network dictates certain *no-signaling* conditions that have to be satisfied by \mathcal{C} in order to be able to correspond to a given network [17, 44, 45]; only events at earlier times can signal to later events, and not vice versa. While we will make this statement precise in the following section, here, we want to represent it graphically by reshaping the network 2.5. The order of events will be represented by the labeling of the corresponding wires, and is such that

$$0^i \prec 1^0 \prec 1^i \prec \dots \prec N-1^i \prec N^0, \quad (2.9)$$

where $a \prec b$ implies that events corresponding to the Hilbert space associated with a can potentially influence events corresponding to the Hilbert space associated with b , but not the other way around. We emphasize that this relation is transitive, *i.e.*, $a \prec b$ and $b \prec c$ implies $a \prec c$.

⁹ The phrasing of this argument makes it seem like there can only be *classical* correlations between choices of preparations and/or measuring instrument. As we will outline in the next section, this is not the case.

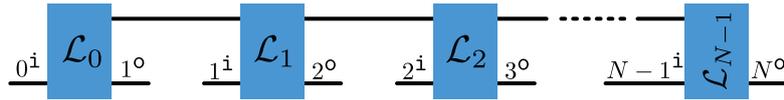


Figure 2.7: Ordered quantum network. This network is topologically equivalent to the one in Fig. 2.5. The ordering of the wires is such that it represents the causal ordering of events in a visually clear way.

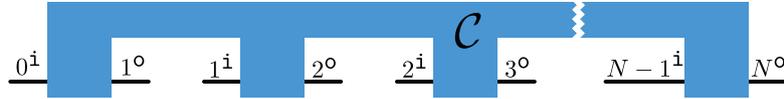


Figure 2.8: Depiction of an ordered quantum network as a comb. Quantum networks can be alternatively represented by comb-like structures, with the understanding, that ‘internally’, this comb has the form of the network depicted in Fig. 2.7.

Taking its temporal order into account, the network of Fig. 2.5 can be reshaped such that all the wires that correspond to earlier events are to the left of wires that correspond to later ones. The corresponding reshaped network is depicted in Fig. 2.7. As in this representation, the network looks like it has teeth and slots, these general temporally ordered networks have been named *quantum combs* [16, 17] (see Fig. 2.8). They are the most general descriptors of quantum processes that are compatible with an underlying temporal order (see also Ch. 4).

Adding preparations and effects to the network now simply consists of attaching their corresponding graphical representation to the correct wires of the comb, and the structure of the resulting network can be read off directly. For example, we can understand the teleportation circuit of Fig. 2.3 in this way; there, the two unitary transformations Σ_j and Σ_j^* can be considered the (simple) underlying comb, and the final transformation circuit is obtained by joining the correct preparations and effects at the corresponding wires (see Fig. 2.9). More generally, the open wires of a comb can be joined with any number of preparations, effects and transformations. For example, in Fig. 2.10, we join a comb with three slots with a preparation, an effect and a transformation to obtain a one-slot comb.¹⁰ We will call the joining of open wires

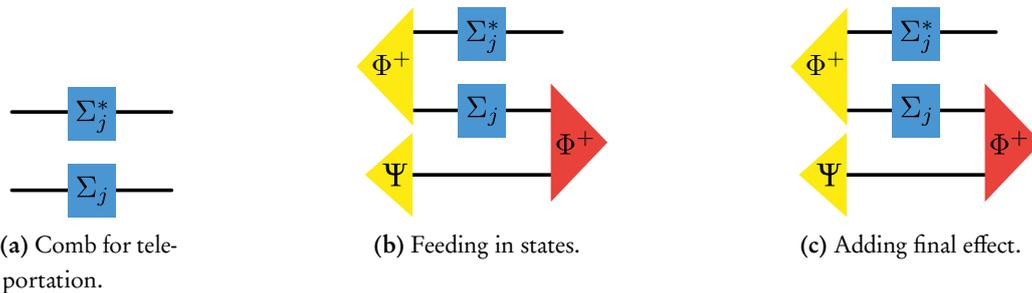


Figure 2.9: Teleportation as the contraction of a comb with preparations and effects. While not the original setup, we can consider the circuit for quantum teleportation as an exercise in contracting a given comb with preparations and effects.

¹⁰ It is important to notice that we are imprecise in nomenclature here. As effects – except for the effect tr – cannot be implemented deterministically, the resulting network is technically not a comb (as it is not TP), but a *tester* [17, 46, 47]. Non-deterministic networks will be discussed in Sec. 2.8. As we only focus on the graphical representation of quantum networks in this section, this distinction is not of importance for the moment.

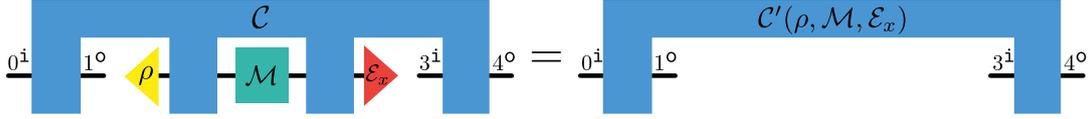


Figure 2.10: Contraction of a two-slot comb with a preparation, a transformation and a measurement. The resulting one-slot comb $\mathcal{C}'(\rho, \mathcal{M}, \mathcal{E}_x)$ depends on the prepared state ρ , the transformation \mathcal{M} and the Effect \mathcal{E}_x (and the original comb \mathcal{C}). We depict the transformation \mathcal{M} in green to distinguish it from the maps that make up the comb \mathcal{C} , and to emphasize that here, we consider the comb as given, and the transformation \mathcal{M} is ‘plugged into it’. For clarity, only the open wires are labeled.

of a comb with the open wires of another object the *contraction* of a comb with said object. Any number of objects can be contracted with each other, as long as the corresponding Hilbert spaces and the temporal orders match. The most general contraction that can be performed is the contraction of a comb (say, \mathcal{C}) with another comb (say, \mathcal{M}), to yield a resulting comb $\mathcal{C}(\mathcal{M})$. This situation is depicted in Fig. 2.11.

Evidently, one could also consider the resulting network a contraction of \mathcal{M} with \mathcal{C} . It depends on the concrete experimental situation, which comb one considers as given, and which one as the one that is ‘plugged into’ said given comb. For example, drawn in the way of Fig. 2.9, state teleportation could be considered as the joining of two combs that Alice and Bob can prepare, respectively, or it could be interpreted as them exploiting a given comb $\Sigma_j \otimes \Sigma_j^*$ with the resources they have – state preparations and measurements. In any given experiment, this question boils down to what degrees an experimenter has control over, and which ones they cannot manipulate. The former degrees of freedom can be used to implement a comb that can be inserted into the given comb, which is made up of the latter degrees of freedom. We will discuss this distinction between environmental (*i.e.*, uncontrollable) degrees of freedom and system (*i.e.*, controllable) degrees of freedom in detail in Ch. 4, where we will employ the comb formalism to describe open quantum system dynamics.¹¹ Before we discuss the mathematical

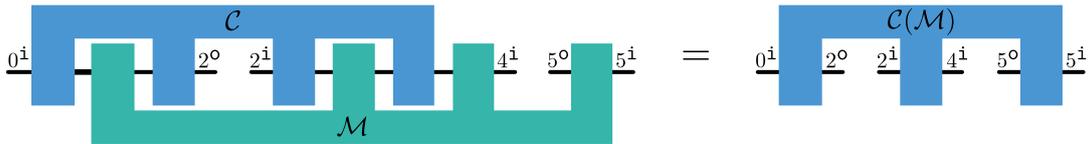


Figure 2.11: Contraction of a comb with another comb. The resulting comb is considered to be a contraction of \mathcal{C} with \mathcal{M} . Equivalently, it could be considered a contraction of \mathcal{M} with \mathcal{C} . The wires are labeled with respect to the comb \mathcal{C} , *i.e.*, i (o) corresponds to input wires (output wires) of \mathcal{C} , but to output wires (input wires) of \mathcal{M} . This convention is continued for the wires that belong to \mathcal{M} only.

structure of combs and quantum networks, we will make their connection to directed acyclic graphs obvious. As these graphs are ubiquitous in the study of causal order [49], this slight detour will prepare the later discussion of temporally ordered processes, and gives further credence to the usage of quantum combs for the analysis of the causal order of processes, that we will carry out in Ch. 6.

¹¹ The discussion in Ch. 4 will mostly be guided by experimental considerations. A very detailed and more rigorous theoretical discussion of the concepts of subsystems and agents can, *e.g.*, be found in [48].

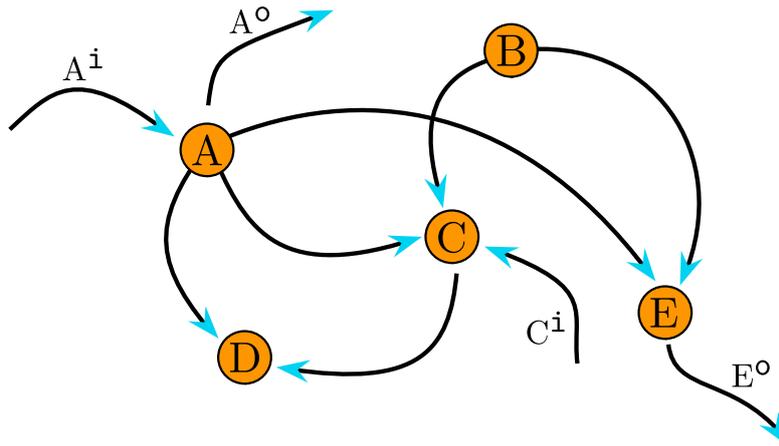


Figure 2.12: Directed acyclic graph (DAG). An arrow between nodes signifies a potential causal influence. For example, A can influence C, but D cannot influence any other node. For clarity, only open arrows are labeled.

2.4 QUANTUM COMBS AND DIRECTED ACYCLIC GRAPHS

Graphically, for a comb, causal ordering means that any operation that is ‘plugged’ into a slot of it can influence the measurement statistics of any measurements performed at slots that are to its right, but not to its left. In this sense, the ordering of the slots provides the order in which information can travel through the network, or, more precisely, tells us what events can signal to other events. So far, we have referred to time as the natural ordering parameter for events; each slot of a comb corresponds to a time t_j and slots that correspond to earlier times can potentially signal to slots at later times. While we shall continue to appeal to time as a natural ordering parameter throughout most of this thesis, we could actually do away with it as a primal concept, and rather employ the more basic idea of causal order. Causal order provides a partial ordering \prec of events, where, as mentioned above, $A \prec B$ implies that events labeled by A can influence events labeled by B , and the labels A and B could correspond to times, laboratories, or locations in space-time.¹² Importantly, if $A \prec B$, then $B \not\prec A$, and, as already mentioned, the relation \prec is transitive. Additionally, we can have $A \not\prec B$ and $B \not\prec A$, *i.e.*, the events A and B are independent of each other.

Such a *partial* order can most easily be represented by a *directed acyclic graph* (DAG), a graph of nodes and directional edges that does not contain closed loops. Evidently, every quantum network has an underlying DAG structure, and ultimately, every DAG can be reshaped to look like the network¹³ of Fig. 2.5. As DAGs are routinely used in the field of causal modeling [49–51], it is worth making this connection explicit.

An example for a DAG with five nodes can be found in Fig. 2.12. Similar to a quantum network, it has arrows that connect events, and open arrows, that correspond to inputs to, and outputs of the DAG. Returning to the above discussion about experimental control, the open arrows/wires can be considered experimentally controllable degrees of freedom,

¹² Whenever there is no risk of confusion, we will often drop the explicit distinction between events and their labels. For example, we will say that $A \prec B$ means that A can influence B

¹³ Potentially, to look exactly the same, one has to absorb preparations and effects into the transformations.

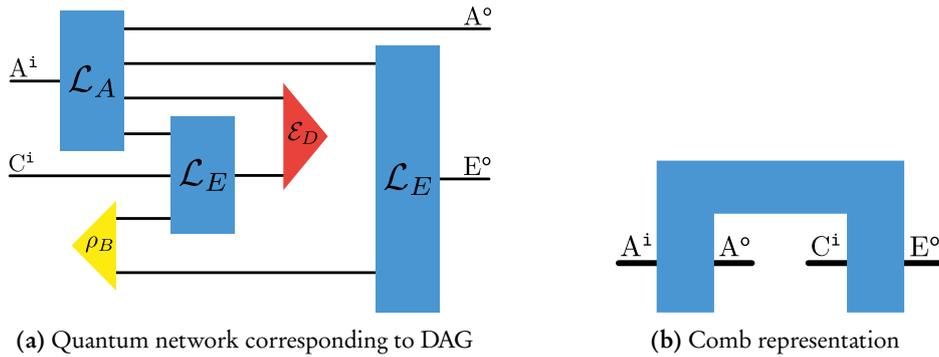


Figure 2.13: Quantum network and comb corresponding to the DAG in Fig. 2.12. (a) Corresponding network, consisting of states ρ , transformations \mathcal{L} and effects \mathcal{E} . In order for the comb to be deterministic, \mathcal{E}_D must correspond to the trace operation (see below). (b) Corresponding comb representation of the DAG of Fig. 2.12.

while all connected arrows/connected wires are out of experimental control. If the DAG represents a quantum mechanical process, each node corresponds to a quantum operation, *i.e.*, a preparation, an effect, a transformation, or a combination thereof. We represent the network 2.12 as a quantum network in Fig. 2.13a, and the corresponding comb in Fig. 2.13b. While DAGs can evidently be used outside of quantum mechanics, we will assume throughout this work that quantum mechanics holds (at least locally), and as such, every DAG has a corresponding quantum network.

More generally, every comb is compatible with a DAG and vice versa [17]. Consequently, combs provide the natural language for quantum (and classical) causal modeling [4, 52]. As we have mentioned, it is not necessary to revert to the concept of time to meaningfully define combs and DAGs. While this distinction between causal order and an arrow of time seems unnecessarily pedantic at this point, keeping it in mind will simplify the generalization of the introduced concepts to situations without predefined global order [52–54], which we will discuss in Ch. 6, and we will make explicit throughout this thesis, whenever causal ordering is not a necessary requirement for results we derive. Having built a graphical intuition of quantum networks, it is now time to introduce an algebraic way to manipulate and contract combs, and express their causal ordering in terms of well-defined constraints on their mathematical structure.

2.5 UNIFIED REPRESENTATION: CHOI-JAMIOŁKOWSKI ISOMORPHISM

Up to this point, we have treated transformations, effects and states on a different mathematical footing. However, employing the Choi-Jamiołkowski Isomorphism (CJI)¹⁴ all of these objects can be considered positive matrices, and their respective properties are encoded in the properties of the corresponding matrices [56, 57]. Analogously, we will see that any comb can be represented as a positive matrix on the correct space, and the causality conditions mentioned

¹⁴ In fact, the Choi and the Jamiołkowski isomorphism differ slightly (see, *e.g.*, [55] for a detailed discussion). The one we shall use throughout this thesis is Choi’s, but we will follow the convention in the literature and call it the Choi-Jamiołkowski isomorphism.

in the previous sections are encoded as trace properties of said matrix. Additionally, this isomorphism allows us to algebraically express concatenation of building blocks of quantum networks in a succinct fashion.

We start by laying out the CJI for quantum maps, and show that states (POVM elements) can be understood as the matrices corresponding to quantum maps with a particular input (output) space. We emphasize that this way of representing quantum maps is not the only one in use, and depending on the respective context, different representations can be more advantageous. Here, the representation in terms of Choi matrices is chosen, as it allows one to consider all building blocks of quantum dynamics as objects of the same type, and their properties can be expressed in the most transparent way in this representation. For a thorough discussion of different representations of quantum maps, and the transformation between them, see for example [8, 13].

Any quantum map $\mathcal{L} : \mathcal{B}(\mathcal{H}^1) \rightarrow \mathcal{B}(\mathcal{H}^0)$ can be mapped isomorphically onto a matrix $L \in \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$ by letting \mathcal{L} act on one half of an (unnormalized) maximally entangled state $|\Phi^+\rangle = \sum_{i=1}^d |i\rangle |i\rangle \in \mathcal{H}^1 \otimes \mathcal{H}^1$, where $\{|i\rangle\}$ is the computational basis of \mathcal{H}^1 :

$$L := \mathfrak{C}[\mathcal{L}] := (\mathcal{L} \otimes \mathcal{I}) [|\Phi^+\rangle\langle\Phi^+|] . \quad (2.10)$$

We will call this matrix the *Choi matrix* or – in slight abuse of nomenclature – *Choi state* of the respective map. As we have already done for the case of effects, we will denote a map and its Choi state by the same letter, the former in calligraphic font, and the latter in upright sans-serif letters. The action of the map \mathcal{L} on a state $\rho \in \mathcal{B}(\mathcal{H}^1)$ can then be expressed in terms of its Choi state as

$$\mathcal{L}[\rho] = \text{tr}_1 \left[\left(\mathbb{1}^0 \otimes \rho^T \right) L \right] , \quad (2.11)$$

where \cdot^T denotes the transposition in the computational basis, tr_1 is the trace over the input space \mathcal{H}^1 and $\mathbb{1}^0$ is the identity matrix on \mathcal{H}^0 . Eq. (2.11) can be verified by direct insertion into Eq. (2.10):

$$\begin{aligned} \text{tr}_1 \left[\left(\mathbb{1}^0 \otimes \rho^T \right) L \right] &= \sum_{i,j} \text{tr}_1 \left[\left(\mathbb{1}^0 \otimes \rho^T \right) \mathcal{L}[|i\rangle\langle j|] \otimes |i\rangle\langle j| \right] \\ &= \sum_{i,j} \langle j| \rho^T |i\rangle \mathcal{L}[|i\rangle\langle j|] = \mathcal{L}[\rho] . \end{aligned} \quad (2.12)$$

By construction, every completely positive map has a corresponding positive Choi matrix and the converse also holds: every positive matrix $L \in \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$ belongs to a map \mathcal{L} that is CP. To see this, we note that, if L is positive, it can be diagonalized as $L = \sum_{\alpha} \lambda_{\alpha} |L_{\alpha}\rangle\langle L_{\alpha}|$, where $\lambda_{\alpha} \geq 0$ and $\langle L_{\alpha}|L_{\beta}\rangle = \delta_{\alpha\beta}$. Using Eq. (2.11) we obtain

$$\begin{aligned} \text{tr}_1 \left[\left(\mathbb{1}^0 \otimes \rho^T \right) L \right] &= \sum_{i,j,\alpha} \lambda_{\alpha} \langle i| \rho^T |j\rangle \langle j|L_{\alpha}\rangle \langle L_{\alpha}|i\rangle \\ &= \sum_{\alpha} \left(\sqrt{\lambda_{\alpha}} \sum_j \langle j|L_{\alpha}\rangle \langle j| \right) \rho \left(\sqrt{\lambda_{\alpha}} \sum_i |i\rangle \langle L_{\alpha}|i\rangle \right) := \sum_{\alpha} L_{\alpha} \rho L_{\alpha}^{\dagger} . \end{aligned} \quad (2.13)$$

Consequently, the action of a map \mathcal{L} that corresponds to a positive Choi matrix L can be written as $\mathcal{L}[\rho] = \sum_{\alpha} L_{\alpha} \rho L_{\alpha}^{\dagger}$. Such a decomposition of the map is called *Kraus decomposition*. Any map

that admits a Kraus decomposition is completely positive [21, 58]. Indeed, let $\rho \in \mathcal{B}(\mathcal{H}^i \otimes \mathcal{H}_a^i)$ be a positive matrix, then there exists $\sqrt{\rho} \geq 0$, such that $\rho = \sqrt{\rho} \sqrt{\rho}$. The action of \mathcal{L} on ρ yields

$$(\mathcal{L} \otimes \mathcal{I}_a)[\rho] = \sum_{\alpha} \left[(L_{\alpha} \otimes \mathbb{1}_a) \sqrt{\rho} \right] \left[\sqrt{\rho} (L_{\alpha}^{\dagger} \otimes \mathbb{1}_a) \right] := \sum_{\alpha} A_{\alpha} A_{\alpha}^{\dagger}. \quad (2.14)$$

Any matrix of the form $A_{\alpha} A_{\alpha}^{\dagger}$ is positive, and so is the sum of positive matrices. Therefore, as \mathcal{H}_a can be of arbitrary size, if \mathcal{L} admits a Kraus decomposition, then it is CP, and \mathcal{L} is completely positive iff its Choi state is positive. This also implies, that in order for \mathcal{L} to be completely positive, it is already sufficient (and necessary) for $\mathcal{L} \otimes \mathcal{I}_d$ to be positive, where \mathcal{I}_d is the identity map on $\mathcal{B}(\mathcal{H}^i)$. In this case the Choi state L is positive, and \mathcal{L} is hence completely positive. For completeness, we mention that the Kraus decomposition of a CP map is not unique, but there is a minimal number of Kraus operators necessary to represent a map. This number – the *Kraus rank* of \mathcal{L} – coincides with the rank of the Choi state L [59].

Besides being CP, quantum channels also preserve trace. Trace preservation of \mathcal{L} means that $\text{tr}(\mathcal{L}[\rho]) = \text{tr}(\rho)$ for all ρ . In terms of Choi states, this requirement translates to

$$\text{tr}(\rho) = \text{tr} \left[\left(\mathbb{1}^o \otimes \rho^T \right) L \right] = \text{tr} \left(\rho^T \text{tr}_o(L) \right). \quad (2.15)$$

As this equality has to hold for all ρ , it implies $\text{tr}_o(L) = \mathbb{1}^i$, where tr_o denotes the trace over \mathcal{H}^o and $\mathbb{1}^i$ is the identity matrix on \mathcal{H}^i . Equivalently, a CP map is trace-preserving iff its Kraus operators satisfy $\sum_{\alpha} L_{\alpha}^{\dagger} L_{\alpha} = \mathbb{1}^i$. With this, we obtain the two properties that define a CPTP map in terms of its Choi state. A map $\mathcal{L} : \mathcal{B}(\mathcal{H}^i) \rightarrow \mathcal{B}(\mathcal{H}^o)$ is CPTP iff its Choi state $L \in \mathcal{B}(\mathcal{H}^o \otimes \mathcal{H}^i)$ satisfies

$$L \geq 0 \quad \text{and} \quad \text{tr}_o(L) = \mathbb{1}^i. \quad (2.16)$$

In what follows – whenever there is no risk of confusion – we drop the explicit distinction between a map and its Choi state both in text and figures.

As we will see below, the requirement of trace preservation can also be interpreted as a causality constraint in the sense, that measurements on the outcome space cannot influence the statistics on the input space. Additionally, if a map is trace-preserving, it means that it can be implemented deterministically; the probability of implementation of a map is given by the trace of its output state (see below), and as $\text{tr}(\mathcal{L}[\rho]) = 1$ for every input state of unit trace, it follows that the probability of implementing a trace-preserving map is 1. Before we give some examples of Choi states of quantum maps, we introduce a useful decomposition of Choi states L of CPTP maps that we will rely on heavily in Ch. 6. As L is positive (and also Hermitian), it can be represented as

$$L = \frac{1}{d_o} \mathbb{1}^o \otimes \mathbb{1}^i + \sum_{\alpha=1}^{d_i^2-1} a_{\alpha} \Gamma_{\alpha}^o \otimes \mathbb{1}^i + \sum_{\beta=1}^{d_i^2-1} \sum_{k=1}^{d_o^2-1} c_{\beta k} \Gamma_{\beta}^o \otimes \Gamma_k^i, \quad (2.17)$$

where $a_{\alpha}, c_{\beta k} \in \mathbb{R}$, $d_x = \dim(\mathcal{H}^x)$ and Γ_b^x are generalized Pauli matrices, *i.e.*, $\Gamma_b^{x\dagger} = \Gamma_b^x$, $\text{tr}(\Gamma_b^x) = 0$ and $\text{tr}(\Gamma_b^x \Gamma_c^x) = \delta_{bc}$ [60]. These matrices, together with the identity matrix form an orthogonal (with respect to the Hilbert-Schmidt inner product) basis of the set of Hermitian matrices. Importantly, there are no terms of the form $\mathbb{1}^o \otimes \Gamma_m^i$ in the decomposition (2.17), as

otherwise the map would not be trace-preserving. We will use the absence of these terms below to prove structural statements about causally ordered combs, as well as causally indefinite processes (see Ch. 6 and App. D.1).

2.6 CHOI-JAMIOŁKOWSKI ISOMORPHISM: EXAMPLES

It is instructive to provide some examples of Choi states of particular maps that we will encounter regularly throughout this thesis. Firstly, the identity map \mathcal{I} corresponds to the unnormalized maximally entangled state

$$\mathfrak{C}[\mathcal{I}] = |\Phi^+\rangle\langle\Phi^+| := \Phi^+. \quad (2.18)$$

As already mentioned, to keep better track of the respective spaces, we consider \mathcal{I} to be a map from $\mathcal{B}(\mathcal{H}^1)$ to $\mathcal{B}(\mathcal{H}^0)$, with $\mathcal{H}^1 \cong \mathcal{H}^0$, and as such $\Phi^+ \in \mathcal{B}(\mathcal{H}^1 \otimes \mathcal{H}^0)$ on the RHS of Eq. (2.18). More generally, any unitary map \mathcal{U} corresponds to an (unnormalized) maximally entangled state

$$\mathbf{U} = (\mathcal{U} \otimes \mathcal{I}) [|\Phi^+\rangle\langle\Phi^+|]. \quad (2.19)$$

Naturally, we have $\text{tr}_0(\mathbf{U}) = \mathbb{1}^1$, and it is easy to see, that the only trace-preserving maps with pure Choi states are unitaries.

Another kind of CP operations that we will encounter frequently are projections. A projection collapses the state ρ onto a pure state Ψ_α corresponding to the outcome α of a (sharp) measurement. This outcome occurs with probability $\langle\Psi_\alpha|\rho|\Psi_\alpha\rangle$. In contradistinction to effects, we consider projections \mathcal{P}_α to ‘feed forward’ the projected state, *i.e.*, $\mathcal{P}_\alpha[\rho] = \langle\Psi_\alpha|\rho|\Psi_\alpha\rangle|\Psi_\alpha\rangle\langle\Psi_\alpha|$, and the corresponding Choi state is given by

$$\begin{aligned} \mathbf{P}_\alpha &= \sum_{ij} |\Psi_\alpha\rangle\langle\Psi_\alpha| \otimes \langle\Psi_\alpha|i\rangle \langle j|\Psi_\alpha\rangle |i\rangle\langle j| \\ &= |\Psi_\alpha\rangle\langle\Psi_\alpha| \otimes |\Psi_\alpha^*\rangle\langle\Psi_\alpha^*| := \Psi_\alpha \otimes \Psi_\alpha^*. \end{aligned} \quad (2.20)$$

As before, we consider the output state (Ψ_α in the above equation) to live on $\mathcal{B}(\mathcal{H}^1)$ and the input state (Ψ_α^* in the above equation) to live on $\mathcal{B}(\mathcal{H}^0)$, with the understanding $\mathcal{H}^1 \cong \mathcal{H}^0$. As such, a projective measurement can be understood as an effect with outcome α , followed by a preparation of the pure state Ψ_α that corresponds to said outcome.

In general, the output space of a map does not have to coincide with its input space. We have already encountered effects $\mathcal{E}_\alpha : \mathcal{B}(\mathcal{H}^1) \rightarrow \mathbb{R}$, that map quantum states to real numbers (probabilities) and ‘destroy’ the state in the process. Consequently, they are quantum maps with a trivial output space, and as such, positivity of effects is equivalent to complete positivity. In terms of their Choi states, the action of an effect is given by

$$\mathcal{E}_\alpha = \text{tr} \left(\rho^T \mathbf{E}_\alpha \right), \quad (2.21)$$

which, unsurprisingly, corresponds to Born’s rule (2.2), and the Choi states $\mathbf{E}_\alpha \in \mathcal{B}(\mathcal{H}^1)$ of effects are – up to a transpose – equal to the corresponding POVM elements.¹⁵ The fact that the

¹⁵ We could have defined POVM elements or the Choi isomorphism with an additional transpose to make POVM elements exactly equal, but we rather adhered to convention.

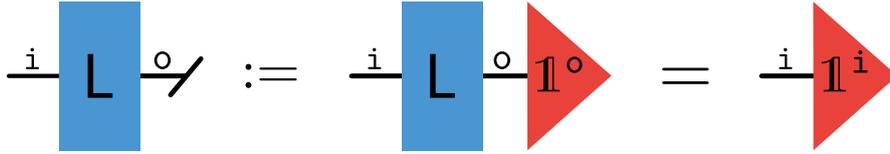


Figure 2.14: Graphical representation of the trace preservation property. Tracing over the output space of a trace-preserving map \mathcal{L} (denoted by $\mathbb{1}^o$) yields an identity $\mathbb{1}^i$ on the input space. In what follows, we will often also use the slash to represent the trace operation.

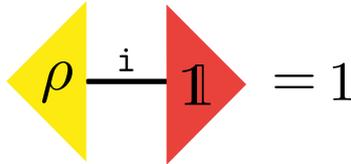


Figure 2.15: Graphical representation of the unit trace of a quantum state as a trace preservation condition.

output space of effects is trivial implies that there is only one trace-preserving effect, namely $E = \mathbb{1}^i$, which corresponds to the trace operation $\mathcal{E}[\rho] = \text{tr}[\rho]$.

Additionally, this is the only effect, that can be implemented deterministically; for any given state ρ , the probability to implement an effect \mathcal{E}_α (or, equivalently, to obtain outcome α when using instrument \mathcal{J}) is given by $\text{tr}(\rho^T E_\alpha)$, and as such, the only effect that can be implemented with certainty, independent of ρ , is $\mathbb{1}$. This representation of the trace operation in terms of the identity matrix, can be used to depict both the trace preservation condition for quantum maps, as well as the unit trace of quantum states, in a straightforward manner (see Figs. 2.14 and 2.15, respectively). This will prove helpful in the derivation of properties of general combs.

Finally, a quantum state ρ can be considered the Choi matrix of a CPTP map $\mathcal{R} : \mathbb{R} \rightarrow \mathcal{B}(\mathcal{H}^i)$, and consequently, the normalization $\text{tr}_i(\rho) = 1$ can be understood as a trace preservation condition¹⁶. To adhere to earlier convention, the output space of \mathcal{R} is labeled by an i . Unlike effects, there are many trace-preserving states – all quantum states – which is intuitively clear, since all quantum states can be prepared deterministically.

Thus, the Choi isomorphism puts all the building blocks of quantum dynamics on the same mathematical footing. Each of them can be considered a positive matrix with particular positivity and trace conditions. In this light, we could have introduced all of these objects as Choi operators of quantum maps in the first place. However, we refrained from doing so, in favor of a more intuitive introduction, and to provide an easier connection to the literature. From now on, we will use *both* the concepts of maps and Choi states to describe dynamics and quantum networks, depending on the context. For example, the communication scheme of Fig. 2.4 can be considered as a concatenation of the preparation \mathcal{R}_X that prepares the state ρ_X , the channel $\mathcal{L} \otimes \cdots \otimes \mathcal{L}$, and the effect \mathcal{E}_x , *i.e.*,

$$\mathbb{P}(x|\mathcal{J}, X) = \mathcal{E}_x \circ (\mathcal{L} \otimes \cdots \otimes \mathcal{L}) \circ \mathcal{R}_X. \quad (2.22)$$

¹⁶ If we strictly adhered to our notational convention, here, quantum states would have to be denoted by R . However, we choose to abide by standard convention.

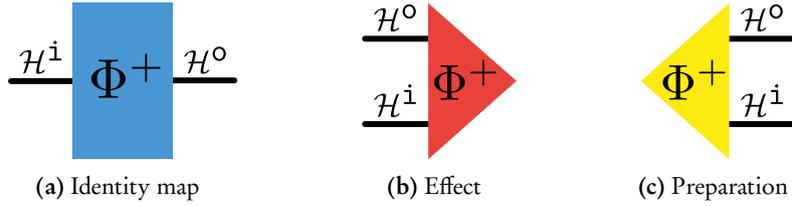


Figure 2.16: Difference between the identity map, a measurement outcome that corresponds to the effect Φ^+ , and a preparation of Φ^+ . Up to normalization, all of these building blocks have the same Choi state $\Phi^+ \in \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$. However, the output space of the identity map is $\mathcal{B}(\mathcal{H}^0)$, the one of the preparation is $\mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$, while that of the effect is trivial. Out of the three, only the identity map and the preparation are trace-preserving. The types of building blocks depicted in each panel are immediately clear from their respective shape and orientation.

Equivalently, this scheme can be described as a contraction of the Choi states of the involved maps:

$$\mathbb{P}(x|\mathcal{J}, X) = E_x \star L \star \cdots \star L \star \rho_X, \quad (2.23)$$

where \star denotes the *link product* [16, 47], which we will introduce in Sec. 2.8 as the operation to compute the Choi state of an arbitrary network of states, transformations and preparations. Before we introduce this powerful mathematical tool, it is necessary to provide an intuitive explanation of the mathematical properties of quantum maps, that will lead us directly to the causality constraints that are imposed on general combs \mathcal{C} .

REMARK 2.1: *Before we advance, a word of caution is in order to avoid confusion. A priori, for a given Choi state $A \in \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$ it is unclear if it corresponds to a map $\mathcal{A} : \mathcal{B}(\mathcal{H}^1) \rightarrow \mathcal{B}(\mathcal{H}^0)$, a preparation $\mathcal{A} : \mathbb{R} \rightarrow \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$, or an effect $\mathcal{A} : \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1) \rightarrow \mathbb{R}$. For example, we have already encountered the (unnormalized) maximally entangled state as the Choi state of the identity channel, as well as the POVM element of a measurement in the Bell basis (see Fig. 2.16). While this ambiguity exists, it will not be a problem in practice, as it will always be clear, both from the ordering of the wires in the corresponding quantum network, and the described underlying physical situation, what type of transformation the respective Choi state corresponds to.*

2.7 TRACE PRESERVATION, CAUSAL ORDERING, AND DETERMINISTIC IMPLEMENTATION

Throughout this thesis, we will frequently discuss causal order, and the deterministic implementation of operations. Both of these properties can be directly related to the properties of Choi states. We have already discussed, that the probability for the implementation of an effect E_α is equal to the probability to obtain the measurement outcome α corresponding to said effect

$$\mathbb{P}(\alpha|\mathcal{J}) = \text{tr}(\rho^T E_\alpha), \quad (2.24)$$

and the only effect that can be implemented with unit probability, independent of the quantum state it acts on, is the trace operation $E = \mathbb{1}^\dagger$. This is also reflected by the condition that all the

POVM elements for a given instrument have to satisfy $\sum_{\alpha} E_{\alpha}^{\text{T}} = \mathbb{1}$; we can be sure that in each run, one of the possible outcomes has to occur.

Analogously, instead of discarding the state after measuring it, we could keep the output state. In this case, each outcome α , given an instrument \mathcal{J} , corresponds to a (trace non-increasing) CP map L^{α} that describes¹⁷ the transformation of the state during measurement, *i.e.*, $\rho' = \text{tr}_1 [(\mathbb{1}_0 \otimes \rho^{\text{T}}) L^{\alpha}]$, where ρ' is a subnormalized quantum state. We deliberately write the label of the outcome as a superscript to distinguish it from the labels for times, that are indicated by subscripts. The probability to implement a CP operation L^{α} is then given by

$$\mathbb{P}(\alpha|\mathcal{J}) = \text{tr}(\rho') = \left[(\mathbb{1}^{\circ} \otimes \rho^{\text{T}}) L^{\alpha} \right] = \text{tr} \left[\rho^{\text{T}} \text{tr}_0 (L^{\alpha}) \right]. \quad (2.25)$$

In general, a CP map cannot be implemented deterministically, and the implementation probability for a given input state ρ can be computed via Eq. (2.25). This equation also tells us that a CP map L^{α} can be implemented deterministically independent of the input state iff $\text{tr}_1 (L^{\alpha}) = \mathbb{1}$, *i.e.*, iff it is trace-preserving. Additionally, by comparison of (2.25) and (2.24), we see the connection between CP maps L^{α} and their corresponding POVM elements: $E_{\alpha} = \text{tr}_0 (L^{\alpha})$; CP maps provide both the outcome probability of a general measurement, as well as the update of the interrogated state upon measurement. On the other hand, POVM elements only provide probabilities, and do not allow one to infer the post-measurement state. As one of the possible outcomes of a measurement has to occur, we can now provide the general definition of an instrument [61, 62]:

DEFINITION 2.1: *An instrument \mathcal{J} is a collection of CP maps $\{L^{\alpha}\}$ that add up to a CPTP map L , *i.e.*, $\sum_{\alpha} L^{\alpha} = L$. Each of the CP maps L^{α} corresponds to one of the possible outcomes of the instrument, and the CPTP map L describes the overall transformation of a state upon interrogation by the instrument \mathcal{J} .*

We emphasize, that a POVM $\{E_{\alpha}\}$ satisfies this definition of an instrument, however, all the CP maps it consists of are effects, *i.e.*, have a trivial output space. We will further refine this definition of an instrument in the following chapters, but for the moment, Def. 2.1 is sufficient. As we will be interested in scenarios where a system of interest can be measured sequentially, *i.e.*, at different points in time, we need the general understanding of instruments provided by Def. 2.1, to be able to account for such experimental situations.

Finally, beyond a requirement for deterministic implementation, trace preservation can be understood as a causality constraint. To see this, consider the experimental situation where Alice prepares states ρ^i with probability p_i and sends them to Bob via a channel \mathcal{K} . Bob uses some instrument \mathcal{J} to measure the states he receives. The probability for Alice to prepare ρ^i and Bob to measure outcome α is given by

$$\mathbb{P}(i, \alpha) = p_i \text{tr} \left[\left(E_{\alpha}^{\text{T}} \otimes (\rho^i)^{\text{T}} \right) \mathcal{K} \right], \quad (2.26)$$

where $\{E_{\alpha}^{\text{T}}\}$ are the POVM elements of Bob's instrument. Causal ordering of this experiment implies that, overall, Alice's preparation probabilities $\{p_i\}$ have to be independent of Bob's

¹⁷ Due to spatial restrictions – there is only so much space around a letter – the positioning of sub- and superscript will be somewhat inconsistent throughout this thesis. Here, L^{α} denotes the CP map corresponding to outcome α (for a given instrument). We will make clear, whenever the sub- and superscripts change their meaning.

measurement outcomes, *i.e.*, $\sum_{\alpha} \mathbb{P}(i, \alpha) = p_i$. As $\sum_{\alpha} E_{\alpha} = \mathbb{1}$, this requirement can only be satisfied independent of the states and probabilities that Alice chooses, iff K is trace-preserving.

When building up quantum networks, the properties of the individual building blocks translate to the properties of the resulting comb C . We will now see how to compute the Choi state of quantum networks, and investigate their properties in detail.

2.8 CONCATENATING BUILDING BLOCKS – THE LINK PRODUCT

Up to this point, we have seen that all the fundamental building blocks of quantum mechanics can be treated within a unified mathematical framework in terms of their Choi states. What is missing so far, is a *composition rule*, that tells us how to compute the Choi state of an arbitrary concatenation of preparations, transformations, and effects. This composition rule is given by the *link product* \star [16, 47], that yields the Choi state of concatenations of quantum maps. For example, if \mathcal{L} and \mathcal{K} are two CP maps, the Choi state of the map $\mathcal{M} = \mathcal{L} \circ \mathcal{K}$ is given by

$$M = \mathfrak{C}[\mathcal{M}] = \mathfrak{C}[\mathcal{L} \circ \mathcal{K}] := L \star K. \quad (2.27)$$

We have already encountered link products in earlier sections. For instance, the action of a channel \mathcal{L} on a quantum state ρ can be considered as a concatenation $\mathcal{L} \circ \mathcal{R}$ of a preparation and a transformation, and the resulting state ρ' is

$$\rho' = \mathfrak{C}[\mathcal{L} \circ \mathcal{R}] = L \star \rho = \text{tr}_i \left[(\mathbb{1}^{\circ} \otimes \rho) L^{\top_i} \right], \quad (2.28)$$

where $\rho \in \mathcal{B}(\mathcal{H}^{\dagger})$, $L \in \mathcal{B}(\mathcal{H}^{\circ} \otimes \mathcal{H}^{\dagger})$, and \cdot^{\top_i} is a transpose with respect to \mathcal{H}^{\dagger} . Eq. (2.28) contains all the ingredients of the general link product: the Choi states of the involved operators are tensored with identity matrices, such that they are of the same dimension and can be multiplied (in this case, ρ is tensored with $\mathbb{1}^{\circ}$); one of the Choi states is transposed with respect to the intersection of Hilbert spaces that both of them are defined on (in this case \mathcal{H}^{\dagger}); and finally the trace is taken with respect to said intersection of Hilbert spaces. Consequently, we have the general definition of the link product, due to Chiribella *et al.* [17]:

DEFINITION 2.2: *Let ℓ and k be two sets, and let $L \in \mathcal{B}(\otimes_{\ell_i \in \ell} \mathcal{H}_{\ell_i})$ and $K \in \mathcal{B}(\otimes_{k_j \in k} \mathcal{H}_{k_j})$. The **link product** $L \star K$ is given by*

$$L \star K = \text{tr}_{\ell \cap k} \left[\left(\mathbb{1}_{k \setminus \ell} \otimes L^{\top_{\ell \cap k}} \right) \left(K \otimes \mathbb{1}_{\ell \setminus k} \right) \right], \quad (2.29)$$

where $x \setminus y = \{i \mid i \in x \text{ and } i \notin y\}$, $\mathbb{1}_{x \setminus y}$ is the identity matrix on $\otimes_{i \in x \setminus y} \mathcal{H}_i$, and $\cdot^{\top_{\ell \cap k}}$ denotes the transpose on the space $\otimes_{i \in \ell \cap k} \mathcal{H}_i$.

By direct insertion, it can be shown, that the link product indeed yields the correct Choi state of the concatenation of two maps, *i.e.*, $\mathfrak{C}[\mathcal{L} \circ \mathcal{K}] = L \star K$ (see [16], Thm. 1). While the algebraic definition is somewhat clunky, it can be easily motivated in graphical terms (see Fig. 2.17). The Hilbert spaces labeled by $\ell \cap k$ correspond to wires that connect the maps \mathcal{L} and \mathcal{K} , while $\ell \setminus k$ ($k \setminus \ell$) labels remaining open wires of the map \mathcal{L} (\mathcal{K}). As is the case of Choi states, the link product $L \star K$ can describe very different networks, depending on what kinds of maps the matrices L and K correspond to. Throughout this thesis, the underlying physical

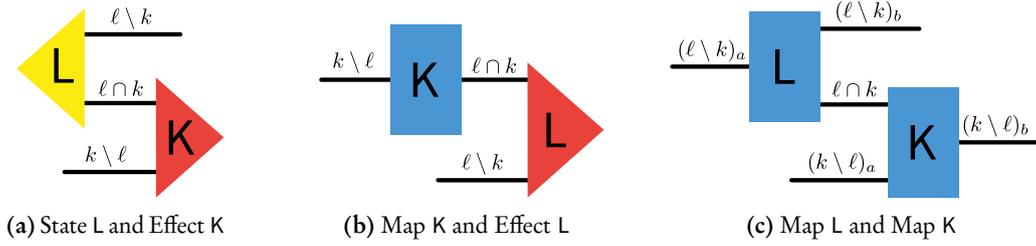


Figure 2.17: Different networks corresponding to the same link product $L \star K$. Note that the ordering of the link product $L \star K$ does not necessarily correspond to the ordering of the underlying network. In (c), we have $(x \setminus y)_a \cup (x \setminus y)_b = x \setminus y$.

situation of a link product will always be clear from context and depicted graphically. Before we discuss its properties in detail, it is instructive to discuss some basic examples of the link product to build intuition. Most of these examples are inspired by [16].

We have already seen that the link product of a channel $L \in \mathcal{B}(\mathcal{H}_0 \otimes \mathcal{H}_1)$ and a state $\rho \in \mathcal{B}(\mathcal{H}_1)$ corresponds to

$$L \star \rho = \mathcal{L}[\rho]. \quad (2.30)$$

Analogously, the link product of an effect $E_\alpha \in \mathcal{B}(\mathcal{H}_1)$ and a state $\rho \in \mathcal{B}(\mathcal{H}_1)$ reproduces the Born rule:

$$E_\alpha \star \rho = \text{tr}(\rho E_\alpha^T), \quad (2.31)$$

which also implies

$$\mathbb{1} \star \rho = \text{tr}(\rho). \quad (2.32)$$

If two Choi states L and K are defined on disjoint Hilbert spaces, *i.e.*, $l \cap k = \emptyset$, the link product reduces to the tensor product

$$L \star K = L \otimes K. \quad (2.33)$$

The properties of the link product follow directly from its definition [16]: it is commutative up to a unitary swap \mathbf{S} on $\mathcal{H}_{l \setminus k} \otimes \mathcal{H}_{k \setminus l}$, *i.e.*, $L \star K = \mathbf{S}(K \star L)\mathbf{S}^\dagger$. As \mathbf{S} simply corresponds to a relabeling of Hilbert spaces, it will henceforth be implied, but omitted in equations. In particular, commutativity of the link product means that the order of the Choi states in the product is not an indicator of the causal order of the underlying network (see Fig. 2.17). Rather, causal ordering will follow from the properties of the individual terms in the link product (see below).

The link product of two Hermitian matrices L and K is also Hermitian, and the link product of two positive matrices is positive. In particular the latter property is intuitively clear, as positive matrices correspond to CP maps, and the resulting link product is the Choi state of a concatenation of CP maps which, naturally, has to be positive. Finally, for all intents and purposes – *i.e.*, in all cases relevant to this thesis – the link product is associative. In detail, if the

sets m, ℓ and k labeling the Hilbert spaces that M, L and K are defined on satisfy $m \cap \ell \cap k = \emptyset$, then

$$(M \star L) \star K = M \star (L \star K). \quad (2.34)$$

As mentioned, for all the cases treated in this thesis, this requirement will be satisfied, so we will assume associativity of the link product from now on. With this product at hand, the Choi states corresponding to any quantum network can simply be expressed as the link product of all involved building blocks. For example, the Choi state corresponding to the communication scheme of Fig. 2.4, *i.e.*, the probability $\mathbb{P}(x|\mathcal{J}, X)$ to measure x , given that the message X was transmitted, can be expressed as

$$\mathbb{P}(x|\mathcal{J}, X) = E_x \star L \star \cdots \star L \star \rho_X, \quad (2.35)$$

where it is understood that all the instances of L are defined on disjoint Hilbert spaces (*i.e.*, $L \star \cdots \star L = L \otimes \cdots \otimes L$). Analogously, the state ξ_{A^0} resulting from the teleportation scheme of Fig. 2.3 is given by

$$\xi_{A^0} = \Phi_{B_1^0 B_2^1}^+ \star \sum_j^{B_1^0 B_1^1} \star \sum_j^{A^0 A^1} \star \Phi_{AB_1^1}^+ \star \Psi_{B_2^1}, \quad (2.36)$$

where, for better orientation, we have labeled the different terms by the spaces they live on, respectively, and the corresponding labels are placed so that their positioning agrees with the previous notation.¹⁸ Finally, the general contraction of two combs C and M (see Fig. 2.11) is simply given by

$$C'(M) = C \star M, \quad (2.37)$$

which demonstrates the notational ease with which general quantum networks can be described using the link product; while its explicit computation necessitates some diligence in the labeling of the involved Hilbert spaces, link products provide a unified, concise, and straightforward way to express the Choi states of quantum networks, and derive their properties.

2.9 DETERMINISTIC QUANTUM NETWORKS AND CAUSAL ORDER

As we have discussed in Sec. 2.7, the trace preservation condition of a map is equivalent to the possibility of its deterministic implementation.¹⁹ In light of this, we will call a quantum network – or rather the comb C that represents it – *deterministic* or *causally ordered*, if it only consists of trace-preserving building blocks. Otherwise, depending on the context, we will call it *probabilistic* or *causally disordered*. Trace preservation of the building blocks induces a hierarchy of trace conditions on the Choi states of deterministic networks.

Employing the link product, we can compute the Choi state $C^{(N)}$ of a quantum network consisting of N channels $\{L_j\}_{j=0}^{N-1}$. The corresponding network will have open wires, that we

¹⁸ It is somewhat inconsistent to denote the respective spaces as sub- as well as superscripts. However, in not doing so, we would break with earlier notation, so it seems the lesser evil to choose a slightly inconsistent notation and mention the inconsistency explicitly. We will have to opt for this solution in several places in this thesis.

¹⁹ Deterministic maps can also be implemented in a probabilistic fashion. We will present one possible scheme to do so in Ch. 6.

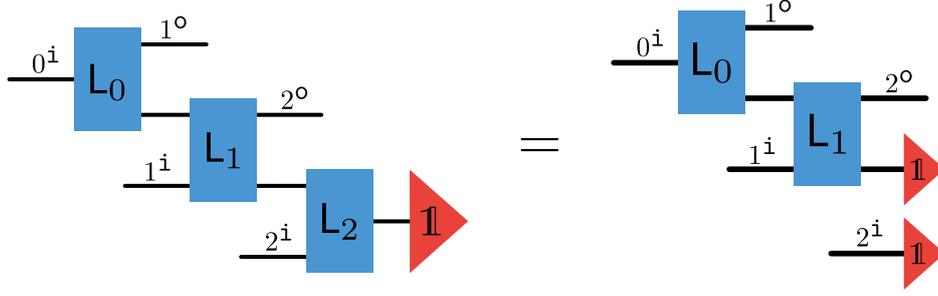


Figure 2.18: Deterministic quantum network with final output for $N = 3$. All the maps L_i are trace-preserving. Tracing over the final output space 3^o (depicted by contraction with the identity matrix in the left panel) leaves an identity on 2^i and a smaller remaining network. The unlabeled wires that connect the respective maps correspond to the Hilbert spaces denoted with a superscript t in the text.

label by $\{0^i, 1^o, \dots, N-1^i, N^o\}$. In accordance with the convention of Sec. 2.3, we assume an ordering of events

$$0^i \prec 1^o \prec 1^i \prec \dots \prec N-1^i \prec N^o, \quad (2.38)$$

with the additional understanding, that all wires labeled by an i point to the left, and all those labeled with an o point to the right. As the respective Hilbert spaces can be trivial (*i.e.*, there is no corresponding wire in the graphical representation), this scenario is a fully general causally ordered quantum network. Now, let $L_j \in \mathcal{B}(\mathcal{H}_{j+1}^o \otimes \mathcal{H}_{j+1}^t \otimes \mathcal{H}_j^i \otimes \mathcal{H}_j^t)$. For $x \neq 0, N$, wires that correspond to \mathcal{H}_x^t connect different maps, and are contracted over. For $x = 0$ and $x = N$, we have $\mathcal{H}_0^t \cong \mathcal{H}_N^t \cong \mathbb{R}$ (see Fig. 2.18). The comb $\mathcal{C}^{(N)}$ of such a network of CPTP maps is given by

$$\mathcal{C} = L_0 \star L_1 \star \dots \star L_{N-1} := \bigstar_{j=0}^N L_j. \quad (2.39)$$

Trace preservation of the maps L_j implies that a trace over the final output space \mathcal{H}_N^o yields an identity matrix on $\mathcal{H}_{N-1}^i \otimes \mathcal{H}_{N-1}^t$, *i.e.*,

$$\mathcal{C}^{(N)} \star \mathbb{1}_{N^o} = L_0 \star L_1 \star \dots \star L_{N-2} \star \mathbb{1}_{N-1^t} \star \mathbb{1}_{N-1^i} := \mathcal{C}^{(N-1)} \star \mathbb{1}_{N-1^i}. \quad (2.40)$$

A graphical representation of this relation can be found in Fig. 2.18 for the case $N = 3$. Evidently, these trace conditions trickle down and hold for every $\mathcal{C}^{(j)}$ ($j \leq N$):

$$\mathcal{C}^{(j)} \star \mathbb{1}_{j^o} = \mathcal{C}^{(j-1)} \star \mathbb{1}_{j-1^i} \Leftrightarrow \text{tr}_{j^o}[\mathcal{C}^{(j)}] = \mathbb{1}_{j-1^i} \otimes \mathcal{C}^{(j-1)}. \quad (2.41)$$

This leads to the general properties of combs that correspond to a deterministic quantum network [17]:

THEOREM 2.1 (DETERMINISTIC COMBS): *A matrix $\mathcal{C}^{(N)} \in \mathcal{B}(\mathcal{H}_N^o \otimes \mathcal{H}_{N-1}^i \otimes \dots \otimes \mathcal{H}_1^o \otimes \mathcal{H}_0^i)$ corresponds to a deterministic quantum network iff $\mathcal{C}^{(N)} \geq 0$ and it satisfies the hierarchy of trace conditions*

$$\text{tr}_{j^o}[\mathcal{C}^{(j)}] = \mathbb{1}_{j-1^i} \otimes \mathcal{C}^{(j-1)}, \quad \forall j \leq N. \quad (2.42)$$

Combs of this form are called deterministic N -combs.

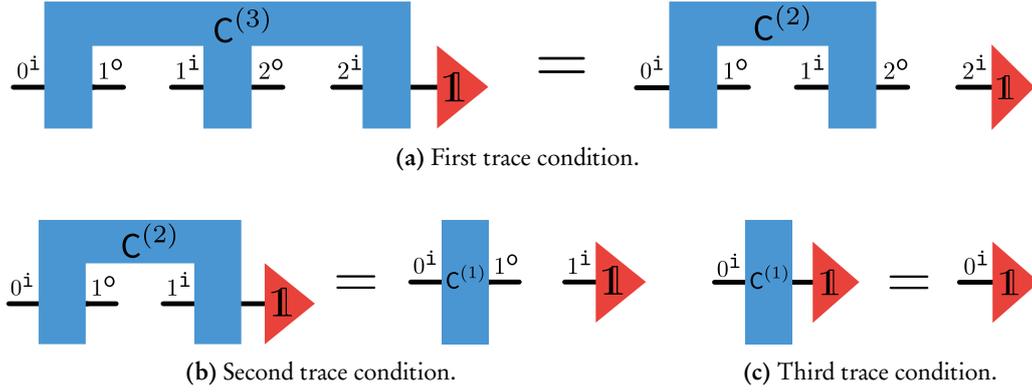


Figure 2.19: Graphical representation of the hierarchy of trace conditions for a deterministic comb $C^{(3)}$. The generalization to more times follows analogously.

From the above discussion, it follows directly that every comb that corresponds to a deterministic network satisfies Eq. (2.42). The converse statement, that every comb that complies with the requirements of Eq. (2.42) can actually be represented by a causally ordered network – and thus by a quantum circuit (see Sec. 4.5) – can be proven by induction. The full proof is provided in [17].

A graphical representation for the hierarchy (2.42) of trace conditions for the case $N = 3$ is given in Fig. 2.19, where the network is – in contrast to the equivalent network in Fig. 2.18 – reshaped in a comb form to underline the causal ordering of events more clearly. In what follows, we will mostly opt for this graphical representation, as it fits in more naturally with the description of open quantum system dynamics, and we will frequently name combs by the number of ‘slots’ rather than the number of output spaces they are defined on. For example the comb $C^{(3)}$ in the top panel of Fig. 2.19 would be a two-slot comb. Since we will always furnish our examples with corresponding graphical representations, there will in general be no risk of confusion.

As already mentioned, each of the Hilbert spaces $\{\mathcal{H}_N^o, \mathcal{H}_{N-1}^i, \dots, \mathcal{H}_1^o, \mathcal{H}_0^i\}$ on which a given comb is defined could be trivial. Thm. 2.1 is therefore fully general and applies to combs corresponding to any conceivable kind of quantum network. Nonetheless, with later subject matters in mind, it is helpful to build some explicit intuition for exemplary cases. Firstly, if the final output space \mathcal{H}_N^o is trivial, *i.e.*, $\mathcal{H}_N^o \cong \mathbb{R}$, we have

$$C^{(N)} = \mathbb{1}_{N-1^i} \otimes C^{(N-1)}, \quad (2.43)$$

which is represented graphically in Fig. 2.20. While this fact can be read off from Eq. (2.42), it can also be inferred directly from the fact that the only trace-preserving effect $E \in \mathcal{B}(\mathcal{H}_N^o)$ is $E = \mathbb{1}_{N^o}$. Secondly, if the initial input space \mathcal{H}_0^i is trivial, we have $C^{(1)} = \rho \in \mathcal{B}(\mathcal{H}_1^o)$, which is represented graphically in Fig. 2.21.

Finally, the set of deterministic combs is well-behaved under composition. This means, that the link product of two combs that have a compatible ordering yields a deterministic comb. We have already seen an example of this in Fig. 2.11, where we contracted a three-slot comb C with another three-slot comb M to obtain a two-slot comb. If the ordering of both combs is



Figure 2.20: Structure of a deterministic comb with trivial final output space. Due to the causality constraints, it decomposes into a tensor product of an identity matrix on the final input space and a causally ordered comb on the remaining spaces.

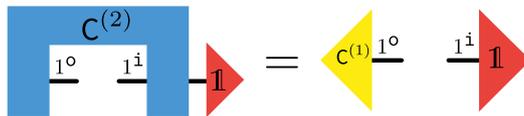


Figure 2.21: Structure of a comb with trivial initial input space. The comb $C^{(1)}$ corresponds to a quantum state $\rho \in \mathcal{B}(\mathcal{H}_1^o)$

compatible (*i.e.*, time flows in the same direction for both combs), then the resulting comb is deterministic. This can be seen directly from the definition of deterministic combs, and the properties of the link product.

On the other hand, contracting combs with incompatible temporal ordering will, in general, not yield physically meaningful results, as it is paramount to creating logically inconsistent closed timelike curves. For example, linking a comb with order $A \prec B$ with a comb that corresponds to an ordering $B \prec A$, would allow to send signals from A to B and from B to A , which, in general, will lead to logical paradoxes. We will discuss in Ch. 6, how to relax the requirement of global causal order, without creating paradoxical situations.

REMARK 2.2: While we followed a constructive approach in terms of quantum networks to arrive at the general definition of a deterministic comb, they can also be introduced based on axiomatic considerations, as the most general admissible quantum maps [17]. Here, admissible means, that they yield a proper quantum output when given a proper quantum input (or a part thereof). In this sense, channels are admissible, as they map quantum states to quantum states, even if they act non-trivially only on a part of the state. The next higher-order map would be a map that, when given a CPTP map as an input yield a CPTP map as an output (such maps are called quantum supermaps [15]). Analogously, a map would be admissible if it took such a supermap – or a part thereof – as an input, and, *e.g.*, yielded a CPTP map as an output. Following this reasoning, one obtains a hierarchy of maps, but it can be shown that this hierarchy collapses at the level of combs, implying that quantum combs are the most general admissible transformations in quantum mechanics.²⁰ Intuitively, this makes direct sense. For example, a supermap is simply a comb with one slot and an initial input and final output wire. A map that maps such a supermap to an object with, say, $N - 1$ slots, would simply be a comb with N slots, that is contracted with the supermap. We have already encountered the mapping of a three-slot comb to a two-slot comb by contraction with another three-slot comb in Fig. 2.11. The general proof for this graphically plausible statement can be found in [17]. Here, we chose to take the constructive route, as it is more intuitive and closer in spirit to the understanding that underlies the description of open quantum system dynamics.

²⁰ This statement holds true under the additional assumption of ‘compatibility with remote connections’ of the maps that map combs onto combs [17], *i.e.*, if these maps are causally ordered themselves. For a more general discussion of such maps, which also goes beyond quantum mechanics, see [63, 64].

As both approaches yield the same definition of combs, we have not compromised on generality by opting for the constructive approach.

Before we conclude this chapter, it is important to – just like in the case of channels – link the trace properties of deterministic combs to the underlying causal ordering.

2.10 DETERMINISTIC COMBS AND CAUSAL ORDERING

As we discussed in Sec. 2.7, the trace preservation condition of quantum channels can be considered a causality constraint; measurements of the outputs of the map can *overall* not influence the probabilities of the inputs, that are freely chosen before the measurement occurs. We emphasize the word ‘overall’ in the previous sentence, as for individual outcomes, it might seem like the measurement outcomes could indeed influence the input probabilities.

For example, consider the simple scenario, where Alice prepares states $\{|0\rangle, |1\rangle\}$ in the Z-basis with probabilities $\{p_0 = \frac{1}{2}, p_1 = \frac{1}{2}\}$ and sends them to Bob via an identity channel. Bob performs measurements in the Z-basis and records the outcomes. Now, if he only considers the runs of the experiment, where he obtained an outcome 0, and asks Alice, which input she prepared for said runs, it looks like Alice prepared state $|0\rangle$ with probability $p_0 = 1$. On the other hand, if Bob conditioned data collection on the outcome 1, it would seem as if Alice prepared state $|1\rangle$ with probability $p_1 = 1$. Consequently, Bob can make it appear like he had an influence on Alice’s input probabilities by conditioning their joint statistics on a particular outcome he obtains, but – quite obviously – overall, *i.e.*, without conditioning, there is no influence of Bob’s measurement on Alice’s input probabilities.²¹ On the other hand, Alice can change Bob’s measurement statistics by changing her preparation procedure. In this sense, causality is a statistical property, and it does not make much sense to ask for the causal order of a process with any individual run of an experiment [49].

This understanding of causality implies that if a process/network is causally ordered $A \prec B \prec \dots \prec N$, then an event K cannot influence the statistics of any of the events $K' \prec K$. We can see straightforwardly that the hierarchy of trace conditions in Eq. (2.42) indeed corresponds to causality constraints. To see this, imagine a four-slot comb defined on times $\{t_1, \dots, t_5\}$ (see Fig. 2.22 for a graphical representation of the described experimental scenario). The measurement statistics of an experimenter who performs measurements at times $\{t_1, t_2\}$ should not depend on anything that happens at later times $\{t_3, t_4, t_5\}$. Specifically, this means that the *choice of instruments* $\mathcal{J}_3, \mathcal{J}_4$, and \mathcal{J}_5 can have no influence on the statistics of outcomes at $\{t_1, t_2\}$. Mathematically, this implies

$$\mathbb{P}(\alpha, \beta | \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4, \mathcal{J}_5) = \mathbb{P}(\alpha, \beta | \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}'_3, \mathcal{J}'_4, \mathcal{J}'_5) = \mathbb{P}(\alpha, \beta | \mathcal{J}_1, \mathcal{J}_2), \quad (2.44)$$

where $\mathbb{P}(\alpha, \beta | \mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4, \mathcal{J}_5)$ denotes the probability to obtain outcome α (β) at t_1 (t_2) given that the instruments $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4$, and \mathcal{J}_5 are used, and Eq. (2.44) has to hold for all

²¹ We will discuss how conditioning can be used to simulate causally disordered processes in Ch. 6.

possible instruments. Expressed in terms of link products, this relation can, equivalently, be written as

$$\begin{aligned} \mathcal{C}^{(5)} \star M_1^{(\alpha)} \star M_2^{(\beta)} \star M_3 \star M_4 \star \mathbb{1}_{5^0} &= \mathcal{C}^{(5)} \star M_1^{(\alpha)} \star M_2^{(\beta)} \star N_3 \star N_4 \star \mathbb{1}_{5^0} \\ \forall \text{ CP maps } M_1^{(\alpha)}, M_2^{(\beta)} \quad \text{and} \quad \forall \text{ CPTP maps } M_3, M_4, N_3, N_4, \end{aligned} \quad (2.45)$$

where $M_1^{(\alpha)}$ and $M_2^{(\beta)}$ are CP maps corresponding to measurement outcomes α and β , M_3 and M_4 (N_3 and N_4) are CPTP maps corresponding to the instruments \mathcal{J}_3 and \mathcal{J}_4 (\mathcal{J}'_3 and \mathcal{J}'_4), and $\mathbb{1}_{5^0}$ is the only trace-preserving effect at t_5 . Now, using the hierarchy of trace conditions (2.42), we can show that Eq. (2.45) is indeed satisfied. For *any* CPTP maps $M_3 \in \mathcal{B}(\mathcal{H}_3^i \otimes \mathcal{H}_3^o)$ and $M_4 \in \mathcal{B}(\mathcal{H}_4^i \otimes \mathcal{H}_4^o)$ we have

$$\begin{aligned} \mathcal{C}^{(5)} \star M_3 \star M_4 \star \mathbb{1}_{5^0} &= \mathcal{C}^{(4)} \star M_3 \star M_4 \star \mathbb{1}_{4^i} = \mathcal{C}^{(4)} \star M_3 \star \mathbb{1}_{4^o} \\ &= \mathcal{C}^{(3)} \star M_3 \star \mathbb{1}_{3^i} = \mathcal{C}^{(3)} \star \mathbb{1}_{3^o} = \mathcal{C}^{(2)} \star \mathbb{1}_{2^o}, \end{aligned} \quad (2.46)$$

where we used the fact that $M_k \star \mathbb{1}_{k^i} = \mathbb{1}_{k^o}$ for CPTP maps M_k . This implies that, independent of the CPTP maps that are plugged into the final slots of the comb $\mathcal{C}^{(5)}$, the resulting comb that acts on $M_1^{(\alpha)}$ and $M_2^{(\beta)}$ is equal to $\mathcal{C}^{(2)} \otimes \mathbb{1}_{2^o}$; put differently, the choice of instruments at later times (in this case, at t_3 , t_4 and t_5) cannot be signaled to earlier times.

More generally, it can be proven in an analogous way that any deterministic *comb* that is performed at later times cannot influence measurement statistics at earlier times (see Fig. 2.23); for example, we could contract a deterministic comb $\mathcal{C}^{(N)} \in \mathcal{B}(\mathcal{H}_N^o \otimes \cdots \otimes \mathcal{H}_0^i)$ with a deterministic comb $M^{(N-n)}$ that is defined on $\mathcal{B}(\mathcal{H}_N^o \otimes \mathcal{H}_{N-1}^i \otimes \cdots \otimes \mathcal{H}_n + 1^o)$, *i.e.*, on the last $N - n$ slots of $\mathcal{C}^{(N)}$, and, in order for the contraction to make sense, the output lines of $\mathcal{C}^{(N)}$ are input lines of $M^{(N-n)}$ (see, for example, Fig. 2.23). Consequently, we have $M^{(N-n)} = \mathbb{1}_{N^o} \otimes M^{(N-n-1)}$, and $\mathcal{C}^{(N)} \star M^{(N-n)}$ yields

$$\begin{aligned} \mathcal{C}^{(N)} \star M^{(N-n)} &= \mathcal{C}^{(N)} \star M^{(N-n-1)} \star \mathbb{1}_{N^o} = \mathcal{C}^{(N-1)} \star M^{(N-n-1)} \star \mathbb{1}_{N-1^i} \\ &= \mathcal{C}^{(N-1)} \star M^{(N-n-2)} \star \mathbb{1}_{N-2^o} = \cdots = \mathcal{C}^{(n+1)} \star M^{(1)} \star \mathbb{1}_{n+1^i} \\ &= \mathcal{C}^{(n+1)} \star \mathbb{1}_{n+1^o} = \mathcal{C}^{(n)} \star \mathbb{1}_{n^i}, \end{aligned} \quad (2.47)$$

where we alternately used the fact that $\mathcal{C}^{(N)}$ and $M^{(N-n)}$ are deterministic combs, but no other properties of $M^{(N-n)}$ were used. As $\mathcal{C}^{(n)}$ is independent of $M^{(N-n)}$, Eq. (2.47) tells us, that, using the deterministic comb $\mathcal{C}^{(N)}$, no information from the future can be sent to the past. Unsurprisingly, if $\mathcal{C}^{(N)}$ is not deterministic – for example, if one of its building blocks is a non-trace-preserving effect or transformation – then it generally does not have a built-in causal order. This is reminiscent of the difference of CP maps and CPTP maps. While the latter preserve trace, and are as such causally ordered, the former do in general not preserve trace, and do not display causal order. However, if CP maps correspond to outcomes of a general measurement, they have to add up to a CPTP map, *i.e.*, causality has to hold on average. Analogously, we can consider non-deterministic combs as the maps corresponding to general, possible temporally correlated measurements [5, 47]. Such a set of positive matrices that add up to a deterministic comb is called a *tester* [17, 46, 47]. The concept of testers extends the idea of instruments to the temporally non-local case, and in this sense, deterministic combs and elements of a tester are the natural generalization of CPTP and CP maps, respectively.



Figure 2.22: Causality constraints for combs. The choice of instruments at the slots corresponding to t_3, t_4 , and at the final time t_5 should not influence the statistics of earlier measurements, represented by the CP maps $M_1^{(\alpha)}$ and $M_2^{(\beta)}$. We choose different colors for the operations that are plugged into $C^{(5)}$ to emphasize that the times $\{t_3, t_4, t_5\}$ are considered future events and should not have any influence on the measurement statistics at $\{t_1, t_2\}$.

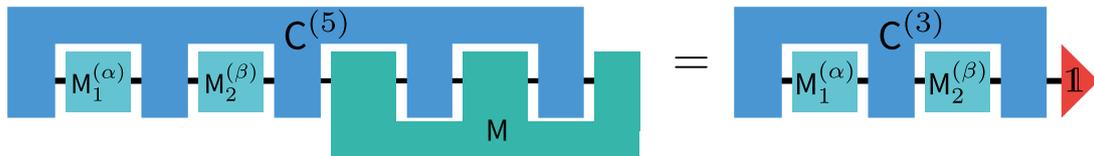


Figure 2.23: General causality constraints for combs. The choice of tester at the slots corresponding to t_3, t_4 , and at the final time t_5 should not influence the statistics of earlier measurements, represented by the CP maps $M_1^{(\alpha)}$ and $M_2^{(\beta)}$.

2.11 HIGHER-ORDER QUANTUM MAPS – SUMMARY

In this chapter, we have introduced the concept of higher-order quantum maps, and their graphical, as well as algebraic, characterization. Even though – apart from some basic examples – we have not yet provided natural physical instances, it should have become obvious that they provide a universal and versatile platform to describe any physical scenario where information is created, transformed and read-out. Arguably, all dynamical situations in quantum mechanics are of this kind, and having this powerful framework at hand will allow us to easily express and manipulate complex physical problems.

One field of application that suggests itself is that of open quantum systems dynamics. Here, the degrees of freedom that can be controlled by the experimenter – dubbed the ‘system’ – are coupled in an uncontrollable, and often times unknown, way to additional degrees of freedom – dubbed the ‘environment’. Due to this coupling, memory can travel forward in time, implying that the dynamics of the system at a point in time can depend on its states at earlier points in time. The description and characterization of such dynamics with memory has long posed a challenge for the open quantum systems community [65, 66] and is often still considered an open problem. Evidently, higher-order quantum maps are perfectly tailored to describe such physical scenarios, and we will make constant use of them for the description of general quantum stochastic processes in the subsequent chapters.

On a more fundamental level, we shall see that a switch of the employed descriptive framework also allows one to resolve open foundational problems at relative ease. Additionally, as we pointed out, quantum combs provide the natural framework to discuss causal order in quantum mechanics and enable a straightforward generalization to situations without a predefined causal order.

Before we come to these, more speculative, discussions, we will first use the theory of higher-order quantum maps for the succinct and consistent description of quantum stochastic processes (and beyond), providing a unified framework for several – seemingly distinct – fields

of physics. While this discussion will still be somewhat abstract, we shall furnish higher-order quantum maps with a clear-cut underlying physical meaning, and experimental implementation in Chs. 4 and 5.

In the previous chapter we introduced higher-order quantum maps as a versatile framework to describe many different physical situations. Here, we will make direct use of them and, in a sense, rediscover them again – albeit with a different motivation, and an explicit underlying physical idea in mind, namely: the description of *quantum stochastic processes*. In the classical domain, stochastic processes are well-understood, both conceptually, with clear and unambiguous definitions for the key concepts. On the other hand, in quantum mechanics, the situation presents itself somewhat murkier.

Fundamentally, an appropriate descriptor of a stochastic process should allow one to correctly predict the probabilities for arbitrary sequences of measurement outcomes, and would thus constitute the maximum amount of information that an experimenter can extract about the dynamics of a system of interest. For example, the full description of a particle undergoing Brownian motion when immersed in some uncontrollable environment is given by the probabilities $\mathbb{P}(x_N, t_N; \dots; x_2, t_2; x_1, t_1)$ to find the particle in the regions $\{x_\alpha\}_{\alpha=1}^N$ when measuring its position at times $\{t_\alpha\}_{\alpha=1}^N$ (see Fig. 3.1); once all of these probabilities are known, there is no more information to be inferred about the process. This description of stochastic processes in terms of joint probability distributions does not translate well to quantum mechanics, where the mere act of measuring alters the state of the interrogated system, and thus influences future dynamics [65, 67, 68].

This experimental influence challenges the idea of a separation between the experimenter and the dynamical process that they intend to probe. The aim of this chapter is to overcome this problem and to provide a descriptor for general *quantum* processes that is independent of potential experimental interventions, and which contains all inferable information about the underlying process. The offered resolution, in turn, will then kill many birds with one stone: it alleviates the conceptual problems in the discussion of quantum processes, provides a unified logical framework for seemingly disparate fields of physics, and enables the unambiguous definition of memory effects in quantum stochastic processes. The former two points will be discussed in the present chapter, while we shall return to the latter one in Ch. 4.

The difficulties that arise when modeling quantum processes can be traced back to the invasiveness of measurements in quantum theory, which appear to introduce a disconnect between finite measurement statistics and the underlying process; experimentally – due to the restriction of temporal resolution and the amount of data that can be detected and stored – we always deduce probability distributions with a finite number of arguments. Nonetheless, one always implicitly assumes that there is an *underlying* process that leads to the experimentally

accessible finite distributions. Put differently, one assumes that there exists an *infinite* joint probability distribution that has *all* the finite ones as marginals. Active influence of the experimental procedure then appears to suggest that there is a fundamental divide between what the experimenter sees, and what the process ‘actually’ is. This kind of divide does not exist in classical physics, where these two points of view – the finite and the infinite one – are reconciled by the Kolmogorov extension theorem (**KET**), which lays bare the minimal requirements for the existence of an underlying process, given a family of measurement statistics for finite sets of times [69–72]. It bridges the gap between experimental reality and mathematically rigorous theoretical underpinnings and, as such, enables the definition of stochastic processes as the limit of finite – and hence observable – objects.

The validity of the KET hinges crucially on the fact that the statistics of observations at a time t do not depend on the kind of measurements that were made at any time $t' < t$. In other words, just like the Leggett-Garg inequalities for temporal correlations [73–75], the KET is based on the assumptions of realism and non-invasive measurements.¹ For example, in a classical stochastic process, measuring the position of a particle undergoing Brownian motion merely reveals information, but does not actively change the state of the particle. On the other hand, the assumptions of non-invasiveness or realism are not fulfilled in many experimental (quantum) scenarios and potential invasiveness leads to a breakdown of the KET, at the cost of a clear connection between an underlying process and its finite time manifestations.

Invasive experimental interventions, the culprit when it comes to posing a problem with respect to describing general stochastic processes, appear naturally in quantum mechanics; in fact, a complete description of quantum processes without interventions is not possible [8]. More generally, traditional descriptions of stochastic processes break down, whenever an experimenter – classical or quantum – chooses to *actively* interfere with a process, which is necessary to uncover its causal structure. For example, instead of just observing the progress of a disease, a pharmacologist tries to find out how the course of a disease changes with the administration of certain drugs, while agent-based modeling investigates how systems behave when they can not only be monitored, but actively influenced [77]. Experimental situations where interventions are actively used to uncover causal relations fall within the fields of *causal modeling* and, in the quantum case, *quantum causal modeling* [42, 49–51].

The fundamental lack of an extension theorem in theories with interventions would be problematic for several reasons: firstly, it would suggest an inconsistency between descriptions of a process for different sets of times; for example, the description of a process for three times t_1, t_2 , and t_3 would *not* already include the description of the process for the two times t_1 and t_3 only. We would need seven different independent descriptors for each of the seven subsets of times to describe all possible events! This lack of consistency would render the study of (quantum) causal models in multi-step experiments impossible; if local interventions lead to a completely different process, it is not meaningful to try to deduce causal relations by means of active manipulations of the system at hand.

¹ In [73, 74], the assumption of *macroscopic* realism is made by Leggett and Garg for the derivation of the eponymous inequality. However, realism (or realism *per se*) is a sufficient assumption (see for example [76] for a detailed discussion), and the one that is most frequently tested for in experiments [75].

Furthermore, the present situation (with no extension theorem) implies an incompatibility between existing frameworks to describe processes with interventions (both classical and quantum) and the classical theory of stochastic processes, even though the former should reduce to the latter in the correct limit. This then suggests that the mere act of interacting with a system over time introduces a fundamental divide between the continuity of physical laws and the finite statistics that can be accessed in reality, thus begging the question: what do we mean by a (quantum) ‘stochastic process’, and how can we reconcile causal modeling frameworks with the idea of an underlying process?

Beyond the foundational considerations, the lack of a full-fledged framework for the treatment of quantum stochastic processes has practical implication for our understanding of memory effects in quantum mechanics. Closed system dynamics, *i.e.*, dynamics where interactions of some accessible system of interest (*i.e.*, the system that the experimenter can interrogate) with its environment are negligible, are mathematically well-understood in quantum mechanics.² However, the situation becomes more complex and subtle, when the system of interest interacts with an environment that is beyond experimental control. Traditionally, in this case, one aims to find the dynamics of the (open) system by solving an equation of motion for the state of the former, which yields families of channels that map initial states of the system to states at later times. Evidently, such methods can only account for two-time correlation functions, which is only sufficient to describe processes with negligible memory effects [19, 71, 78, 79], *i.e.*, *Markovian* processes, but generally insufficient to fully describe an arbitrary (non-Markovian) quantum process [8, 67, 71]. Situations with non-negligible memory effects require explicit accounting for multi-time correlations [65, 80] which already points us in the direction of higher-order quantum maps as the ideal tool to circumvent the problems one encounters as soon as memory effects play a role for the dynamics of a quantum system.

In this chapter, based on the findings of [4], we shall answer all of the above questions (except for the definition of memory effects), by generalizing the KET to the framework of (quantum) causal modeling, thus closing the apparent divide between the finite and the continuous point of view. This generalization, in turn, will put the existing theories of stochastic processes (classical, quantum, and beyond), and the theory of causal modeling on the same unified footing. The breakdown of the KET is thus revealed as a breakdown of (current) formalism only, not a fundamental property of quantum processes. Additionally, even though we will predominantly phrase our results in terms of the comb formalism that we encountered in the previous chapter, they also apply to causally disordered processes, as well as potential post-quantum theories that satisfy some minimal assumptions of OPTs. Finally, the generalization of the KET to quantum mechanics will provide a direct way to construct generative models of quantum stochastic processes, which we shall introduce and discuss in detail.

This chapter lays the foundations for the discussion of quantum stochastic processes and establishes quantum combs as their most general descriptors. The investigation of the *properties* of general quantum stochastic processes, like, *e.g.*, the presence of memory effects, is relegated to subsequent chapters, where we will investigate their physical origin, as well as their experimental reconstruction and the characterization and detection of non-Markovianity.

² The fact that the mathematical framework to describe closed system dynamics is well-understood does, of course, not mean that the corresponding dynamics are easy to compute in general.

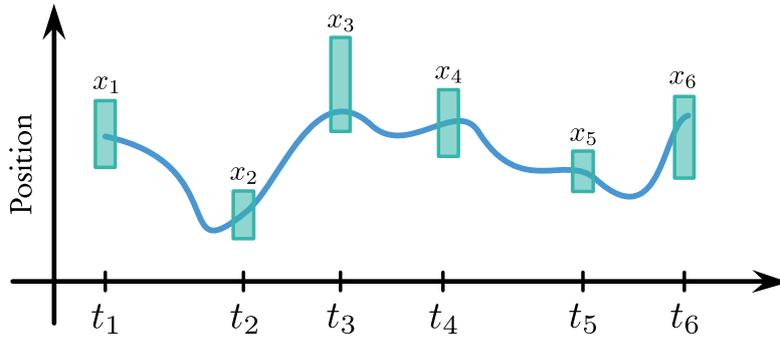


Figure 3.1: Example trajectory of a particle undergoing Brownian motion. For six times $\{t_6, \dots, t_1\}$, the underlying process is fully described by the joint probabilities $\mathbb{P}(x_6, t_6; \dots; x_1, t_1)$ to find the particle in the regions $\{x_6, \dots, x_1\}$ at times $\{t_6, \dots, t_1\}$.

3.1 CLASSICAL STOCHASTIC PROCESSES – BASIC CONCEPTS

Stochastic processes are ubiquitous in nature. Their theory is used, among other applications, to model the stock market, predict the weather, describe transport processes in cells and understand the random motion of particles suspended in a fluid [81, 82]. Intuitively, when we speak of stochastic processes, we often mean joint probability distributions of possible measurement outcomes at a finite set of times: the probability for a stock to have prices $\$1$, $\$2$ and $\$3$ on three subsequent days, or the probability to find a particle undergoing Brownian motion in regions x_2 and x_1 when measuring its position at times t_2 and t_1 . This finite intuition notwithstanding, we always assume that there is an *underlying* process that leads to the observed finite joint distributions. Discussing the properties of this underlying process and connecting it to the observed finite distributions is the goal of this and the subsequent sections.

In order to treat them more rigorously, and to facilitate the transition to the quantum regime, it is necessary to define some basic concepts pertaining to the theory of classical stochastic processes (see, for example, Ref. [71] for a more comprehensive discussion). Firstly, in order to be able to meaningfully attribute probabilities to events, it is useful to define the concept of **σ -algebras** [72]:

DEFINITION 3.1 (σ -ALGEBRA): Let Ω be a set. A σ -algebra on Ω is a collection F of subsets of Ω , such that

- $\Omega \in F$ and $\emptyset \in F$.
- If $f \in F$, then $\Omega \setminus f \in F$.
- F is closed under (countable) unions and intersections, i.e., if $f_1, f_2, \dots \in F$, then $\bigcup_{j=1}^{\infty} f_j \in F$ and $\bigcap_{j=1}^{\infty} f_j \in F$.

We will call such a pair (F, Ω) a *measurable space*. Intuitively, the *sample space* Ω can be considered the space of all possible outcomes, while the σ -algebra F is the set of all possible events that can be detected. For example, when measuring the position of a particle in a plane, the space of all possible outcomes is $\Omega = \mathbb{R}^2$, while the σ -algebra F depends on the instrument

that is employed to detect the particle. If the instrument can only resolve if the particle is in the positive half-plane $x \geq 0$ or the negative half-plane $x < 0$, we have $F = \{\mathbb{R}_{x \geq 0}^2, \mathbb{R}_{x < 0}^2, \mathbb{R}^2, \emptyset\}$. Throughout this thesis, we will assume that the number of detectable outcomes is finite, *i.e.*, $|F| < \infty$. A σ -algebra can be understood as allocating a set f_j to every possible outcome j .

Having defined σ -algebras, we can introduce the second important ingredient for the theory of stochastic processes, **probability measures**:

DEFINITION 3.2 (PROBABILITY MEASURE): *Let (F, Ω) be a measurable space. A probability measure $\mathbb{P} : F \rightarrow \mathbb{R}$ is a real-valued function that satisfies*

- $\mathbb{P}(\Omega) = 1$.
- $\mathbb{P}(f) \geq 0$ for all $f \in F$
- \mathbb{P} is additive for (countable) unions of disjoint events, *i.e.*, $\mathbb{P}\left(\bigcup_{j=1}^{\infty} f_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(f_j)$ for $f_j \in F$ and $f_j \cap f_{j'} = \emptyset$ when $j \neq j'$.

As the name suggests, \mathbb{P} maps each event f_j to its corresponding probability.³ The definition of a σ -algebra is in clear analogy to that of a POVM [61, 83, 84], and, more generally, to the definition of instruments we gave in the previous chapter; roughly speaking, each of the CP maps of an instrument \mathcal{J} corresponds to an element of the σ -algebra, *i.e.*, a detectable/measurable event, that is defined by what \mathcal{J} can resolve.

With these concepts at hand, we are now in the position to define the concept of a **probability space** [72]:

DEFINITION 3.3 (PROBABILITY SPACE): *A probability space is a triple (\mathbb{P}, Ω, F) , of a probability measure \mathbb{P} , and a measurable space (Ω, F) , with $\mathbb{P}(\Omega) = 1$.*

Quite naturally, $\mathbb{P} : F \rightarrow [0, 1]$ yields the probability to obtain outcome f in F . For the example above, \mathbb{P} would yield the probability to find the particle in one of the half-planes. In quantum mechanics, $\mathbb{P}(f)$ is computed via the Born rule $\text{tr}(E_f^T \rho)$, where E_f^T is the POVM element corresponding to the outcome f .

The extension to sequences of measurement outcomes at multiple times – *i.e.*, the definition of stochastic processes – is now straightforward. Before we do so, a brief remark is necessary to avoid potential confusion. In the literature, stochastic processes are generally defined in terms of *random variables* [70, 71]. For completeness, we provide this definition in App. A.1. However, all of our results will be phrased with respect to joint probability distributions, without making explicit use of the concept of random variables. Therefore, it would lead us unnecessarily astray to express our definition of a classical stochastic process in terms of them. For all intents and purposes, there is no difference between our approach and the one generally found in the literature.

³ Strictly speaking, we should have used a different symbol for a probability measure defined in this way, as we already used \mathbb{P} in the previous chapter for probabilities obtained via the Born rule. However, as we shall see, both concepts coincide (in the right sense) in quantum mechanics.

DEFINITION 3.4 (CLASSICAL STOCHASTIC PROCESSES ON FINITE SETS OF TIMES): A stochastic process on times $t_\alpha \in \Lambda_k$ with $|\Lambda_k| = k < \infty$ is a triple $(\mathbb{P}_{\Lambda_k}, \Omega_{\Lambda_k}, F_{\Lambda_k})$ of a sample space

$$\Omega_{\Lambda_k} = \times_{\alpha \in \Lambda_k} \Omega_\alpha, \quad (3.1)$$

a σ -algebra F_{Λ_k} on Ω_{Λ_k} , and a probability measure \mathbb{P}_{Λ_k} on F_{Λ_k} with $\mathbb{P}_{\Lambda_k}(\Omega_{\Lambda_k}) = 1$.

The symbol \times denotes the Cartesian product for sets. For compactness, whenever there is no risk of confusion, we will use $\alpha \in \Lambda_k$ to denote labels of times in Λ_k .⁴ Each Ω_α corresponds to a sample space at t_α , and the probability measure $\mathbb{P}_{\Lambda_k} : F_{\Lambda_k} \rightarrow [0, 1]$ maps any sequence of outcomes at times $\{t_\alpha\}_{\alpha \in \Lambda_k}$ to its corresponding probability of being measured. Importantly, the employed instruments do not have to be the same at each time steps. For instance, in the example above, the spatial resolution of the measurement device could be distinct for different times.

While generally, one would identify Λ_k with a chronological set of times – and we do so frequently throughout this chapter to build intuition – we will often leave its physical origin undefined, in order to also allow for the definition of stochastic processes without a global temporal order. In what follows, we shall employ the convention, that, unless stated otherwise, the set Λ_k does not necessarily have to display a chronological order.

In less technical terms, a stochastic process, as we define, it is given by joint probability distributions $\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1}) := \mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k})$ over sequences of measurement outcomes that take values $\{x_{j_\alpha}\}$ at times/labels $t_\alpha \in \Lambda_k$, where Λ_k is a set with cardinality $|\Lambda_k| = k$ [71].⁵ Once \mathbb{P}_{Λ_k} is known, there is no more information that can be learnt about the process (for the given measurement devices).

The slightly technical definition of stochastic processes in terms of σ -algebras has the advantage of generalizing straightforwardly to the case of quantum stochastic processes; as we will see, the role of elements $f \in F_{\Lambda_k}$ will be played by CP maps $M \in \otimes_{\alpha \in \Lambda_k} \mathcal{B}(\mathcal{H}_\alpha^i \otimes \mathcal{H}_\alpha^o)$, while the role of \mathbb{P}_{Λ_k} will be assumed by a positive linear functional, \mathcal{C}_{Λ_k} , that maps sequences of CP maps – corresponding to sequences of measurement outcomes at times t_α in Λ_k – to their associated probabilities. In quantum mechanics, \mathcal{C}_{Λ_k} is a comb, and the corresponding probability is given by $\mathcal{C}_{\Lambda_k} \star_{\alpha \in \Lambda_k} M_{x_{j_\alpha}}$, where \star is shorthand for a link product over all labels $\alpha \in \Lambda_k$ of times in Λ_k (see below).

The above definition of stochastic processes naturally generalizes to sets of times Λ , where $|\Lambda|$ can be infinite:

DEFINITION 3.5 (CLASSICAL STOCHASTIC PROCESSES): A stochastic process on times $t \in \Lambda$ is a triple $(\mathbb{P}_\Lambda, \Omega_\Lambda, F_\Lambda)$ of a sample space

$$\Omega_\Lambda = \times_{\alpha \in \Lambda} \Omega_\alpha, \quad (3.2)$$

a σ -algebra F_Λ on Ω_Λ , and a probability measure \mathbb{P}_Λ on F_Λ with $\mathbb{P}_\Lambda(\Omega_\Lambda) = 1$.

⁴ Here, we break with the convention of the previous chapter, and label times by α (instead of j), to distinguish them clearly from k , the cardinality of the set of times the stochastic process is defined on.

⁵ Throughout this chapter, we will employ the convention that subscripts signify the time as well as the respective outcome. For example, x_{j_α} signifies a measurement outcome at time t_α , labeled by j_α .

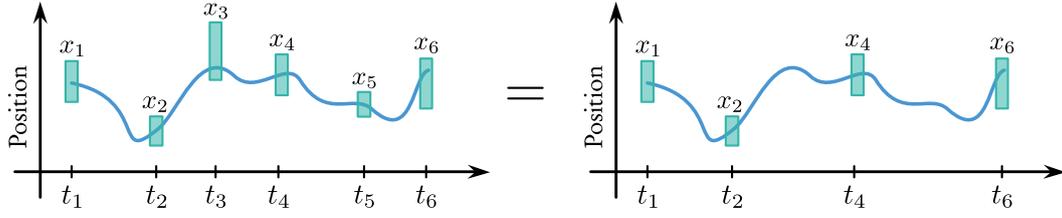


Figure 3.2: Consistency condition for classical stochastic processes. The correct joint probability $\mathbb{P}_{\Lambda_{k'}}$ can be obtained from any joint probability distribution \mathbb{P}_{Λ_k} with $\Lambda_k \supseteq \Lambda_{k'}$ by summing over the excessive times. Here, $\Lambda_k = \{t_6, \dots, t_1\}$ and $\Lambda_{k'} = \{t_6, t_4, t_2, t_1\}$, and $\mathbb{P}_{\Lambda_{k'}} = \sum_{j_3, j_5} \mathbb{P}_{\Lambda_k}$.

Notably, \mathbb{P}_Λ is not an experimentally reconstructible quantity unless $|\Lambda|$ is finite. Nonetheless, we implicitly assume its existence when probing stochastic processes on finite sets of times. Connecting finite joint probability distributions to the concept of an underlying process is the main achievement of the Kolmogorov extension theorem, as we will discuss in detail in the subsequent sections.

At this point, it is important to shed some light onto the assumptions that go into the discussion of stochastic processes in classical physics. Implicitly, in classical physics, we assume that the interrogation of the system overall leaves its state unchanged; put differently, we assume that there is no difference between averaging over the measurement outcomes at t_α , and *not* performing a measurement at t_α at all. Consequently, the stochastic process on any subset of times $\Lambda_{k'} \subseteq \Lambda_k$ can be obtained from \mathbb{P}_{Λ_k} ; for instance, if $\Lambda_k = \{t_k, \dots, t_1\}$, the distribution over all but the α^{th} time is found by marginalizing the larger distribution: $\mathbb{P}_{\Lambda_k \setminus \{t_\alpha\}}(x_{j_k}, \dots, x_{j_{\alpha+1}}, x_{j_{\alpha-1}}, \dots, x_1) = \sum_{j_\alpha} \mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k})$, where \sum_{j_α} denotes the sum over all outcomes labeled by j_α (see Fig. 3.2). As we shall see below, this condition is not satisfied by quantum processes, raising numerous complications in the definition and description of quantum stochastic processes.

3.2 σ -ALGEBRAS, POVMS AND INSTRUMENTS

In the previous section, we have hinted at the fact that POVMs and instruments are the natural extension of σ -algebras to the quantum mechanical setting. Before discussing causal modeling and quantum stochastic processes in the subsequent sections, it is instructive to define POVMs and instruments in a slightly more rigorous manner than we have so far. In doing so, we make the connection between the respective concepts manifest, which, in turn, we can use as the guiding principle to a resolution of the difficulties that arise when describing processes with interventions.

Previously, we defined a POVM as a set of positive operators $\{E_j\}$ that satisfy $\sum_j E_j = \mathbb{1}$. While this definition is correct, it obfuscates its natural connection to σ -algebras and probability measures. We have the following, more rigorous definition [61, 83, 85]:

DEFINITION 3.6 (POSITIVE OPERATOR VALUED MEASURE): *Let (F, Ω) be a measurable space. A Positive Operator Valued Measure (POVM) $\mathcal{G} : F \rightarrow \mathcal{B}(\mathcal{H})$ is a mapping from the σ -algebra F to operators on \mathcal{H} , such that*

- $\mathcal{G}[\Omega] = \mathbb{1}$

- $\mathcal{G}[f_j] = E_j \geq 0$ for all $f_j \in F$
- $\mathcal{G}[\bigcup_{j=1}^{\infty} f_j] = \sum_{j=1}^{\infty} \mathcal{G}[f_j]$, for $f_j \in F$, and $f_j \cap f_{j'} = \emptyset$ for $j \neq j'$.

Several remarks about this definition are in order. In classical physics, an instrument defines what measurement outcomes can be resolved, and as such partitions the sample space into a σ -algebra. In quantum mechanics, this is not enough. A single degree of freedom can be interrogated in different, non-commuting ways, and consequently, it is not sufficient to only specify the possible measurement outcomes, but the respective instrument that was used to interrogate the system of interest is of importance.

For instance, when measuring the spin of a spin- $\frac{1}{2}$ particle in the Z- and X- basis, the possible outcomes can be labeled by $\{0, 1\}$ in either case, but the corresponding measurements and POVM elements differ. We have already seen that for a given quantum state ρ , the probability for an outcome j , when using the instrument \mathcal{J} , is computed via Born's rule $\mathbb{P}(j|\mathcal{J}) = E_j \star \rho$. In this sense, Born's rule can be considered as a mapping from a σ -algebra to probabilities *via a POVM* that allows one to account for the different ways in which a quantum mechanical degrees of freedom can be measured. As such, Def. 3.6 in conjunction with Born's rule is the natural extension of the definition of a probability space in classical physics.

Analogously, if the transformation of the state upon receiving outcome j is of importance – as is evidently the case in the study of stochastic processes – the definition of a POVM given above is straightforwardly extended to the rigorous definition of an instrument [62]:

DEFINITION 3.7 (INSTRUMENTS): *Let (F, Ω) be a measurable space. An instrument $\mathcal{J} : F \rightarrow \mathcal{B}(\mathcal{H}^0 \otimes \mathcal{H}^1)$ is a mapping from the σ -algebra F to operators on $\mathcal{H}^1 \otimes \mathcal{H}^0$, such that*

- $\mathcal{J}[\Omega]$ is a CPTP map M
- $\mathcal{J}[f_j] = M_j \geq 0$ for all $f_j \in F$
- $\mathcal{J}[\bigcup_{j=1}^{\infty} f_j] = \sum_{j=1}^{\infty} \mathcal{J}[f_j]$, for $f_j \in F$ and $f_j \cap f_{j'} = \emptyset$ for $j \neq j'$.

This definition of an instrument coincides with the previous understanding of an instrument as a collection of CP maps that add up to a CPTP map, but has the advantage, that it now makes the connection to classical stochastic processes straightforward. Importantly, the σ -algebra in the above definition could correspond to a multi-time experiment, and as such, it also applies to testers. In slight abuse of notation, whenever possible, we will continue to identify \mathcal{J} with a collection of CP maps, instead of a mapping from a σ -algebra to $\mathcal{B}(\mathcal{H}^1 \otimes \mathcal{H}^0)$.

As we have seen, a classical stochastic process on a set of times Λ_k is defined as a map $\mathbb{P} : F_{\Lambda_k} \rightarrow [0, 1]$, where F_{Λ_k} is a σ -algebra on $\times_{\alpha \in \Lambda_k} \Omega_{\alpha}$. The definition of a quantum stochastic process then follows in an analogous way; we will argue below that, with the definition of an instrument at hand, a quantum stochastic process must be a mapping of the form [62, 86, 87]

$$\mathcal{C}_{\Lambda_k} : \bigotimes_{\alpha \in \Lambda_k} \mathcal{B}(\mathcal{H}_{\alpha}^1 \otimes \mathcal{H}_{\alpha}^0) \rightarrow \mathbb{R}, \quad (3.3)$$

such that \mathcal{C}_{Λ_k} yields a probability whenever it acts on a tester. In the subsequent sections, we rigorously discuss this concept of quantum stochastic processes on finite sets of times Λ_k , and

show that it answers the question of how to properly generalize the idea of an underlying process in quantum mechanics (and beyond).

Before we do so, we have to make sure, though, that we accurately understand why this question has to be asked in the first place, and what properties of general stochastic processes necessitate new answers. To this end, we will first embed classical stochastic processes into the broader framework of classical causal modeling, *i.e.*, classical processes with interventions, and show why fundamental properties of classical processes fail to hold as soon as we allow for active experimental influence. While – at first sight – seemingly constituting a slight conceptual detour, this discussion will lead us straight to the heart of the matter, quantum stochastic processes.

3.3 CLASSICAL CAUSAL MODELING

Classical stochastic processes are traditionally described by joint probability distributions for measurement outcomes at different times. However, observing the statistics for measurement outcomes reveals *correlations* between events, but no information about *causal* relations. For instance, correlations of two events A and B could stem from A influencing B , B influencing A , or both A and B being influenced by an earlier event C [49–51]. Both of these situations (direct cause and common cause) are depicted graphically in the DAG of Fig. 3.3, where events at A_3 and B_4 , as well as at A_4 and B_3 are correlated, respectively. However, in the former case, the correlations stem from A_3 causally influencing B_4 , while in the latter, they are a result of the common cause A_2 .

More tangibly, reiterating an example from Ref. [51], events A and B could, for example, be the occurrence of sunburns and the sales of ice cream, respectively. While these two variables are highly correlated, this correlation alone would not fix a causal relation between them. Inferring the causal structure of a process is the aim of causal modeling. Here, active interventions are used to uncover how different events can influence each other. In the example above, one could suspend the sale of ice cream to see how it affects the occurrence of sunburns, and would find out that ice cream sales have no direct effect on sunburns (and vice versa, as the correlations of ice cream sales and sunburns stem from a common cause, the weather, and not from any direct causal relation).

To uncover these causal relations, we have to go beyond passive observation, and allow for active interventions. Mathematically, this means that causal modeling for k events A_k, \dots, A_1 necessitates the collection of all joint probability distributions $\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1} | y_{\ell_k}, \dots, y_{\ell_1}) := \mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k} | \mathcal{J}_{\Lambda_k})$ to measure the outcomes x_{j_k}, \dots, x_{j_1} given that the *interventions* $y_{\ell_k}, \dots, y_{\ell_1}$ were performed. Here, Λ_k is a general set of labels for events and *a priori*, there is no particular order imposed on its elements. For example, Λ_k could contain labels for different laboratories where experiments are performed, and in each laboratory $\alpha \in \Lambda_k$, the experimenter can perform measurements (with outcomes x_{j_α}) and interventions (labeled by y_{ℓ_α}). \mathcal{J}_{Λ_k} are then the instruments that were used in each of these laboratories, and, they can be seen as rules for how to intervene upon seeing a particular outcome.

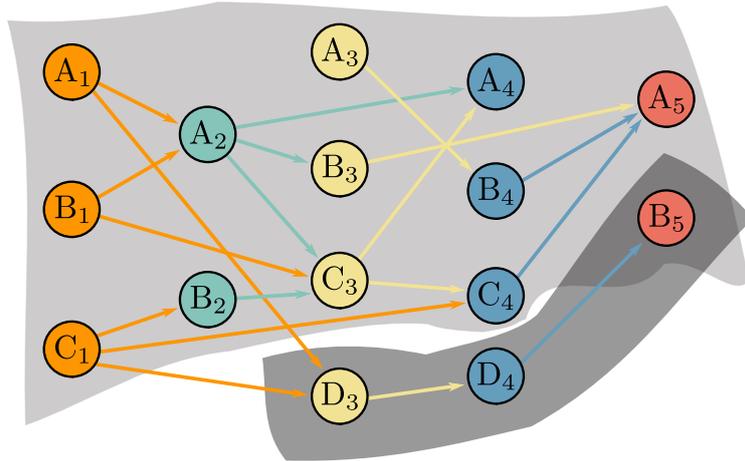


Figure 3.3: (Quantum) Causal network. Performing different interventions allows for the causal relations between events to be probed. For example, in the figure the event B_1 directly influences the events C_3 and A_2 , while A_3 influences only B_4 . Depending on the degrees of freedom that can be accessed by the experimenter, these causal relations can or cannot be detected. For instance, the influence of A_1 on D_3 could not be discovered if only the degrees of freedom in the light gray area were experimentally accessible. The statistics of events do in general not satisfy the requirements of the KET. For example, the events D_3, D_4 , and B_5 could be successive (e.g., at times t_3, t_4 and t_5) spin measurements in Z-, X- and Z-direction, respectively. Summing over the results of the spin measurement in X-direction at t_4 would not yield the correct probability distribution for two measurements in Z-direction at t_3 and t_5 only (see also Sec. 3.6). On the other hand, independent of the accessible degrees of freedom, the *generalized* extension theorem (GET) discussed in Sec. 3.8 holds for *any* process.

For instance, when investigating Brownian motion, an instrument could be a deterministic replacement rule: upon finding the particle at x_{j_α} , replace it by a particle at y_{ℓ_α} . It could also be probabilistic: upon finding the particle at x_{j_α} , replace it by a particle at y_{ℓ_α} with probability $p(y_{\ell_\alpha} | x_{j_\alpha})$. In anticipation of our generalization to quantum stochastic processes, we adopt Choi matrices to represent the corresponding two instruments mathematically. It is straightforward to see that the former is a collection $\{M_{y_{\ell_\alpha} x_{j_\alpha}}\}$ of CP maps, where

$$M_{y_{\ell_\alpha} x_{j_\alpha}} = |y_{\ell_\alpha}\rangle\langle y_{\ell_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|, \quad (3.4)$$

while the latter consists of CP maps $M_{x_{j_\alpha}}$ of the form

$$M_{x_{j_\alpha}} = \sum_{\ell_\alpha} p(y_{\ell_\alpha} | x_{j_\alpha}) |y_{\ell_\alpha}\rangle\langle y_{\ell_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}| := \rho_{x_{j_\alpha}} \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|, \quad (3.5)$$

where $|x_{j_\alpha}\rangle$ ($|y_{\ell_\alpha}\rangle$) is a definite state vector corresponding to the measurement outcome (repreparation) x_{j_α} (y_{ℓ_α}) at t_α , and $\rho_{x_{j_\alpha}}$ is a classical state, *i.e.*, it is diagonal in the classical basis $\{|y_{\ell_\alpha}\rangle\}$. As mentioned, the number of possible measurement outcomes and reпреparations can differ, but for compactness we will assume them to be equal for most of this chapter. The generalization to differing numbers of inputs and outputs is always possible without any technical difficulties.

Both sets of CP maps $\{M_{y_{\ell_\alpha} x_{j_\alpha}}\}$ and $\{M_{x_{j_\alpha}}\}$ defined above add up to a CPTP map, and thus constitute an instrument. Of course, we could discuss classical causal modeling without the explicit reference to the CP maps that correspond to the individual interventions, but it will prove helpful to do so, in order to transition to the quantum case at ease.

Unlike in quantum mechanics, in classical physics, it is also possible to measure the system of interest at each time t_α without actively intervening. We dub the corresponding instrument the idle instrument

$$\mathcal{J}_\alpha = \mathbf{id}_\alpha = \{|x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|\}_{j_\alpha=1}^d. \quad (3.6)$$

Overall, its action does not change the state of the interrogated system. In detail, setting $M_{\text{id}} = \sum_{j_\alpha} |x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|$, we see that $M_{\text{id}} \star \rho = \rho$ for every state ρ that is diagonal in the basis the measurements are performed in.

With this, we see that classical stochastic processes form a subset of classical causal modeling; they correspond to the probability distributions one obtains by employing the idle instrument at every time t_α :⁶

$$\mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k} | \mathbf{id}_{\Lambda_k}) = \mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k}), \quad (3.7)$$

where \mathbf{id}_{Λ_k} denotes the idle instrument at each of the events in Λ_k . While this statement is well-known, and somewhat obvious, it is crucial to understand its implications. In classical physics, and experimenter can choose to not disturb the state of the system of interest, when interrogating it. Fundamentally, this is why the theory of stochastic processes and causal modeling are distinct in classical physics and can be discussed independently. In quantum mechanics, this is not the case, and we shall see that the set of quantum stochastic processes coincides with that of quantum causal models. Basic properties of stochastic processes break down in the classical realm as soon as active interventions are considered, and the situation is even more problematic in quantum mechanics, where interventions cannot be avoided. To see this, it is now – finally – time to discuss the Kolmogorov extension theorem (KET), which lays the foundation for our understanding of classical stochastic processes.

3.4 THE KOLMOGOROV EXTENSION THEOREM (KET)

In practice, when probing a stochastic process, we can only do so for finite sets of times. The KET is concerned with the question of what properties the resulting family $\{\mathbb{P}_{\Lambda_k}\}_{\Lambda_k \subset \Lambda}$ of finite joint probability distributions must satisfy in order for an underlying process \mathbb{P}_Λ to exist. As such, it connects the experimentally accessible quantities $\{\mathbb{P}_{\Lambda_k}\}$ with the theoretical concept of an underlying process, and thus provides an operationally meaningful definition of what we mean by a classical stochastic process. It is important to emphasize that in what follows, we will be somewhat lax with respect to nomenclature. A stochastic process – both on finite as well as infinite sets of times – is a triple of a probability distribution, a sample space and a σ -algebra. However, for brevity we shall identify the stochastic processes we consider with the corresponding joint probability distributions alone, and omit the sample space and the σ -algebra, which will always exist.

⁶ Here, again, we use the reference to time as an aid to build intuition, not because it is important that the set the stochastic process is defined on is temporally ordered.

If the experimentally obtained family $\{\mathbb{P}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda}$ of finite joint probability distributions stems from an underlying stochastic process on the (finite, countably or uncountably infinite) set Λ , then all of them are marginals of one probability distribution \mathbb{P}_Λ . In detail, we have

$$\mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k}) = \sum_{\Lambda \setminus \Lambda_k} \mathbb{P}_\Lambda(\mathbf{x}_\Lambda) := \mathbb{P}_\Lambda^{\Lambda_k}(\mathbf{x}_{\Lambda_k}) \quad \forall \Lambda_k \subseteq \Lambda, \quad (3.8)$$

where \mathbf{x}_{Λ_k} is the subset of \mathbf{x}_Λ corresponding to the times Λ_k , $\sum_{\Lambda \setminus \Lambda_k}$ denotes a sum over outcomes at all times that are part of $\Lambda \setminus \Lambda_k$ (i.e., all the times that lie in Λ but not in Λ_k), and $\mathbb{P}_\Lambda^{\Lambda_k}$ denotes the restriction of \mathbb{P}_Λ to the times Λ_k . In case the set Λ is infinite, the marginalization procedure can correspond to an integral over the times in $\Lambda \setminus \Lambda_k$ (though, to avoid introducing too much notation, we will still use $\sum_{\Lambda \setminus \Lambda_k}$ to represent it). For example, if the process we are interested in is the Brownian motion of a particle, \mathbb{P}_Λ would be the probability density of all possible trajectories that the particle could take in the time interval Λ , and all finite distributions could, in principle, be obtained from \mathbb{P}_Λ .

If an underlying process exists, then Eq. (3.8) implies *consistency conditions* (also often called *compatibility* or *Kolmogorov conditions*) between probability distributions for any two finite subsets of times $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$. In this case, \mathbb{P}_{Λ_k} has to be a marginal of \mathbb{P}_{Λ_K} ; if $\Lambda \supseteq \Lambda_K \supseteq \Lambda_k$, then

$$\Lambda \setminus \Lambda_k = (\Lambda_K \setminus \Lambda_k) \cup (\Lambda \setminus \Lambda_K), \quad (3.9)$$

and consequently

$$\mathbb{P}_{\Lambda_k} = \sum_{\Lambda \setminus \Lambda_k} \mathbb{P}_\Lambda = \sum_{\Lambda_K \setminus \Lambda_k} \sum_{\Lambda \setminus \Lambda_K} \mathbb{P}_\Lambda = \sum_{\Lambda_K \setminus \Lambda_k} \mathbb{P}_{\Lambda_K}. \quad (3.10)$$

Expressed in the notation introduced above, we have $\mathbb{P}_{\Lambda_k} = \mathbb{P}_{\Lambda_K}^{\Lambda_k}$ for all $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$. Intuitively, this means that \mathbb{P}_{Λ_K} , the descriptor of the stochastic process on the times Λ_K , contains all information about subprocesses on fewer times.

While an underlying process leads to a family of compatible finite probability distributions, the KET shows that the converse is also true. Any family of consistent probability distributions – in the sense of Eq. (3.10) – implies the existence of an underlying process. Specifically, the Kolmogorov extension theorem [69–72] defines the minimal properties finite probability distributions have to satisfy in order for an underlying process to exist:

THEOREM 3.1 (KOLMOGOROV EXTENSION THEOREM): *Let Λ be a set of times. For each finite $\Lambda_k \subseteq \Lambda$, let \mathbb{P}_{Λ_k} be a (sufficiently regular) k -step joint probability distribution. There exists an underlying stochastic process \mathbb{P}_Λ that satisfies $\mathbb{P}_{\Lambda_k} = \mathbb{P}_\Lambda^{\Lambda_k}$ for all finite $\Lambda_k \subseteq \Lambda$ iff $\mathbb{P}_{\Lambda_k} = \mathbb{P}_{\Lambda_K}^{\Lambda_k}$ for all $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$.*

In other words, if a family of joint probability distributions on finite sets of times satisfies a consistency condition (as well as an additional minor regularity property [71, 72]), then there is an underlying stochastic process on Λ that has the distributions $\{\mathbb{P}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda}$ as marginals. As stated above, the KET defines the notion of a classical stochastic process in an operational way, and reconciles the existence of an underlying process with its finite time manifestations.

Importantly, in the (physically relevant) case where Λ is an infinite set, the probability distribution \mathbb{P}_Λ can generally not be experimentally accessed. For example, for Brownian

motion, the set Λ could contain all times in the interval $[0, T]$ and each realization \mathbf{x}_Λ would represent a possible continuous sequence of outcomes over this time interval, which we previously dubbed a trajectory of the particle. While we assume the existence of these underlying trajectories (and hence the existence of \mathbb{P}_Λ), in experiments concerning Brownian motion, one can only access \mathbb{P}_{Λ_k} , *i.e.*, their manifestations on finite sets Λ_k . The KET bridges the gap between the finite experimental reality and the underlying infinite stochastic process. Additionally, it also enables the *modeling* of stochastic processes: Any mechanism that leads to finite joint probability distributions that satisfy a consistency condition is ensured to have an underlying process. For example, the proof of the existence of Brownian motion relies on the KET as a fundamental ingredient [88–91].

We can now better appreciate earlier statements about the role of invasiveness: Loosely speaking, the KET holds for classical stochastic processes, because there is no difference between ‘doing nothing’ and conducting a measurement but ‘not looking at the outcomes’ (*i.e.*, summing over the outcomes at a time). Put differently, the validity of the KET is based on the fundamental assumption that the interrogation of a system does not, on average, influence its state. Consequently, marginalization is the correct way to obtain descriptors for fewer times and any classical stochastic process leads to compatible finite joint probability distributions; this compatibility is independent of whether the system was observed or not, and the converse also holds.

This fails to be true in causal modeling scenarios and, in particular, in quantum mechanics. Intuitively, both active and unavoidable interventions, as they appear in classical and quantum causal modeling, overall change the state of the system they act upon, and influence future statistics. Collecting statistical data for given instruments then does *not* allow one to predict what would have happened if no intervention had been performed. In mathematical terms, the overall CPTP map corresponding to an instrument \mathcal{J}_α at t_α does generally not coincide with the ‘do-nothing’ map of the underlying theory. Averaging \mathbb{P}_{Λ_k} over the outcomes at a time t_α then only allows one to predict the statistics of measurements given that \mathcal{J}_α was implemented at t_α , but not the statistics of measurements if the ‘do-nothing’ operation was ‘performed’ at t_α .

To better pinpoint the breakdown of the KET in theories with interventions, we shall discuss it in some more detail in the following two sections, before showing that higher-order quantum maps are the correct conceptual tool to overcome the problems posed by the invasiveness of many physical theories.

3.5 THE KET AND CLASSICAL CAUSAL MODELING

As we have discussed above, satisfaction of the compatibility property of joint probability distributions for different sets of times hinges on the fact that observations in classical physics do not alter the state of the system that is being observed. In contrast to passive interrogations, that merely reveal information, active interventions, like they are used in the case of causal modeling, on average change the state of the interrogated system. Thus the future statistics after

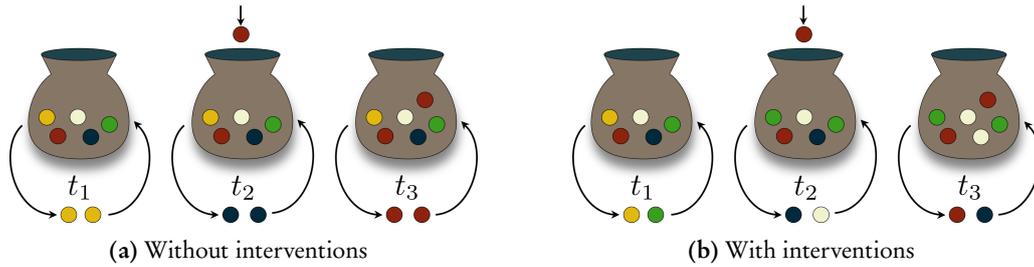


Figure 3.4: (a) Classical process without intervention. A possible three step process is the drawing with replacement of colored balls from an urn. In our example, independent of the actions of the experimenter, a red ball drops into the urn at t_2 (this could, e.g., represent the interaction with an uncontrollable environment.). The experimenter can deduce the joint probability distribution $\mathbb{P}_{\{t_3, t_2, t_1\}}(c_{j_3}, c_{j_2}, c_{j_1})$, to draw different sequences of colors. $\mathbb{P}_{\{t_3, t_2, t_1\}}$ contains all distributions for fewer times, for example $\mathbb{P}_{\{t_3, t_1\}}$, $\mathbb{P}_{\{t_3, t_2\}}$, and $\mathbb{P}_{\{t_1\}}$. (b) Classical process with intervention. Instead of putting the same ball back in the urn, the experimenter could exchange it with a different color (for example, upon drawing yellow, they could replace it with green at t_1 , replace blue with white at t_2 and replace red with blue at t_3). The respective replacement rules are encapsulated in the instruments \mathcal{J}_3 , \mathcal{J}_2 , and \mathcal{J}_1 . Now, from the probability distribution $\mathbb{P}_{\{t_3, t_2, t_1\}}(c_{j_3}, c_{j_2}, c_{j_1} | \mathcal{J}_3, \mathcal{J}_2, \mathcal{J}_1)$, it is generally not possible to deduce probability distributions for fewer steps, like, e.g., $\mathbb{P}_{\{t_3, t_1\}}(c_{j_3}, c_{j_1} | \mathcal{J}_3, \mathcal{J}_1)$, or $\mathbb{P}_{\{t_3, t_2\}}(c_{j_3}, c_{j_2} | \mathcal{J}_3, \mathcal{J}_2)$. This lack of consistency can not be remedied by simple relabeling of the times due to the red ball that drops into the urn at t_2 .

an intervention crucially depends on how the system was manipulated and the prerequisite of compatible joint probability distributions is generally not fulfilled anymore.

Consider, for example, the case of a pharmacologist who tries to understand the effect of different drugs they developed on a disease. In our simplified example, let the disease have two different symptoms S_a and S_b , and denote the absence of symptoms S_c . Whenever the pharmacologist observes S_a , they administer drug D_a , whenever they observe symptom S_b they administer drug D_b , and whenever they observe S_c they do nothing; this choice of actions defines an instrument \mathcal{J} . Running their trial with sufficiently many patients, the pharmacologist can deduce probability distributions for the occurrence of symptoms over time, given the drugs that were administered. For example, if the drugs were administered on three consecutive days, they would have obtained a probability distribution of the form $\mathbb{P}_{\Lambda_3}(s_{j_3}, s_{j_2}, s_{j_1} | \mathcal{J}_3 = \mathcal{J}, \mathcal{J}_2 = \mathcal{J}, \mathcal{J}_1 = \mathcal{J})$, where $s_{j_\alpha} \in \{S_a, S_b, S_c\}$, and the instruments (*i.e.*, the drug administration rule) are the same each day. However, this data would not allow them to find out what would have happened, had they *not* administered drugs on day two, *i.e.*, $\sum_{j_2} \mathbb{P}_{\Lambda_3}(s_{j_3}, s_{j_2}, s_{j_1} | \mathcal{J}_3, \mathcal{J}_2, \mathcal{J}_1) \neq \mathbb{P}_{\Lambda_3 \setminus \{t_2\}}(s_{j_3}, s_{j_1} | \mathcal{J}_3, \mathcal{J}_1)$; intermediate interventions change the state of the interrogated system, and hence the future statistics that are being observed. Nonetheless, it would seem odd in this situation to give up the idea of an underlying course of the disease that reacts in a well-defined way to the administration of drugs, simply because the compatibility condition does not hold anymore. For another, more numerically tangible example of the breakdown of consistency conditions in classical physics, see Fig. 3.4.

It is instructive to reiterate the underlying mathematical reason for the respective existence and absence of Kolmogorov conditions in classical stochastic processes and more general theories. As mentioned, a measurement at t_α in the computational basis yielding outcome x_{j_α}

is represented by the CP map $M_{x_{j_\alpha}} = |x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|$, and the corresponding overall CPTP map is

$$D = \sum_{j_\alpha} |x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|, \quad (3.11)$$

which we will call the *completely dephasing* map, as its action on an arbitrary state ρ consists of eliminating all the off-diagonal terms in the computational basis, *i.e.*,

$$\mathcal{D}[\rho] = \sum_{j_\alpha} \langle x_{j_\alpha} | \rho | x_{j_\alpha} \rangle |x_{j_\alpha}\rangle\langle x_{j_\alpha}| = \sum_{j_\alpha} \rho_{x_{j_\alpha} x_{j_\alpha}} |x_{j_\alpha}\rangle\langle x_{j_\alpha}|. \quad (3.12)$$

In general, when coherences play a role for the dynamics (see Ch. 5), \mathcal{D} does not coincide with the ‘do-nothing’ operation \mathcal{I} , but it does so on the set of states that are diagonal in the measurement basis, which is exactly the set of states we are concerned with in classical physics. In this sense, projective measurements in the computational basis are *non-invasive* on the set of classical states. Consequently, collecting data for measurement outcomes in the computational basis allows one to make predictions about what would have happened, had one *not* performed an operation at a given time step. In Ch. 5, we shall turn this reasoning on its head and use the requirement of non-invasiveness of classical measurements as a definition of the set of classical processes.

On the other hand the CPTP map corresponding to a classical instrument *with active interventions* does not leave states in the classical basis unchanged. For example, for the instrument that prepares a fresh state y_{ℓ_α} upon measuring outcome x_{j_α} (for an illustration, see Fig. 3.4), the overall CPTP map is given by

$$M = \sum_{\ell_\alpha j_\alpha} |y_{\ell_\alpha}\rangle\langle y_{\ell_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|, \quad (3.13)$$

and in general $M \star \rho \neq \rho$, even for states ρ that are diagonal in the classical basis. Consequently, classical causal modeling does generally not display containment properties. Now, after having discussed the classical scenario and the breakdown of the KET as soon as active interventions are introduced, it is time to use the lessons we learned and apply them to the quantum realm.

3.6 QUANTUM MECHANICS AND COMPATIBILITY CONDITIONS

A comprehensive description of quantum mechanical processes must necessarily account for the fundamental invasiveness of measurements, which renders the KET invalid for the same reason that some choices of intervention do in the case of classical causal modeling. To see how even projective measurements in quantum mechanics lead to families of probability distributions that do not satisfy the compatibility conditions of the KET, we consider the following concatenated Stern-Gerlach experiment: Let the initial state of a spin- $\frac{1}{2}$ particle be $|+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$, where $|\uparrow\rangle$ and $|\downarrow\rangle$ are the spin-up and spin-down state in the Z -direction, respectively. Now, we measure the state sequentially in the Z -, X - and Z -direction at times t_1, t_2 and t_3 , and for simplicity we assume that the dynamics between measurements is trivial, *i.e.*, the identity map. These measurements have the possible outcomes $\{\uparrow, \downarrow\}$ (for the measurement in Z -direction

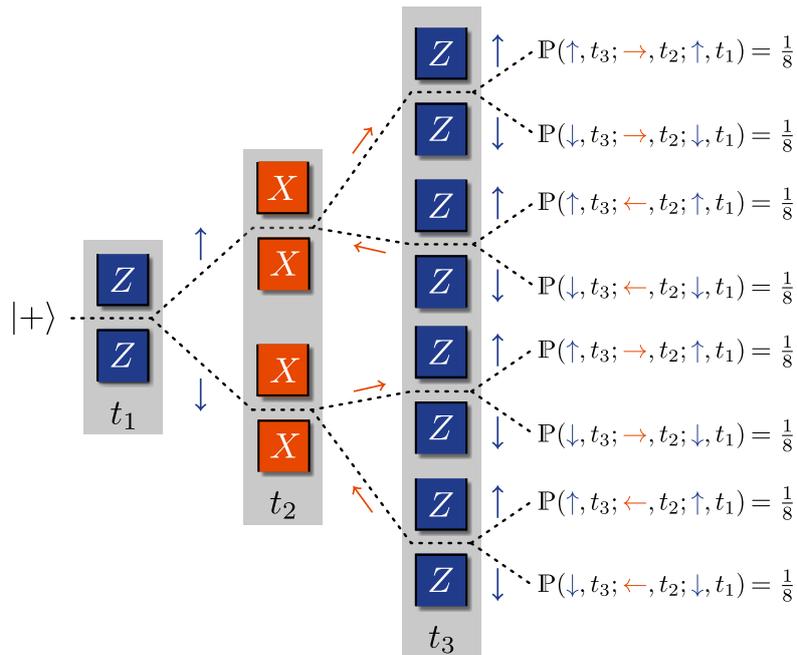


Figure 3.5: Concatenated Stern-Gerlach experiment. The initial state $|+\rangle$ is measured sequentially at times t_1, t_2 , and t_3 in the Z-, X-, and Z-direction, and trivial dynamics is assumed in between measurements. For compactness, the dependence of the joint probabilities on the employed instrument is omitted.

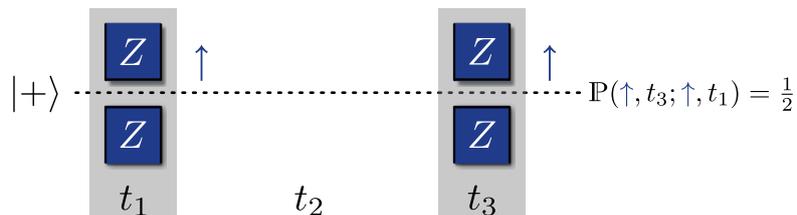


Figure 3.6: Concatenated Stern-Gerlach experiment (two-step). The initial state $|+\rangle$ is measured sequentially in the Z-direction at times t_1 and t_3 , but *not* at time t_2 .

and $\{\rightarrow, \leftarrow\}$ (for the measurement in X-direction). It is easy to see that the probability for any possible sequence of outcomes is equal to $1/8$ (see Fig. 3.5). For instance, we have

$$\mathbb{P}_{\Lambda_3}(\uparrow, \rightarrow, \uparrow | \mathcal{J}_z, \mathcal{J}_x, \mathcal{J}_z) = \mathbb{P}_{\Lambda_3}(\uparrow, \leftarrow, \uparrow | \mathcal{J}_z, \mathcal{J}_x, \mathcal{J}_z) = \frac{1}{8}, \quad (3.14)$$

where \mathcal{J}_z and \mathcal{J}_x represent the instruments used to measure in the Z- and X- direction respectively, and $\Lambda_3 = \{t_3, t_2, t_1\}$. Summing over the outcomes at time t_2 , we obtain, for example, the marginal probability $\mathbb{P}_{\Lambda_3}^{\{t_1, t_3\}}(\uparrow, \uparrow | \mathcal{J}_z, \mathcal{J}_z) = 1/4$. However, by considering the case where the measurement is *not* made at t_2 , it is easy to see that $\mathbb{P}_{\{t_3, t_1\}}(\uparrow, \uparrow | \mathcal{J}_z, \mathcal{J}_z) = 1/2$ (see Fig. 3.6). The intermediate measurement changes the state of the system, and the corresponding probability distributions for different sets of times are no longer compatible [43, 65].

While, in this example, we have assumed trivial dynamics in between measurements, and different instruments at different times, we could readily construct an example where the measuring devices are the same at every time, but the corresponding joint probability distributions

still fail to satisfy consistency conditions. For example, a dynamics between measurements given by a Hadamard gate, respectively, would lead to a violation of Kolmogorov conditions in the above example, even if all the measurements were performed in the Z -basis. More generally, the breakdown of consistency conditions is the norm in quantum mechanics, rather than the exception [65, 80].

It is insightful to highlight the close relation of this breakdown of consistency, and the violation of Leggett-Garg inequalities in quantum mechanics [73, 75]. The assumption of consistency between descriptors for different sets of times that is crucial for the derivation of the KET subsumes the assumptions of realism and non-invasive measurability that are the basic principles leading to the derivation of Leggett-Garg inequalities: While realism implies that joint probability distributions for a set of times can be expressed as marginals of a joint probability distribution for more times, non-invasiveness means that all finite distributions are marginals of the *same* distribution. For example, the two-step joint probability distributions $\mathbb{P}_{\{t_2, t_1\}}$, $\mathbb{P}_{\{t_3, t_2\}}$, and $\mathbb{P}_{\{t_3, t_1\}}$, that are considered in the Leggett-Garg scenario are all marginals of a three-step distribution $\mathbb{P}_{\{t_3, t_2, t_1\}}$. As soon as non-invasiveness and/or realism do not hold anymore, the KET can fail and Leggett-Garg inequalities can be violated.

Nevertheless, there must be some compatibility between descriptors for different sets of times; the breakdown of the KET should be a problem of the formalism rather than a physical fact. We will see in the next section that a change of perspective enables one to recover a generalized consistency condition and to prove an extension theorem in quantum mechanics and any theory with interventions.

3.7 QUANTUM STOCHASTIC PROCESSES

The conceptual problem of the absence of Kolmogorov conditions, and the lack of a KET in quantum mechanics and other theories with interventions can be remedied by assuming the standpoint of quantum causal modeling, and choosing a description of such stochastic processes that explicitly takes interventions and their corresponding change of the system into account. In quantum mechanics, interventions are represented by CP maps, and a descriptor of a general stochastic process has to – in clear analogy to the classical case – allocate the correct probabilities to any possible sequence of CP maps. Unsurprisingly, such a description is provided by higher-order quantum maps. While we have already alluded to this fact, we are now in a position to fully understand the answers it provides, and with the corresponding framework at hand, it is not only possible to recover a compatibility property that is satisfied by *any* process with interventions, but also to prove a generalized extension theorem that provides the foundation for the mathematically rigorous discussion of quantum stochastic processes. In quantum mechanics, CP maps comprise the most general transformations an experimenter can perform. In principle, in a theory beyond quantum mechanics, the possible operations that are implementable by an experimenter could be of a different mathematical structure, without altering the arguments we make below. Consequently, to keep the ensuing discussion as general as possible, and as such also applicable to potential post-quantum theories,

we shall predominantly phrase the subsequent results in terms of maps instead of Choi states, but the relation to combs is always clear.

Sticking with the language familiar to quantum mechanics – but keeping in mind that a generalization to OPTs is straightforward – each realization of an experiment corresponds to a sequence of CP maps that transform the system at a series of times Λ_k , and the set of possible CP maps that could be applied is dictated by the choice of instruments used to interrogate the system in question. Then, a quantum process is *fully* characterized once *all* of the probabilities $\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1} | \mathcal{J}_{k'}, \dots, \mathcal{J}_1)$ for each such sequence and all possible instruments are known.

Written more succinctly, a k -step quantum stochastic process is fully characterized by an object \mathcal{C}_{Λ_k} that maps sequences of CP maps to probabilities, *i.e.*,

$$\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1} | \mathcal{J}_{k'}, \dots, \mathcal{J}_1) = \mathcal{C}_{\Lambda_k}[\mathcal{M}_{x_{j_k}}, \dots, \mathcal{M}_{x_{j_1}}]. \quad (3.15)$$

See Fig. 3.7 for a graphical representation. *A priori*, the map \mathcal{C}_{Λ_k} does not have to be linear. However, under the assumption that the underlying theory is a probabilistic one – which applies to classical physics, quantum mechanics, and OPTs – it can be shown, that \mathcal{C}_{Λ_k} indeed has to be a linear positive functional [62, 86, 87, 92]; the proof of this fact follows the same line as the general arguments for the linearity of quantum maps we outlined in Sec. 2.1.

In this sense, the action of \mathcal{C}_{Λ_k} represents a Born rule for temporal processes and is a positive multilinear functional that can be reconstructed in a finite number of experiments [2, 8, 31, 32]. As we have seen, in quantum mechanics, \mathcal{C}_{Λ_k} constitutes a special case of a quantum comb [87], and we have

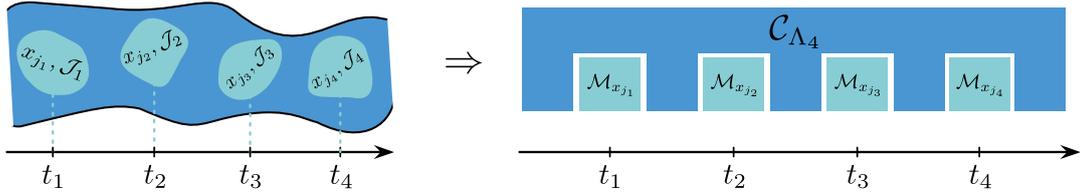
$$\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1} | \mathcal{J}_{k'}, \dots, \mathcal{J}_1) = \mathcal{C}_{\Lambda_k} \star_{\alpha \in \Lambda_k} \mathcal{M}_{x_{j_\alpha}} \quad (3.16)$$

Even though this fact appears natural, it is worth recalling that we originally introduced combs as arising from general quantum networks, while here we are looking for mathematical objects that can describe experimentally accessible joint probability distributions. We will make this connection more transparent in the subsequent chapters. Here, it arises directly from the requirements of linearity and positivity that we impose on \mathcal{C}_{Λ_k} .

Additionally, it is important to keep in mind that the theory of combs, and in particular their representation in terms of Choi states, assumes the validity of quantum mechanics, while the arguments we make throughout this chapter only necessitate that \mathcal{C}_{Λ_k} acts linearly on the space of implementable sequences of operations.⁷ Nonetheless, as it is helpful to build intuition, we shall keep using the terminology of quantum mechanics and thus, we will call \mathcal{C}_{Λ_k} a k -slot comb.

The comb \mathcal{C}_{Λ_k} contains all the multi-time correlations necessary to fully characterize a k -step quantum process and as such allows one to unambiguously quantify memory effects in quantum mechanics (see Ch. 4). Importantly, while the CP maps $\mathcal{M}_{x_{j_\alpha}}$ change the state of the system that is interrogated, they do *not* change the k -time process given by \mathcal{C}_{Λ_k} . Loosely speaking, the comb contains all parts of the dynamics that are not manipulated by and/or accessible to the experimenter. This is analogous to the way in which the preparation of an initial state and the measurement of the final state in quantum process tomography do not

⁷ We also assume that all \mathcal{C}_{Λ_k} are bounded, which is trivially satisfied, as the relevant image space of \mathcal{C}_{Λ_k} is $[0, 1]$.



$$= \mathcal{C}_{\Lambda_4}[\mathcal{M}_{x_{j_4}}, \dots, \mathcal{M}_{x_{j_1}}] = \mathbb{P}_{\Lambda_4}(x_{j_4}, \dots, x_{j_1} | \mathcal{J}_4, \dots, \mathcal{J}_1)$$

Figure 3.7: Graphical representation of a four step process. For $\Lambda_4 = \{t_4, \dots, t_1\}$, the process – represented by the blue shape on the left – is interrogated four times using instruments $\mathcal{J}_4, \dots, \mathcal{J}_1$ with outcomes x_{j_4}, \dots, x_{j_1} . This situation can be represented diagrammatically as a comb \mathcal{C}_{Λ_4} acting on the corresponding operations $\mathcal{M}_{x_{j_4}}, \dots, \mathcal{M}_{x_{j_1}}$. \mathcal{C}_{Λ_4} encodes all multi-time correlations between observables at the times Λ_4 , and is independent of what instruments are used to probe it.

influence the underlying dynamics (*i.e.*, the CPTP map connecting input and output state [19, 79, 93]). As such, our description of stochastic processes allows one to clearly delineate between what constitutes the process, and what constitutes interventions, which is not possible in different – incomplete – approaches to quantum stochastic processes [71, 80].

Just as in the classical case, the knowledge of all relevant joint probability distributions (*i.e.*, the knowledge of \mathcal{C}_{Λ_k}) allows one to deduce causal relations between the k events in Λ_k . This forms the basis of the field of quantum causal modeling [50, 51]. Evidently, both classical causal modeling as well as classical stochastic processes are included in the quantum causal modeling framework as special cases, which gives our descriptor \mathcal{C}_{Λ_k} the universality we were aiming fore.

In more detail, whenever a system is measured and prepared in a fixed basis (using a classical instrument), and the identity operation \mathcal{I} and the completely dephasing operation \mathcal{D} act indistinguishably (from the perspective of a classical observer), then the resulting set of joint distributions is consistent with a classical causal model. We will further elaborate on processes that appear classical for an observer that probes them with classical instruments in Ch. 5. As we have seen, classical stochastic processes are a special case of this setup, where all the instruments \mathcal{J}_α are the idle instrument $\mathcal{J}_{\text{id}} = \{|x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|\}_{j_\alpha=1}^d$ and the families of joint probability distributions obtained via

$$\mathbb{P}_{\Lambda_k}(x_{j_k}, \dots, x_{j_1} | \mathcal{J}_k, \dots, \mathcal{J}_1) = \mathcal{C}_{\Lambda_k} \star_{\alpha \in \Lambda_k} (|x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|) \quad (3.17)$$

satisfy Kolmogorov conditions. The naturally arising question, if the KET can be proven in an alternative way, using the comb framework, will be answered in the affirmative below (see Sec. 3.8.2). Now, we will use our extended understanding of stochastic processes, to recover compatibility conditions for general processes, and to prove the generalized extension theorem, that guarantees their existence.

3.8 GENERALIZED EXTENSION THEOREM (GET)

The complete description of general processes on finite numbers of times that we developed above allows us to straightforwardly formalize generalized Kolmogorov conditions between

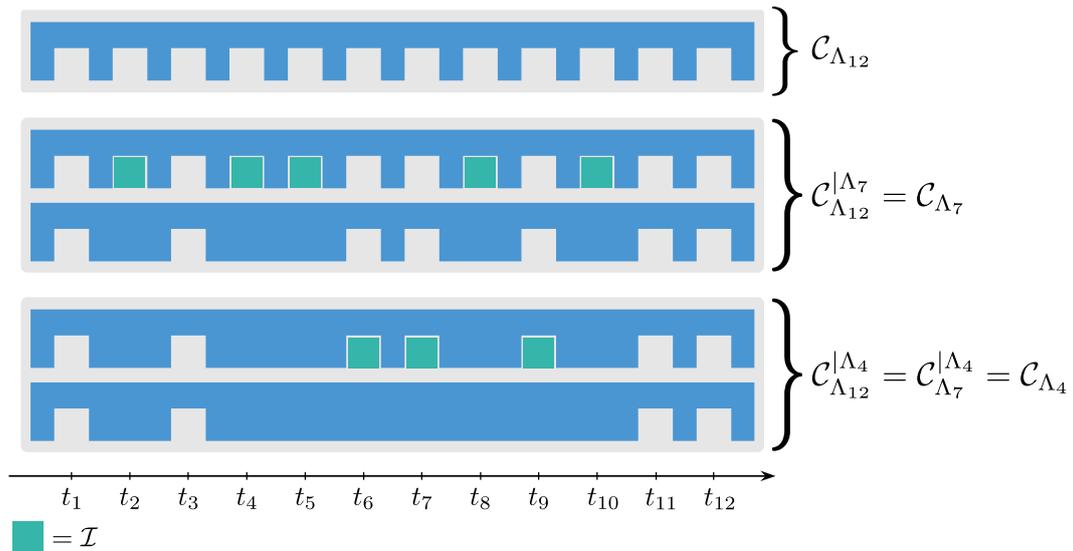


Figure 3.8: Consistency condition for combs. If there is an underlying process, any comb \mathcal{C}_{Λ_k} can be obtained from \mathcal{C}_{Λ_K} by letting \mathcal{C}_{Λ_K} act on the identity map at the excessive times. Here, for the sets of times $\Lambda_{12} = \{t_{12}, \dots, t_1\}$, $\Lambda_7 = \{t_{12}, t_{11}, t_9, t_7, t_6, t_3, t_1\}$ and $\Lambda_4 = \{t_{12}, t_{11}, t_3, t_1\}$, the containment of the comb \mathcal{C}_{Λ_7} in $\mathcal{C}_{\Lambda_{12}}$ and the containment of \mathcal{C}_{Λ_4} in both $\mathcal{C}_{\Lambda_{12}}$ and \mathcal{C}_{Λ_7} is depicted.

processes defined on different sets of times. Naturally, if there is an underlying process, then families of combs for different sets of times satisfy the following consistency condition [32]: for any two sets of times $\Lambda_k \subseteq \Lambda_K$, the comb \mathcal{C}_{Λ_k} can be obtained from \mathcal{C}_{Λ_K} by letting the latter act on identity operations \mathcal{I}_α at times $t_\alpha \in \Lambda_K \setminus \Lambda_k$, *i.e.*,

$$\mathcal{C}_{\Lambda_k}[\cdot] = \mathcal{C}_{\Lambda_K}[\mathcal{I}_{\Lambda_K \setminus \Lambda_k} \cdot] := \mathcal{C}_{\Lambda_K}^{\Lambda_k}[\cdot], \quad (3.18)$$

where we have employed the shorthand notation $\mathcal{I}_{\Lambda_K \setminus \Lambda_k}$ to signify that the identity operation was ‘implemented’ at each time $t_\alpha \in \Lambda_K \setminus \Lambda_k$, and \cdot is a placeholder for operations at the remaining times (in this case operations at times in Λ_k). The graphical representation of Eq. (3.18) is depicted in Fig. 3.8.

It is important to note the difference between Eq. (3.18) and the consistency condition for classical stochastic processes, stemming from the stronger notion of ‘doing nothing’ in the quantum case. If there is an underlying process, any comb can be obtained from one that applies to a larger set of times by letting it act on the identity map at the excessive times. This is by no means the same as computing the marginals of families of probability distributions that have been obtained for a fixed set of measurement instruments; rather, marginalization over a time t_α in this sense amounts to letting the comb act on M_α , the CPTP map corresponding to the instrument \mathcal{I}_α . By now, it should be clear why this difference does not occur in classical stochastic processes; there, averaging over a time step t_α is equivalent to letting the respective comb act on the completely dephasing map D_α . As the action of this map cannot be distinguished from the action of the ‘do-nothing’ map \mathcal{I}_α for states that are diagonal

in the measurement basis, marginalization and not performing an operation lead to the same statistics.⁸

On the other hand, for more general processes, this distinction can be made, and the correct way of ‘marginalization’ is obtained by letting the comb act on identity maps at the excessive time steps, instead of completely dephasing maps (or any other CPTP map). Put differently, the way of marginalization we provide in Eq. (3.18) is also correct for classical stochastic processes, but not the other way round.

We recover descriptors for different sets of times that are compatible with each other only when we switch to a causal modeling description of the process, *i.e.*, only when we actively account for interventions. From this, we obtain the main result of this chapter, the generalized extension theorem (GET) [4]:

THEOREM 3.2 (GENERALIZED EXTENSION THEOREM): *Let Λ be a set of times. For each finite $\Lambda_k \subseteq \Lambda$ let \mathcal{C}_{Λ_k} be a k -slot comb. There exists a general stochastic process \mathcal{C}_Λ that satisfies $\mathcal{C}_{\Lambda_k} = \mathcal{C}_\Lambda^{\downarrow \Lambda_k}$, as defined in Eq. (3.18), for all finite $\Lambda_k \subseteq \Lambda$, iff $\mathcal{C}_{\Lambda_k} = \mathcal{C}_{\Lambda_K}^{\downarrow \Lambda_k}$ for all finite $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$.*

The proof can be found in App. A.2. It proceeds analogously to that of the original Kolmogorov extension theorem, presented, for example, in [72]; the consistency property is used to uniquely define a comb $\mathcal{C}_\Lambda^\sharp$ on a ‘sufficiently large’ container space, that contains all finite combs as ‘marginals’ in the correct sense. Due to its linearity and boundedness $\mathcal{C}_\Lambda^\sharp$ can then be extended to a linear functional $\bar{\mathcal{C}}_\Lambda^\sharp$ that fulfills the properties of the comb \mathcal{C}_Λ on the closure of said container space.

As in the classical case [71], the underlying stochastic process \mathcal{C}_Λ is – unlike $\bar{\mathcal{C}}_\Lambda^\sharp$ – not necessarily unique, which means that there potentially exist many different \mathcal{C}_Λ that have the family of combs $\{\mathcal{C}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda_K}$ on finite sets of times as marginals. Since the action of all of these possible \mathcal{C}_Λ coincides with the unique $\bar{\mathcal{C}}_\Lambda^\sharp$ on a sufficiently large space, and hence yields the correct finite combs \mathcal{C}_{Λ_k} , this non-uniqueness cannot be detected experimentally and does not constitute a practical problem.

As alluded to throughout this chapter, even though we have phrased it in the language of quantum mechanics, there is nothing particularly quantum mechanical about the GET. The proof of the theorem only uses the linearity and boundedness of the maps \mathcal{C}_{Λ_k} , as well as their compatibility. Consequently, it holds for *any* probabilistic theory (importantly including those with interventions) that can be phrased in terms of *linear* functionals acting on sequences of interventions.

Furthermore, the input and output spaces of the CP maps the comb acts on do not have to be of the same dimension. If they differ, the identity map used for the consistency condition has to be slightly generalized: A CPTP map⁹ $\mathcal{M}_\alpha : \mathcal{B}(\mathcal{H}_\alpha^o) \rightarrow \mathcal{B}(\mathcal{H}_\alpha^i)$, is implemented via a

⁸ In fact, we can relax this statement, and consider a process to be classical, if the action of \mathcal{I} and \mathcal{D} can not be distinguished by performing *classical measurements*. We discuss this in Ch. 5. Here, this technicality has no influence on the motivations we provide.

⁹ Somewhat counter-intuitively, \mathcal{M}_α maps from a space labeled by o to a space labeled by i . This is in order to abide by the notational conventions we introduced in the previous chapter, where we defined inputs and outputs with respect to the comb, and not the operations that are plugged into it.

corresponding unitary U_α , a *fixed* ancillary state $\eta_\alpha \in \mathcal{B}(\mathcal{H}_{A_\alpha})$, and a partial trace tr_{B_α} that is such that the resulting output state

$$\mathcal{M}_\alpha[\rho] = \text{tr}_{B_\alpha} \left[U_\alpha (\rho \otimes \eta_\alpha) U_\alpha^\dagger \right] \quad (3.19)$$

lies in $\mathcal{B}(\mathcal{H}_\alpha^1)$ [94] (see Ch. 4 for a discussion of the unitary dilation of CPTP maps and combs). With this, we can define a *generalized identity map* $\mathcal{I}_\alpha^{(o \rightarrow i)}[\rho] = \text{tr}_{B_\alpha} (\rho \otimes \eta_\alpha)$ and the GET still applies, when marginalization is understood with respect to $\mathcal{I}_\alpha^{(o \rightarrow i)}$. Consequently, our theorem accounts for the case where particles are created/annihilated in the process, as well as the case where different degrees of freedom are manipulated at each time t_α , or where the number of measurement outcomes and active interventions differ.

More fundamentally, in the derivation of the GET, we make the implicit assumption that probabilities only depend on the respective CP maps that were implemented, but not on the particular instrument that was used to implement them. This property has been dubbed ‘instrument non-contextuality’ [24, 87] or ‘operational instrument equivalence’ [43]. In principle, our derivation could be straightforwardly adapted to any theory, where this assumption is not satisfied anymore, but probabilities are still a linear function of the maps *and* their respective contexts (*i.e.*, the respective instrument). Instead of the identity map, one would then use a pair $(\mathcal{I}, \mathcal{J}_\mathcal{I})$ of identity map *and* identity context for marginalization, and the GET would still hold.

Even though we have not yet explicitly discussed the physical emergence of quantum stochastic processes, the GET tells us what minimal properties they must satisfy, and what mathematical framework we have to employ to describe them. In this sense, just like the KET for classical stochastic processes, the GET provides the fundamental theorem for the discussion of quantum stochastic processes and establishes combs as their universal descriptor. As we will see in the following chapter, combs can also be established in a less axiomatic, but more physically motivated way as the framework to treat general quantum dynamics.

As in the classical case, the proof of the GET does not assume an *a priori* temporal ordering. The sets Λ_k could be sets of times, but also labels of different laboratories without a well-defined order. We have the following remark:

REMARK 3.1: *The proof of the GET does not assume any ordering of the sets Λ_k , and only uses the generalized containment property (3.18) as its main ingredient.*

Consequently, the GET also applies to causally non-separable processes [54, 95], as the descriptors for different sets of laboratories would still satisfy a compatibility condition. However, these processes do not have a deterministic Stinespring dilation [53], *i.e.*, there is – to date – no known (non-probabilistic) dynamical mechanism that leads to them, which makes the interpretation of an underlying ‘process’ much less clear in the absence of a definitive causal ordering. While we shall briefly remark on this fact in the conclusion of this chapter, the full exploration of this interpretation is left as an open question for future work. Next, we will see that the distinction between stochastic processes and causal modeling does not exist in the general case.

3.8.1 *Quantum Stochastic Processes and Quantum Causal Modeling*

We have seen that classical stochastic processes are a strict subset of the set of processes that make up the theory of classical modeling, and both of them are subsets of quantum causal modeling. It remains to discuss the role of quantum stochastic processes in this hierarchy.

Given the results of the previous discussion, we can cut this analysis short: in quantum mechanics, it is unavoidable to employ a description that takes interventions into account, when attempting to obtain a consistent description of a quantum process; if one wants to properly define quantum stochastic processes, one is directly forced to use a framework that can account for all possible interventions. Consequently, we have the following proposition:

PROPOSITION 3.1: *The theory of quantum causal modeling and the theory of quantum stochastic processes are equivalent.*

In contrast to the classical case, the set of quantum causal models does not just contain the set of quantum stochastic processes, but coincides with it.

Besides its appeal in terms of a full classification of causal modeling and stochastic processes, this proposition has an additional consequence: it implies that the breakdown of the KET in quantum mechanics is *fundamental*, while, in principle, it can be removed in classical processes with interventions by changing perspective. In the latter case, a *super-observer*, one that observes both the experimenter manipulating the system of interest as well as the stochastic process itself, would obtain families of joint probability distributions that display a compatibility property in the sense of the KET. Put differently, for classical processes, by incorporating the experimenter and their choices of instruments into the stochastic process, the KET always applies on a higher level.

In quantum mechanics, this is generally not true. No matter the level at which a super-observer observes a process, the respective joint probability distributions do not satisfy a compatibility property, and the KET fails to hold. This fundamental breakdown of the KET in quantum mechanics is also mirrored by no-go theorems that show that non-contextual theories cannot reproduce the predictions of quantum mechanics; for many of these theorems, the notion of ontic latent variables [96, 97] or ontic processes [43] are introduced, and the basic assumption is made that the distributions over observable outcomes can be obtained by marginalization of a larger joint distribution over the values of the ontic variable. Subsequently, it is shown that, together with other assumptions, this prerequisite fails to reproduce predictions made by quantum mechanics. The GET dictates how to correctly compute marginals in quantum mechanics, such that all resulting probability distributions ‘fit together’ and are the marginals of one common comb \mathcal{C}_Λ . It is therefore conceivable that a derivation in spirit of the aforementioned references that starts from the assumption of compatibility in the sense of the GET would lead to theories that can indeed reproduce quantum mechanics.

We reiterate that classical stochastic processes are a very special subset of general stochastic processes, namely, the ones where the experimenter can only perform projective measurements in a fixed basis, and the resulting joint probability distributions satisfy Kolmogorov conditions. To conclude the discussion of the general extension theorem, we now show explicitly that it contains the KET as a corollary.

3.8.2 $GET \Rightarrow KET$

Our generalized extension theorem applies to a strictly larger class of theories than the standard KET and includes the latter as a corollary. We have the following proposition:

PROPOSITION 3.2: *The GET implies the KET.*

The detailed proof of this statement can be found in App. A.3. There, we show that a family $\{\mathbb{P}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda}$ of compatible finite probability distributions can be mapped onto a family of combs $\mathcal{C}_{\Lambda_k}^{\text{cl}}$ that satisfy the consistency condition of the GET. The existence of the underlying process $\mathcal{C}_{\Lambda}^{\text{cl}}$ then ensures the existence of a joint probability distribution \mathbb{P}_{Λ} that has all finite ones as marginals.

While the original version of the KET does not hold for quantum processes, it is important to keep in mind that the breakdown of the compatibility property of joint probability distributions is not a signature of quantum mechanics *per se*; as we have already seen, any framework that allows for interventions will exhibit this feature. The GET provides a proper theoretical underpinning for the corresponding experimental situations. On the other hand, the breakdown of the compatibility property can happen in quantum mechanics even if only projective measurements in a fixed basis $\{|x_{j_a}\rangle\}$ are performed [65, 80].

As already mentioned, the absence of compatibility is tantamount to the absence of either realism, or non-invasiveness (or both). Consequently, it can be used as a definition of non-classicality, as proposed in Ref. [80]. There, the authors employ the breakdown of the consistency condition on the level of probability distributions, when measuring in a fixed basis, as a means to define the non-classicality of *Markovian*, *i.e.*, memoryless, processes. In Ch. 5, we shall follow this definition to analyze the set of classical *non-Markovian* processes by means of higher-order maps.

3.9 RELATION TO PREVIOUS WORK

The proof of the GET does not rely on any particularities that are exclusive to quantum mechanics or our formulation thereof. As such, our extension theorem constitutes a sound basis for the description of *any* conceivable (classical, quantum or beyond) theory of stochastic processes with interventions – independent of the employed framework. Here, we want to embed this result into the larger context of the existing literature on quantum stochastic processes (and beyond).

While we referred throughout to the framework of higher-order quantum maps, our results apply equally to any other framework for describing quantum processes as linear functionals. The original idea to phrase quantum stochastic processes in this way dates back to Lindblad [62] and Accardi *et al.* [86, 92]; more recent examples of mathematical objects and frameworks (often the same thing under a different name) given a firm theoretical foundation by the GET include: *process tensors* [8, 31, 32] and *causal automata/non-anticipatory channels* [18, 98], which describe the most general open quantum processes with memory (see Ch. 4); *causal boxes* [99] that enter into quantum networks with modular elements; *operator tensors* [100, 101] and *superdensity matrices* [102], employed to investigate quantum information in general relativistic

space-time; and, finally, *process matrices*, used for quantum causal modeling [42, 50, 51, 54]. For classical stochastic processes, as well as the causal modeling framework we discussed, our result applies to the theory of ε -transducers used within the theory of computational mechanics [103, 104] to describe processes with active interventions.

Our theorem proves the existence of a container space for all of the aforementioned frameworks and allows for their complete and consistent representation in the continuous time limit, thus providing an overarching theorem for probabilistic theories with interventions. This is of particular importance for the field of open quantum mechanics where the lack of an extension theorem has been a roadblock to obtaining a framework that coincides with classical descriptions in the correct limit [65]. Here, switching perspective allows one to describe both classical as well as quantum open systems in a unified framework. This fact has recently been used to obtain an unambiguous definition of non-Markovianity in quantum mechanics that coincides with the classical one in the correct limit [105] (we shall encounter this definition in Ch. 4).

The GET goes beyond previous attempts to generalize the KET for quantum mechanics. An extension theorem for POVMs was derived in Ref. [85] and was used in Ref. [106] to show the existence of an ‘infinite composition’ of an instrument. This extension theorem is, however, limited to particular cases of POVMs, and not general enough to provide an underpinning for the description of stochastic processes with interventions.

More generally, a version of the KET for quantum processes was derived in Ref. [86]. In this work, the authors showed that any quantum stochastic process can be reconstructed ‘up to equivalence from a projective family of correlation kernels’. By decomposing the control operations $\mathcal{M}_{x_{j\alpha}}$ into their component Kraus operators, it can explicitly be shown that these correlation kernels correspond to combs, and consequently, for quantum processes, the GET is equivalent to Thm. 1.3 in Ref. [86]. However, the mathematical structure of the latter does not tie in easily with recently developed frameworks for the description of quantum (or classical) causal modeling, nor does it lend itself in a straightforward way to the discussion of their key properties. The structural features of combs render the investigation of fundamental features of a process, like their non-Markovianity [31, 105], their causal structure [50, 54, 99], and their classicality, tractable.

Our formulation has the advantage that combs are defined in a clear-cut *operational* way, and allow for a generalized Stinespring dilation [17, 32], which makes their interpretation in terms of open quantum system dynamics straightforward. Additionally, even though the GET is stated for combs that map sequences of CP maps to probabilities, its proof also applies – with slight modifications – to general quantum combs (*i.e.*, maps that map combs onto combs [16, 17, 63, 64]).

3.10 QUANTUM STOCHASTIC PROCESSES – OUTLOOK

The generalized extension theorem we derived in this chapter lays the foundation for the discussion of quantum stochastic processes. Beyond this theoretical appeal, it also allows one – just like the KET for classical stochastic processes (see, *e.g.*, Ch. IX of [91]) – to check if a

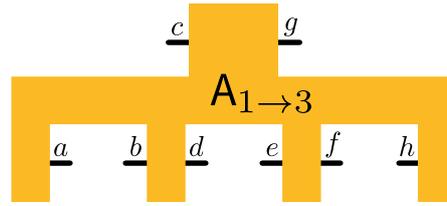


Figure 3.9: Graphical representation of a $1 \rightarrow 3$ adapter $A_{1 \rightarrow 3}$. Any slot of a comb it is plugged into is split into three slots.

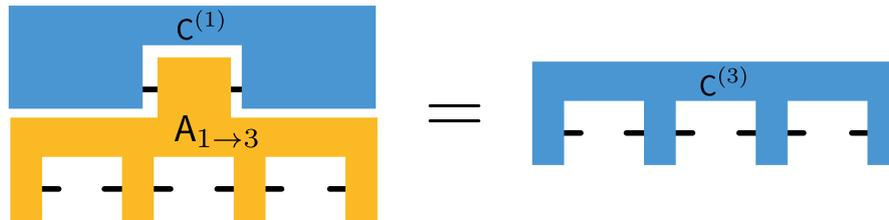


Figure 3.10: Action of an $1 \rightarrow 3$ adapter on a one-slot comb $C^{(1)}$. The resulting three-slot comb $C^{(3)}$ should contain $C^{(1)}$ as a marginal in the sense of the GET.

proposed *generative model* actually leads to a stochastic process; specifically, it does so, if all resulting descriptors on finite sets of times are compatible with each other. With this idea in mind, we propose a toolbox of generative models for quantum stochastic processes, that relies on the GET as its only ingredient, and does not require an underlying Hamiltonian for its derivation.

In classical physics, a generative model defines, how to get from joint probability distributions for a set of times to a family of probability distributions for a larger set of times. For example, for a Markovian process, one could provide conditional probabilities $p(y|x)$ to measure y given that the last outcome was x . With this, and an initial distribution $p(x)$, all multi-time distributions can be constructed [71, 107]. We shall take this idea as a guiding principle – albeit not in a too literal sense – to construct combs that apply to sets of times Λ_K starting from combs defined on fewer times Λ_k . Notably, this construction will be somewhat cumbersome when done algebraically, but very clear when represented graphically, and it is advisable to always consult the corresponding figures to understand the intuitive meaning behind the equations we encounter in this section.

One possible way to construct compatible families of combs is to define a $1 \rightarrow 3$ adapter $A_{1 \rightarrow 3}$, that can be plugged into the slot of a comb and yields a comb with three slots (see Figs. 3.9 and 3.10 for a graphical representation). As we discussed in the previous chapter, such an adapter is, itself, a quantum comb. Now, we can identify the middle slot of $A_{1 \rightarrow 3}$ (*i.e.*, the slot with the wires labeled by d and e in Fig. 3.9) with the original slot the adapter was plugged into, and the remaining two slots correspond to new times. This choice of an adapter is not the only possible one, but suggests itself naturally, due to its simplicity and symmetry. In order for $A_{1 \rightarrow 3}$ to lead to a compatible family of combs, the comb that results from applying the adapter $A_{1 \rightarrow 3}$ has to contain the original comb in the sense of Eq. (3.18), *i.e.*, the old comb and the new comb have to be compatible with each other.

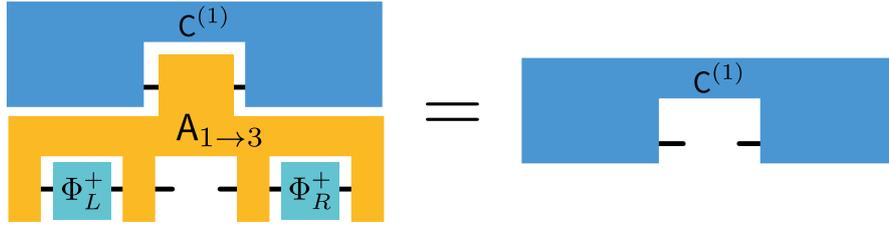


Figure 3.11: Consistency condition for adapters. The ‘old’ comb $C^{(1)}$ has to be contained in the ‘new’ comb, *i.e.*, in the one resulting from by applying the adapter. Here, $C^{(1)}$ is obtained by plugging identities into the slots corresponding to the excessive times of the ‘new’ comb.

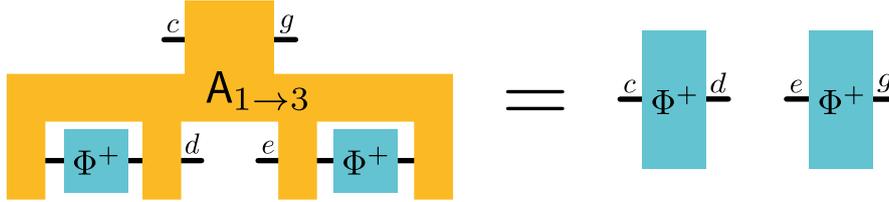


Figure 3.12: Requirement for universal $1 \rightarrow 3$ adapters. In order to yield a family of consistent combs, the adapter has to leave the original comb unchanged, when identity maps are plugged into its left and right slot.

For example, if we start with a one-slot comb $C^{(1)}$, the comb $C^{(3)} = C^{(1)} \star A_{1 \rightarrow 3}$ resulting from applying the adapter, has to satisfy

$$C^{(3)} \star \Phi_L^+ \star \Phi_R^+ = C^{(1)}, \quad (3.20)$$

where Φ_L^+ (Φ_R^+) is the Choi state of the identity map that is plugged into the left (right) slot of $C^{(3)}$ (see Fig. 3.11).

If we want this relation to hold for *any* comb $C^{(1)}$, *i.e.*, if we want $A_{1 \rightarrow 3}$ to be universal, then, using the labeling of Hilbert spaces employed in Fig. 3.12, $A_{1 \rightarrow 3}$ has to satisfy

$$A_{1 \rightarrow 3} \star \Phi_{ba}^+ \star \Phi_{hf}^+ = \Phi_{dc}^+ \otimes \Phi_{ge}^+, \quad (3.21)$$

where Φ_{XY}^+ is the (unnormalized) maximally entangled state on $\mathcal{H}_X \otimes \mathcal{H}_Y$ (see Fig. 3.12). If Eq. (3.21) is satisfied, then the adapter $A_{1 \rightarrow 3}$ leaves the comb it is applied to unchanged when identity maps are plugged into its left and right socket. Consequently, when applying $A_{1 \rightarrow 3}$ to a comb, the resulting object is consistent with the original one in the sense of Eq. (3.18). Now, with such a proper $1 \rightarrow 3$ adapter at hand, we can construct families of compatible combs, starting from a comb $C^{(1)}$, by successively applying the adapter $A_{1 \rightarrow 3}$ to each slot of the comb (see Fig. 3.13 for a graphical representation). Consequently, we obtain combs $C^{(1)}, C^{(3)}, C^{(9)}, \dots$, and we have

$$C^{(3n)} = C^{(n)} \star_{\alpha=1}^n A_{1 \rightarrow 3}^{(\alpha)}, \quad (3.22)$$

where $A_{1 \rightarrow 3}^{(\alpha)}$ is an adapter that fits in the α^{th} slot of $C^{(n)}$. If $A_{1 \rightarrow 3}^{(\alpha)}$ satisfies (3.21), then the resulting family $\{C^{(3n)}\}_{n=1}^{\infty}$ satisfies generalized consistency conditions, and we know from the GET, that this construction actually corresponds to the construction of a quantum stochastic process.

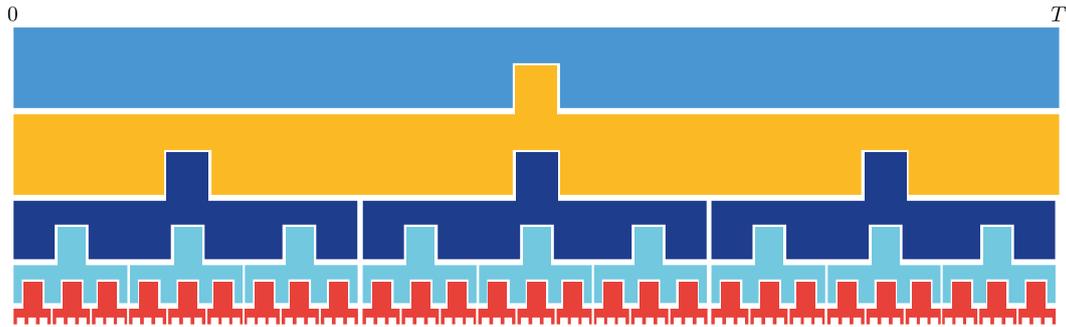


Figure 3.13: Construction of a family of consistent combs. In each new layer, the number of slots is multiplied by three, and each previous layer can be obtained by plugging identity operations into the excessive slots. By proper labeling, all the resulting slots can, for example, correspond to times in the interval $[0, T]$.

By proper labeling, all the resulting slots can correspond to times in the interval $[0, T]$, and the underlying stochastic process C_Λ would be defined on $\Lambda = [0, T]$.

Even though we have discussed the basic properties of $A_{1 \rightarrow 3}$, for an explicit construction, one would still have to fix the internal causal order of the adapter. *Per se*, in Fig. 3.9 the causal ordering between the wires labeled by c and g and the remaining wires is *a priori* unclear, and can be chosen freely.¹⁰ As the ensuing detailed discussion of the properties of adapters $A_{1 \rightarrow 3}$ is somewhat technical, we relegate it to App. A.4. Importantly, though, this discussion shows, that such adapters actually exist, and are not limited to trivial objects.

Quite obviously, the construction we proposed in this section is by no means the only one that yields a generative model for quantum stochastic processes. For example, we have chosen a $1 \rightarrow 3$ adapter for simplicity and symmetry, but any other kind of adapter with the correct consistency properties would work just as well. Additionally, it is somewhat naïve to employ the same adapter at each iteration step, and to obtain more meaningful result, one should potentially renormalize the adapters used for each layer of the construction process. Nonetheless, this procedure – even in this simple form – provides a powerful toolbox for the generation of quantum stochastic processes. This is of particular interest, as the GET, while guaranteeing the *existence* of an underlying process, says nothing about its properties. A construction of the form that we proposed might allow one to make explicit statements about the field theory that corresponds to the limit $n \rightarrow \infty$ in a similar manner as properties of conformal field theories (CFTs) are obtained from an analysis of the corresponding multi-scale entanglement renormalization ansatz (MERA) representation [108–110]. A thorough analysis of this connection would exceed the scope of this thesis, though, and has to be relegated to future work.

3.11 GENERAL STOCHASTIC PROCESSES – SUMMARY

In this chapter, we have examined the foundations of the theory of classical stochastic processes. While the Kolmogorov extension theorem constitutes the foundation for this theory, it does

¹⁰ On the other hand, the ordering of the wires labeled by a, b, d, e, f , and h is fixed and goes from left to right in Fig. 3.9.

not hold in quantum mechanics, or any other theory that allows for active interventions. This breakdown goes hand in hand with the violation of Leggett-Garg inequalities: the violation of such an inequality *always* implies that compatibility conditions are not satisfied, and hence the KET does not hold.

The only escape from this lack of rigorous foundations in any theory of stochastic processes that goes beyond classical physics without interventions is to actively consider experimental influences to obtain a consistent definition and rigorous mathematical foundation for quantum stochastic processes. Put differently, without taking interventions into account, there is no way to consistently define quantum stochastic processes. In this sense, two seemingly different frameworks – the framework of causal modeling, and the theory of quantum stochastic processes – are actually two sides of the same coin.

In hindsight, the route we took to prove the generalized extension theorem seems natural; once one has the framework of quantum combs at one’s disposal and understands the concepts of instruments as the natural way to tie in the peculiarities of quantum mechanics into the theory of stochastic processes, there is no other logical way to tackle quantum stochastic processes. However, without this theoretical machinery, there appears to be no clear resolution to the breakdown of the KET in quantum mechanics [65], which, for example, is mirrored by the proliferation of inequivalent ‘definitions’ of non-Markovianity in the quantum realm (see, e.g., Ref. [66] for a comprehensive review).

The GET can be considered as an ‘umbrella theorem’, that constitutes the foundation for the plethora of equivalent frameworks that are in use to describe quantum processes. Additionally, while the tools we use to describe quantum dynamics provided our guiding principles, the arguments we employed were not inherently quantum mechanical. Due to the linearity of mixing, *any* meaningful description of a stochastic process – quantum or not – must be expressible in terms of a linear function on the space of locally accessible operations [8]. The proof of the GET is versatile enough to account for any framework that aims to describe stochastic processes, and hence provides a sound mathematical underpinning for all of them, contextual or not.

The roadblocks encountered when describing quantum processes in terms of joint probability distributions can be remedied by changing perspective; while the evolution of a density matrix over time does not contain enough statistical information for consistency properties to hold [65], considering a quantum stochastic process as a linear functional acting on sequences of CP maps, allows one to formulate a fully-fledged theory. In the limit of continuous time, the sequence of CP maps becomes a continuous driving/control of the system of interest. Thus, the GET provides the theoretical foundation for these experimental scenarios, which is important for development of quantum technologies. Likewise, just as in the case of classical stochastic processes, the GET provides a toolbox for the modeling of quantum stochastic processes; any mechanism that leads to consistent families of combs automatically defines an underlying process. Here, we have but scratched the surface of the idea of adapters, that provide the versatile means to construct quantum stochastic processes.

Finally, while we have mostly discussed temporally ordered processes, in principle, even causally disordered processes could be described by families of functionals that satisfy a consistency requirement (Λ would then be thought of as a set of labels for different laboratories).

However, there is no deterministic Stinespring dilation for causally disordered processes [53], which makes an intuitive interpretation of an ‘underlying’ causally disordered process somewhat difficult. On the other hand, there are dilations that include post-selection [3, 15, 52, 111], and it is reasonable to conjecture that an underlying causally disordered stochastic process would be equivalent to post-selection on a class of trajectories resulting from continuous weak measurement.

Up to this point, we have discussed both higher-order quantum maps, as well as quantum stochastic processes, from a very axiomatic point of view, by considering the logical structure that has to be employed to describe them and by explicitly deriving the fundamental theorem for their discussion. What we are still lacking is a clear picture of the physical scenarios that lead to the probability distributions that can be investigated experimentally. Asking this question directly leads to the theory of open quantum system dynamics, where one tries to model the dynamics of a system of interest that is coupled to an unknown and uncontrollable environment. We shall now go on to discuss the description of open quantum systems from the vantage point of higher-order quantum maps and the operational understanding of quantum stochastic processes we have developed throughout this chapter. This discussion will shed light on the question what can, in principle, be learnt about open system dynamics, and how can the ideas of memory and memory length be generalized from the classical to the quantum case.

OPEN QUANTUM SYSTEM DYNAMICS

So far, our discussion of physical processes has predominantly taken place on axiomatic grounds. However, this is generally not the vantage point of an experimenter. Rather, they would choose to describe their experiment in terms of degrees of freedom that they can control, those that they cannot control, and an underlying dynamics that leads to the observed measurement statistics. What we are thus still missing is a dynamical mechanism that yields the respective joint probability distributions that we discussed in the previous chapter.

It is now time to ‘look under the hood’ of stochastic processes, and approach them with a distinctly physical, rather than conceptual, motivation. Consequently, here, we shall discuss the connection of the theoretical framework of general stochastic processes and the standard picture of open quantum system dynamics, *i.e.*, the dynamics of a system of interest coupled to an uncontrollable and experimentally inaccessible environment. Unsurprisingly, we will see that we had the correct framework all along: every open quantum dynamics leads to a quantum comb, and vice versa.

Before arriving at this point, we will take a slight detour to make sure that we actually understand its necessity. To do so, we will start by giving a short overview over the traditional description of open quantum systems. On the one hand, this discussion will provide additional arguments for the use of higher-order quantum maps in the treatment of open dynamics, independent of the theory of stochastic processes. For example, while traditional approaches are sufficient to describe memoryless (*i.e.*, Markovian) processes [19, 71, 78, 79], they generally break down as soon as memory effects play a non-negligible role [8, 67, 71]. Additionally, memory effects lead to apparent paradoxical violations of fundamental physical and information-theoretic bounds [28–30, 112]. The switch to higher-order quantum maps – and the accompanying switch to a distinctly operational description – resolves these (and many other) issues by constructing bounds that can properly account for initial correlations and general memory effects [113], and by taking multi-time correlations into consideration [32, 105]. This mirrors the situation of the previous chapter, where the lack of Kolmogorov conditions was also merely a problem of formalism, and could be overcome by changing perspective and introducing higher-order quantum maps as the descriptor of general stochastic processes.

The theory of open quantum systems is a powerful application of the comb formalism. In the logic of our discussion, it will provide intuition for the *experimental/tomographic reconstruction* of quantum combs, a topic we have not touched upon so far in our treatment of their mathematical properties. In this sense, the theory of open quantum systems naturally establishes combs as an experimentally accessible entity that overcomes fundamental problems

in the description of open system dynamics and provides them with a physical reality that goes beyond their theoretical appeal.¹

Introduced in the order we chose, it might appear as if we considered open quantum system dynamics as an interesting field because it allows for the usage of higher-order quantum maps. However, we should understand this story the other way round; any device that aims to harness quantum effects by storing, manipulating or controlling quantum states will inevitably interact in an uncontrollable way with its environment and thus be plagued by noise. Consequently, understanding and characterizing the dynamics of quantum systems that are coupled to their surroundings is a necessary prerequisite for the efficient implementation of new quantum technologies. In particular, with the increase of miniaturization and read-out frequencies in mind, the development of tools that are tailored to characterize open system dynamics with non-negligible memory effects, and quantify said memory effects once the characterization is obtained, is of crucial importance both from a foundational as well as a technological point of view.

Traditionally, the dynamics of open quantum systems are found by solving an equation of motion for the reduced state of the former, and are described by channels that map initial (input) states of the system to final (output) states, which can be reconstructed experimentally. Here, we will predominantly be concerned with the latter aspect of the field of open quantum system dynamics: In the presence of an uncontrollable environment that interacts with the degrees of freedom that we can access, how can we experimentally reconstruct a descriptor of the dynamics? While a breakdown of reconstructability, as we will encounter it in the discussion of traditional approaches to open quantum system dynamics, also implies an inadequacy of the corresponding equations of motion, we will not discuss these *master equations* in detail in this chapter (for an overview, see, e.g., Ref. [71]).

Here, we shall rather pinpoint the problems and shortcomings that traditional approaches are plagued by. The operational way out of the encountered problems is provided by a perspectival readjustment: instead of describing open quantum system dynamics as the time evolution of the state of the system of interest, in the presence of system-environment correlations – the regime where the traditional approach breaks down – one rather has to understand it as a mapping defined on the space of operations that the experimenter can implement [31, 32]. The resulting descriptor can be reconstructed experimentally, independent of the existence of memory effects and/or initial correlations, and allows one to answer natural questions in the field of open quantum system dynamics, like, e.g., the definition and characterization of memory effects.

We shall discuss both of these questions in detail in this chapter. In particular, the investigated framework will be used to analyze existing standard tools that are employed to probe the existence of memory effects in quantum processes, which will lead to the main result of this chapter (based on Ref. [1]): an alternative, operationally clear-cut definition of CP divisibility, and the comprehensive analysis of the temporal correlations that this witness of memory effects is sensitive to. This investigation, in turn, will provide a quantitatively tangible interpretation

¹ Evidently, quantum combs have a clear-cut physical and tomographic meaning independent of the area of open quantum system dynamics. As the latter is the main focus of this thesis, here, we use it as a natural underlying physical picture of quantum networks.

of many of the measures that are in use for the detection of memory effects, as most of them rely on the breakdown of CP divisibility. We round off the discussion of memory effects by providing the generalization of memory length, *i.e.* Markov order, to the quantum case (based on Refs. [5, 6]).

Subsequent to this analysis, in the following chapter, we will bring the developed framework closer to experimental reality, by tailoring it to the situation, where the experimenter has only limited control over the system of interest.

4.1 OPEN QUANTUM SYSTEM DYNAMICS – THE TRADITIONAL APPROACH

In what follows, we will consider an experimenter, who can probe degrees of freedom (the *system of interest*) that are coupled to an environment which is out of their experimental control. Thus, the system of interest is considered ‘open’ as it can interact with external degrees of freedom. For example, such a system could constitute qubits that one wishes to use for computations in a quantum computer by implementing sequences of quantum gates and a final readout, constituting a set of control operations. The interaction with the environment introduces errors to the computation, and their understanding is of utmost importance in order to be able to correct them.

Arguably, from the perspective of this experimenter, the most general, operationally meaningful description of a quantum process is a mapping from experimentally controllable inputs to final output states ρ' (which, in the example above, would contain the result to the computation). Indeed, once the experimenter knows what the state of their system of interest is at the end of their respective experiment, given the experimental parameters they chose, they know everything that can be learnt about the underlying dynamics.

Depending on the experimental setup in question, these controllable inputs could, for example, be initially prepared system states, initial preparation operations, sequences of local operations, or both initial states *and* sequences of local operations. Due to the linearity of quantum mechanics (in the sense that we discussed in Ch. 2), for each of these cases, the map describing the process can be reconstructed experimentally by measuring the final states corresponding to a complete basis of the inputs.

The final state can – in principle – be reconstructed via *quantum state tomography* (QST), *i.e.*, by inferring the probabilities $P_j = \text{tr}(\rho' E_j^T)$ for a set $\{E_j^T\}$ of POVM elements that spans the space ρ' is defined on [19] (such a POVM is called *informationally complete* (IC)).

To see how this abstract notion of open system dynamics plays out in experimental reality, and to introduce some of the key ideas of this chapter, consider the simplest scenario of open quantum system dynamics – leading to the well-known case of quantum channels: Let s be a system of interest that can be experimentally accessed. It is coupled in an uncontrollable way to its environment e . At a time t_1 the system is uncorrelated with e , which is in a fixed state $\eta_e \in \mathcal{B}(\mathcal{H}_e^1)$. Together, s and e are considered closed. Consequently, they undergo a unitary evolution $\mathcal{U} : \mathcal{B}(\mathcal{H}_s^1 \otimes \mathcal{H}_e^1) \rightarrow \mathcal{B}(\mathcal{H}_s^0 \otimes \mathcal{H}_e^0)$ for, say, a time Δt , that is generated

by the system-environment Hamiltonian H_{se} and can – in principle – be derived from the Schrödinger equation²:

$$\mathcal{U}[\rho_{se}] = \mathcal{U}\rho_{se}\mathcal{U}^\dagger = e^{-iH_{se}\Delta t}\rho_{se}e^{iH_{se}\Delta t}. \quad (4.1)$$

As system and environment are uncorrelated at t_1 , the experimenter can freely prepare system states $\rho_s^{(j)} \in \mathcal{B}(\mathcal{H}_s^i)$, and reconstruct (*i.e.*, perform QST) the system state $\rho_s^{(j)'} \in \mathcal{B}(\mathcal{H}_s^o)$ at $t_2 = t_1 + \Delta t$. The output state $\rho_s^{(j)'}$ corresponding to the input state $\rho_s^{(j)}$ is given by (see Fig. 4.1)

$$\rho_s^{(j)'} = \text{tr}_e \left[\mathcal{U} \left(\rho_s^{(j)} \otimes \eta_e \right) \mathcal{U}^\dagger \right], \quad (4.2)$$

where the final trace over the degrees of freedom of the environment reflects the fact that at t_2 , the experimenter only measures the system and disregards the environment. Expressed in terms of link products, Eq. (4.2) reads

$$\rho_s^{(j)'} = \mathbb{1}_e^o \star \mathcal{U} \star \rho_s^{(j)} \star \eta_e, \quad (4.3)$$

where, following our notational convention, \mathcal{U} is the Choi state of \mathcal{U} . Compressing all the elements of the description that the experimenter has no control over (again, see Fig. 4.1), Eq. (4.2) can be rewritten as

$$\rho_s^{(j)'} = \mathcal{L}[\rho_s^{(j)}], \quad (4.4)$$

where $\mathcal{L} : \mathcal{B}(\mathcal{H}_s^i) \rightarrow \mathcal{B}(\mathcal{H}_s^o)$ is the linear map that maps all initial states $\rho_s^{(j)}$ to the correct corresponding output states $\rho_s^{(j)'}$. From Eq. (4.3) we can ‘read off’ the Choi state of \mathcal{L} :

$$\mathcal{L} = \mathbb{1}_e^o \star \mathcal{U} \star \eta_e, \quad (4.5)$$

where $\mathbb{1}_e^o$ corresponds to the final trace over the environment. This implies that \mathcal{L} is CPTP, as all the elements in the link product of Eq. (4.5) are CPTP. Naturally, it can also be shown directly from Eq. (4.2) that \mathcal{L} is CPTP [21, 114].

While in Ch. 2 we introduced the idea of complete positivity and trace preservation from axiomatic considerations, here, we see that we obtain this property directly from the assumption of unitarity of the *se* dynamics, initial independence of the system and its environment, and the fact that the state of the environment is finally traced over.

Due to linearity, \mathcal{L} is unambiguously defined by its action on a basis $\{\rho_s^{(j)}\}_{j=1}^{d_s^2}$ of $\mathcal{B}(\mathcal{H}_s^i)$, where d_s is the dimension of \mathcal{H}_s^i ; the number of d_s^2 basis elements stems from the fact that the quantum states on \mathcal{H}_s^i span the space of $d_s \times d_s$ Hermitian matrices, which is d_s^2 -dimensional. Every input state³ $\rho \in \mathcal{B}(\mathcal{H}_s^i)$ can be decomposed as a real linear combination $\rho = \sum_{j=1} r_j \rho^{(j)}$, and hence the action of \mathcal{L} on ρ is given by

$$\mathcal{L}[\rho] = \sum_{j=1} r_j \mathcal{L}[\rho^{(j)}] = \sum_{j=1} r_j \rho^{(j)'}. \quad (4.6)$$

² Throughout this thesis, we will consider all Hamiltonians to be time-independent. The generalization to the time-dependent case can either be achieved straight forwardly, or removed argumentatively, by considering the environment large enough.

³ For compactness, from now on, we will often drop the subscript *s* and reserve the symbol ρ exclusively for system states.

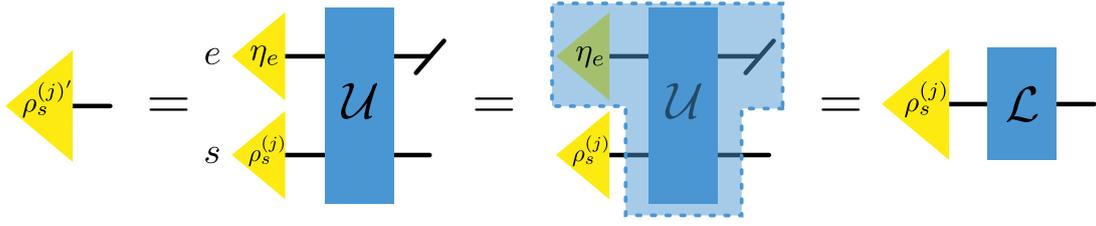


Figure 4.1: Quantum channel resulting from open dynamics. Initially (at t_1), system and environment are uncorrelated. They undergo unitary evolution together, and finally, at time t_2 , the environment is discarded. As is convention in open quantum system dynamics, the trace is denoted by a slash instead of the effect $\mathbb{1}$. The channel \mathcal{L} represents ‘everything that is outside of experimental control’. Drawing boxes around the uncontrollable parts of the dynamics will prove to be a powerful graphical tool throughout this chapter.

Once the output states $\rho^{(j)'}$ for a basis of input states are known, the map describing the underlying process is entirely defined. This fact forms the basis of quantum process tomography (see, e.g., Ref. [79, 93]) where the map \mathcal{L} is reconstructed experimentally by determining the output states $\rho^{(j)'} = \mathcal{L}[\rho^{(j)}]$ for a set of d_s^2 linearly independent input states and employing linear inversion techniques.

In detail, let $\{\Delta_k\}_{k=1}^{d_s^2}$ be the *dual set* [68] to the basis $\{\rho_s^{(j)}\}_{j=1}^{d_s^2}$, i.e., $\text{tr}(\rho^{(j)} \Delta_k^\dagger) = \delta_{jk}$ (for comprehensiveness, an explicit construction of $\{\Delta_k\}_{k=1}^{d_s^2}$ for any given basis is provided in App. B.1). With this, \mathcal{L} can be written as [8, 31, 32]

$$\mathcal{L} = \sum_{j=1}^{d_s^2} \rho^{(j)'} \otimes \Delta_j^*. \quad (4.7)$$

Indeed, by insertion, we see that the matrix $\mathcal{L} \in \mathcal{B}(\mathcal{H}_s^0 \otimes \mathcal{H}_s^1)$ defined in this way yields the correct output state for any input state $\rho = \sum_{j=1}^{d_s^2} r_j \rho^{(j)}$:

$$\mathcal{L} \star \rho = \text{tr}_i \left[\left(\mathbb{1}_s^0 \otimes \rho^T \right) \mathcal{L} \right] = \sum_{j,k=1}^{d_s^2} r_j \text{tr} \left(\Delta_k^\dagger \rho^{(j)} \right) \rho^{(k)'} = \sum_{j=1}^{d_s^2} r_j \rho^{(j)'} = \mathcal{L}[\rho], \quad (4.8)$$

where we have used the linearity of \mathcal{L} . Importantly, the quantum channel \mathcal{L} can be experimentally reconstructed without any knowledge of the unitary system-environment dynamics, and, once reconstructed, allows one to predict the correct output state for any input state.

Reconstructing \mathcal{L} for different times Δt , one obtains a family of maps $\{\mathcal{L}_{\Delta t}\}$ that correctly describe the dynamics from t_1 to all times $t_1 + \Delta t$, and there exists a wealth of master equation approaches to derive $\mathcal{L}_{\Delta t}$ for various physical scenarios [71]. Most notably among them the Gorini-Kossakowski-Sudarshan-Lindblad (**GKSL**)⁴ equation for the case of Markovian dynamics [62, 115, 116], and the Nakajima-Zwanzig equation that can account for memory effects [117, 118]. As we will predominantly be concerned with the reconstruction and the mathematical properties of open processes, and not with their derivation from microscopic or phenomenological models, we will not discuss these methods in this thesis. Additionally, the aforementioned master equation approaches do not, in general, allow one to derive multi-time correlation functions, which is one of the main foci of our treatment of open system dynamics.

⁴ This equation has also been independently discovered by Franke in 1976 [115]. So far, his name has not made it into the famous acronym, though.

Returning to the reconstruction of \mathcal{L} , the experimental procedure outlined above relies fundamentally on the fact that the system and the environment are initially in a product state. This, in turn, implies that the system state can be varied without affecting the state of the environment, and can thus be considered a controllable input to the process. As soon as initial system-environment correlations are present, this assumption does not hold, and a reconstruction of the dynamics in the way described above is not possible anymore. Consequently, when experimenters began reconstructing quantum gates – the fundamental elements of a quantum computer – in the late 1990s and early 2000s, they routinely did not obtain completely positive maps [119–121] for various experimental reasons, but predominantly due to initial correlations between the system of interest and its environment [112, 122].

To make matters worse, as most clearly elucidated by Pechukas in his seminal paper [25] and in a subsequent exchange between him and Alicki [27, 123] – and later generalized in [124, 125] – a linear map from input to output states is CP iff there are *no* initial *se* correlations, which seems to imply that in the presence of initial correlations, either complete positivity or linearity have to give. Both from an experimental, as well as a theoretical standpoint, this state of affairs is problematic. On the one hand, complete positivity is a useful property – giving up complete positivity means giving up the Holevo quantity [28], data processing inequality [29], and entropy production inequality [30] – and a CP description naturally predicts the physical fact that one always reconstructs positive probabilities (even for correlated preparations).

On the other hand, dropping linearity is not a viable option either: complete tomography is not possible when the dynamics is nonlinear – at least not in a finite number of experiments. Additionally, a breakdown of linearity would challenge the probabilistic structure of quantum mechanics.

Faced with this choice, many researchers have opted to relinquish complete positivity of dynamics in favor of a framework for open dynamics based on *not completely positive* (NCP) maps [124, 126]. In brief, NCP maps are linear maps that preserve positivity for some subset of the space of system density operators, but fail to do so on the remaining set. While mathematically well-defined (though not unique), the NCP framework lacks a clear link to the operational reality of quantum dynamics, and we will not consider them in detail (see [8] for a more in-depth discussion), as it proves unnecessary to give up either complete positivity or linearity [31]. We shall demonstrate this, using the theory of higher-order quantum maps, in the following section.

Importantly, given that system-environment correlations can be considered as a memory of past interactions, the breakdown of formalism we just alluded to implies that the traditional approach to the description of open quantum system dynamics cannot properly account for memory effects, *i.e.*, it does not work in the non-Markovian regime. Additionally, even if it could be employed to properly describe situations with initial correlations (which it cannot), the resulting maps \mathcal{L} would only account for two-point correlations. For example, with \mathcal{L} at hand, one could predict the probabilities of measurement outcomes at $t_1 + \Delta t$, given that the state ρ was prepared at t_1 . However, it does not allow one to predict the joint probabilities for measurements at, say, $t_1 + \frac{\Delta t}{2}$ and $t_1 + \Delta t$, and this situation cannot generally be accounted for, if both $\mathcal{L}_{\frac{\Delta t}{2}}$ and $\mathcal{L}_{\Delta t}$ are known (see, *e.g.*, [127] and Sec. 4.8.4).

Quantum channels – or families thereof – are, by design, not sufficient to describe multi-time quantum stochastic processes of the kind we introduced in the previous chapter.⁵ Unsurprisingly then, attempts to model memory effects in quantum mechanics by investigating the dynamics of the state of the system of interest, have yielded a ‘zoo’ of non-equivalent incompatible measures and witnesses of non-Markovianity, that lack a clear operational motivation [129–143] (for a recent review of the treatment of memory effects in quantum mechanics that goes beyond these witnesses and provides a comprehensive hierarchy of notions of non-Markovianity, see [66]).

Nonetheless, the reconstruction procedure outlined above points into the right direction: an experiment is fully described, once the outputs for a full basis of inputs can be predicted. All of the aforementioned problems can be overcome by considering the correct objects to be the inputs of the dynamics in the presence of initial correlations or memory effects. This distinctly operational approach naturally leads to a description of open quantum system dynamics in terms of higher order quantum maps.

4.2 INITIAL CORRELATIONS AND SUPERCHANNELS

Above, we considered open quantum system dynamics as a mapping from inputs to outputs, where the inputs correspond to what can be freely prepared by the experimenter, without changing uncontrollable external parameters, and outputs to what the experimenter can measure at the end of the experiment. If system and environment are initially uncorrelated, the inputs of the dynamics are initial states of the system, as they can be prepared by the experimenter without affecting the environment. However, this is not generally true.

Generally, an experimenter would prepare an initial state by applying a control operation $\mathcal{M}^{(j)} : \mathcal{B}(\mathcal{H}_{s_1}^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_1}^i)$, that leaves the system in a known state $\rho^{(j)}$. In anticipation of our ensuing discussion of the multi-time scenario, here, we employ the convention that $\mathcal{H}_{s_\alpha}^x$ corresponds to the system Hilbert space \mathcal{H}_s^x at time t_α .⁶ These control operations can be anything that is admissible in quantum mechanics, including unitary transformations, projective measurements, projective measurements followed by a unitary transformation and everything in between [144].

Now, if the initial system-environment state ρ_{se} is correlated, the control operation $\mathcal{M}^{(j)}$ will inevitably change the state of the environment. The prepared system states cannot be considered as the input of the process anymore, as the experimenter cannot create them without changing the process. As alluded to above, in this case the (fictitious) mapping from input states to output states would be non-linear, as this dynamics would depend on the state that was prepared. As such, it would not be a meaningful description of the process. On the other hand, the experimenter has full control – independent of uncontrollable parameters – over the *operations* they use to manipulate the initial state, which implies that the control operations,

⁵ This statement should not be confused with the collections of CPTP maps that have recently been used to fully describe open quantum system dynamics in [7, 128] in the presence of initial *se* correlations.

⁶ Somewhat counterintuitively, $\mathcal{M}^{(j)}$ maps from a space labeled by *o* to one labeled by *i*. As in the previous chapter, this peculiarity stems from the fact that we abide by the convention to denote spaces from the perspective of the process. As outputs of $\mathcal{M}^{(j)}$ are inputs of the process (see, e.g., Fig. 4.2), the labeling has to be chosen in this way for consistency.

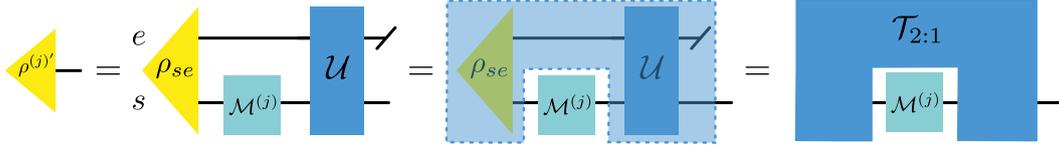


Figure 4.2: Resulting superchannel from an initially correlated system-environment state and open quantum system dynamics. The control operation $\mathcal{M}^{(j)}$ is implemented at time t_1 , and the final state is measured at time t_2 . The superchannel $\mathcal{T}_{2:1}$ contains all parts of the dynamics that cannot be controlled by the experimenter, and acts on all the inputs that they can prepare independently of the dynamics.

rather than the initial state of the system should be considered the input of the process when initial correlations play a non-negligible role. Importantly, in addition to being the correct description, this switch of perspective recovers both linearity *and* complete positivity.

In detail, for a control operation $\mathcal{M}^{(j)}$, the corresponding resulting output state $\rho^{(j)'}$ at the conclusion of the experiment is computed via (see Fig. 4.2)

$$\rho^{(j)' } = \text{tr}_e \left(\mathcal{U} \left[(\mathcal{M}^{(j)} \otimes \mathcal{I}_e) [\rho_{se}] \right] \right), \quad (4.9)$$

or, expressed in terms of link products,

$$\rho^{(j)' } = \mathbb{1}_e^o \star \mathcal{U} \star \mathcal{M}^{(j)} \star \rho_{se} = \mathcal{M}^{(j)} \star (\mathbb{1}_e^o \star \mathcal{U} \star \rho_{se}) := \mathcal{M}^{(j)} \star \mathcal{T}_{2:1}. \quad (4.10)$$

Eq. (4.10) tells us that the output state $\rho^{(j)'}$ corresponding to a control operation $\mathcal{M}^{(j)}$ can be computed as the result of a linear map $\mathcal{T}_{2:1}$ acting on $\mathcal{M}^{(j)}$, *i.e.*, $\rho^{(j)' } = \mathcal{T}_{2:1}[\mathcal{M}^{(j)}]$. As all the elements it is made up of are CPTP, $\mathcal{T}_{2:1} = \mathbb{1}_e^o \star \mathcal{U} \star \rho_{se}$ is also CP and trace preserving in a well-defined sense (see below). $\mathcal{T}_{2:1}$ is called a *superchannel* [31], and once reconstructed, it allows one to predict the correct output state for any experimentally realizable input, *i.e.*, for any control operation $\mathcal{M}^{(j)}$ that the experimenter can perform at the beginning t_1 of the dynamics. The subscripts of $\mathcal{T}_{2:1}$ are chosen in anticipation of its generalization to several times, and signify that the superchannel describes the dynamics from t_1 to t_2 .

As for the case of channels, due to linearity, the superchannel $\mathcal{T}_{2:1}$ can be reconstructed by measuring the outputs (states $\rho^{(j)'}$) corresponding to a basis of inputs (operations $\mathcal{M}^{(j)}$), without any knowledge of the underlying system-environment dynamics \mathcal{U} , or the initial *se* state ρ_{se} . In detail, we have

$$\mathcal{T}_{2:1} = \sum_{j=1}^{d_s^4} \rho^{(j)' } \otimes \mathfrak{m}^{(j)*}, \quad (4.11)$$

where $\{\mathfrak{m}^{(j)}\}_{j=1}^{d_s^4}$ is the set of duals to the basis $\{\mathcal{M}^{(j)}\}_{j=1}^{d_s^4}$, *i.e.*, $\text{tr}(\mathfrak{m}^{(\ell)\dagger} \mathcal{M}^{(j)}) = \delta_{j\ell}$. The number of basis elements derives itself from the fact that the matrices $\mathcal{M}^{(j)}$ span the space $\mathcal{B}(\mathcal{H}_1^1 \otimes \mathcal{H}_1^0)$, which is d_s^4 dimensional. The superchannel $\mathcal{T}_{2:1} \in \mathcal{B}(\mathcal{H}_{s_2}^o \otimes \mathcal{H}_{s_1}^1 \otimes \mathcal{H}_{s_1}^0)$ is a $d_s^3 \times d_s^3$ matrix, that, in accordance with its causality conditions, satisfies

$$\mathcal{T}_{2:1} \star \mathbb{1}_{s_2}^o = \mathbb{1}_{s_1}^1 \otimes \rho_s, \quad (4.12)$$

where $\rho_s = \text{tr}_e(\rho_{se})$. Consequently, $\mathcal{T}_{2:1}$ is a special case of a deterministic one-slot quantum comb.

The fact that Eq. (4.11) indeed yields the correct superchannel can be seen directly by insertion. When applied to an arbitrary CP map $\mathcal{N} = \sum_{\ell} n_{\ell} \mathcal{M}^{(\ell)}$, we obtain

$$\begin{aligned} \mathbb{T}_{2:1} \star \mathcal{N} &= \sum_{\ell} n_{\ell} \text{tr}_{1 \circ 1^1} \left[\left(\mathbb{1}_{s_2}^{\circ} \otimes \mathbf{M}^{(\ell)\text{T}} \right) \mathbb{T}_{2:1} \right] \\ &= \sum_{\ell} n_{\ell} \rho^{(\ell)'} = \sum_{\ell} n_{\ell} \mathcal{T}_{2:1}[\mathcal{M}^{(\ell)}] = \mathcal{T}_{2:1}[\mathcal{N}], \end{aligned} \quad (4.13)$$

where we have used the linearity of $\mathcal{T}_{2:1}$. As $\mathcal{T}_{2:1}$ maps all conceivable inputs to the correct corresponding outputs, it contains all information about the process that can be inferred from operations on the system alone, and thus plays the same role as quantum channels for the case of an uncorrelated initial se state. Importantly, $\mathbb{T}_{2:1}$ contains this special case; setting $\rho_{se} = \rho_s \otimes \eta_e$, we obtain

$$\mathbb{T}_{2:1} = \mathbb{1}_e^{\circ} \star \mathbf{U} \star \rho_{se} = \mathbb{1}_e^{\circ} \star \mathbf{U} \star \rho_s \star \eta_e = \mathbf{L} \otimes \rho_s, \quad (4.14)$$

where $\mathbf{L} = \mathbb{1}_e^{\circ} \star \mathbf{U} \star \eta_e$ is a quantum channel of the form we discussed in the previous section, and Eq. (4.10) yields [31]

$$\rho^{(j)'} = \mathbf{M}^{(j)} \star \mathbf{L} \star \rho_s = (\mathbf{M}^{(j)} \star \rho_s) \star \mathbf{L} = \mathcal{L}[\mathcal{M}^{(j)}[\rho_s]]. \quad (4.15)$$

When there are no initial correlations – or rather, no correlations that play a role for the open dynamics – then the superchannel is of product form, and vice versa. When we analyze memoryless processes, we will see, that this is a special case of the general structural properties of Markovian dynamics.

We can see from Eq. (4.12) that the superchannel $\mathbb{T}_{2:1}$ allows one to deduce the initial state ρ_s of the system (even in the presence of initial se correlations). This fact can also be seen on more intuitive grounds. The reconstruction of $\mathbb{T}_{2:1}$ necessitates the implementation of d_s^4 linearly independent CP maps. Not all of these maps can be trace preserving; due to their trace constraint, there are only $d^4 - d^2$ linearly independent CPTP maps that act on $\mathcal{B}(\mathcal{H}_s)$ which means that the remaining linearly independent basis elements are d_s^2 non trace-increasing CP maps. This set of CP maps must be IC, and as such, recording the probabilities of occurrence for each of them amounts to full tomography of the initial system state ρ_s . In turn, $\mathbb{T}_{2:1}$ contains all information about ρ_s . Importantly, though, the actual reconstruction of $\mathbb{T}_{2:1}$ does not have to be carried out in the above way, *i.e.*, with $d_s^4 - d_s^2$ CPTP maps and d_s^2 CP maps. Any set of d_s^4 linearly independent maps will yield exactly the same superchannel $\mathbb{T}_{2:1}$, but the reconstruction *always* necessitates at least d_s^2 maps that are *not* trace preserving. Recently, these theoretical insights have been used to experimentally reconstruct superchannels in the laboratory [145].

4.3 SUPERCHANNELS, LINEARITY, COMPLETE POSITIVITY, AND TRACE PRESERVATION

Switching perspective and considering implementable operations, instead of initial states as the inputs of the dynamics remedies the problems encountered when initial system-environment correlations play a role. On the mathematical side, it yields a descriptor $\mathbb{T}_{2:1}$ of the dynamics, that is linear, CP and TP. It is worth commenting on each of these properties and their meaning for superchannels to build intuition for the subsequent sections.

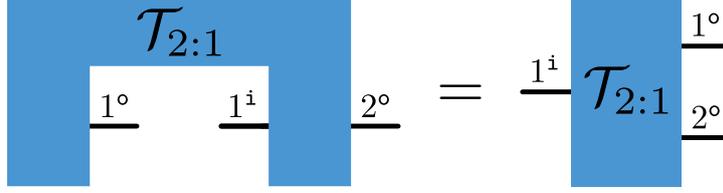


Figure 4.3: Different ways of considering $\mathcal{T}_{2:1}$, as a map $\mathcal{T}_{2:1} : \mathcal{B}(\mathcal{H}_{s_1}^i \otimes \mathcal{H}_{s_1}^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_2}^o)$ (left) and a map $\mathcal{T}_{2:1} : \mathcal{B}(\mathcal{H}_{s_1}^i) \rightarrow \mathcal{B}(\mathcal{H}_{s_1}^o \otimes \mathcal{H}_{s_2}^o)$ (right). Both complete positivity as well as trace preservation can be understood in both cases, however the former case is better adopted to open system dynamics.

We can see easily, that without a switch of perspective, linearity cannot generally be upheld (see also [8, 146] for a more detailed discussion of this point); for example, consider control operations that prepare initial states by performing projective measurements, *i.e.*, $\mathcal{M}^{(j)}[\rho_s] = \langle j | \rho_s | j \rangle | j \rangle \langle j |$. If the initial system-environment state ρ_{se} is correlated, then the resulting state after the action of the control operation is given by

$$\mathbb{M}^{(j)} \star \rho_{se} = |j\rangle \langle j| \otimes \langle j | \rho_{se} | j \rangle, \quad (4.16)$$

and the corresponding final system state $\rho_s^{(j)'}$ reads

$$\mathbb{1}_e^o \star \mathbb{U} \star \mathbb{M}^{(j)} \star \rho_{se} = |j\rangle \langle j| \star (\mathbb{1}_e^o \star \mathbb{U} \star \langle j | \rho_{se} | j \rangle) := |j\rangle \langle j| \star \mathbb{L}^{(j)} = \mathcal{L}^{(j)}[|j\rangle \langle j|]. \quad (4.17)$$

As the map $\mathcal{L}^{(j)}$ generally depends on j , the above equation would correspond to a non-linear dynamics. Importantly, if ρ_{se} is of product form, then $\mathbb{L}^{(j)} = \langle j | \rho_s | j \rangle (\mathbb{1}_e^o \star \mathbb{U} \star \eta_s)$ is – after normalization by the success probability $\langle j | \rho_s | j \rangle$ – the same for every projective control operation, and linearity is recovered. The superchannel is not plagued by an apparent non-linearity, independent of whether or not there are initial correlations, as it considers the correct objects to be the input of the dynamics.

Complete positivity of $\mathcal{T}_{2:1}$ is somewhat less intuitive. As we discussed in Ch. 2, by reordering the wires, we can regard $\mathcal{T}_{2:1}$ as a mapping $\mathcal{T}_{2:1} : \mathcal{B}(\mathcal{H}_{s_1}^i) \rightarrow \mathcal{B}(\mathcal{H}_{s_1}^o \otimes \mathcal{H}_{s_2}^o)$ (see Fig.4.3 for an illustration). Considered in this way, complete positivity means⁷ that $(\mathcal{T}_{2:1} \otimes \mathcal{I}_a)[\rho_{s_1 a}^i] \geq 0$ for any positive $\rho_{s_1 a}^i \in \mathcal{B}(\mathcal{H}_{s_1}^i \otimes \mathcal{H}_a)$ and any size of the ancilla a .

However, this understanding of complete positivity is neither very illuminating, nor does it correspond to how we think about the superchannel in the context of open quantum system dynamics. Here, $\mathcal{T}_{2:1}$ is considered a mapping from CP maps to final output states. Regarded in this way, complete positivity (*i.e.*, positivity of $\mathbb{T}_{2:1}$) means that when $\mathcal{T}_{2:1}$ acts on the system part of a CP map $\widetilde{\mathcal{M}}^{(j)} : \mathcal{B}(\mathcal{H}_{s_1}^o \otimes \mathcal{H}_a^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_1}^i \otimes \mathcal{H}_a^i)$, then the resulting map $(\mathcal{T}_{2:1} \otimes \mathcal{I})[\widetilde{\mathcal{M}}^{(j)}] = \widetilde{\mathcal{M}}^{(j)'}$ is a map $\widetilde{\mathcal{M}}^{(j)'}$: $\mathcal{B}(\mathcal{H}_a^o) \rightarrow \mathcal{B}(\mathcal{H}_a^i \otimes \mathcal{H}_{s_2}^o)$ that is completely positive (see Fig. 4.4). It is straightforward to see that $\mathcal{T}_{2:1}$ is indeed CP in this sense. We have

$$\widetilde{\mathbb{M}}^{(j)'} = \mathbb{T}_{2:1} \star \widetilde{\mathbb{M}}^{(j)}, \quad (4.18)$$

and we already know that, if $\mathbb{T}_{2:1}$ and $\widetilde{\mathbb{M}}^{(j)}$ are positive, then their link product $\widetilde{\mathbb{M}}^{(j)'}$ is also positive, *i.e.*, it corresponds to a completely positive map.

⁷ Strictly speaking, we would have to introduce a new symbol for this mapping, as $\mathcal{T}_{2:1}$ was introduced as a map from CP maps to outputs. However, this differentiation is not crucial for the discussion of complete positivity, and we use the same letter to keep notational overhead to a minimum.

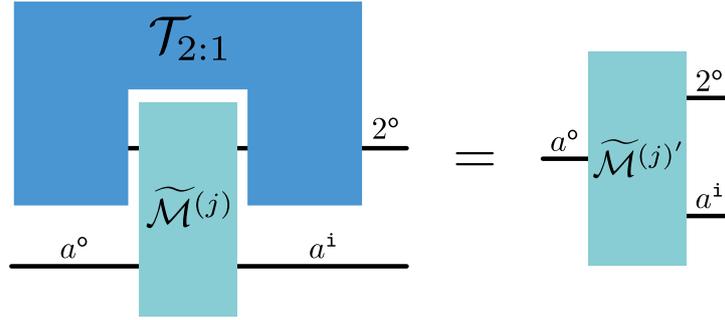


Figure 4.4: Complete positivity of $\mathcal{T}_{2:1}$. When acting on a part of a CP map $\widetilde{\mathcal{M}}_\alpha$, the resulting map $\widetilde{\mathcal{M}}'_\alpha$ is also CP, independent of the size of the involved ancilla.

Analogously, trace preservation of $\mathcal{T}_{2:1}$ can also be considered in two different ways. On the one hand, when considered as a mapping $\mathcal{T}_{2:1} : \mathcal{B}(\mathcal{H}_{s_1}^i) \rightarrow \mathcal{B}(\mathcal{H}_{s_1}^o \otimes \mathcal{H}_{s_2}^o)$ (see Fig. 4.3), we have

$$\mathrm{tr}(\mathcal{T}_{2:1}[\rho_{s_1}^i]) = \mathrm{tr}(\rho_{s_1}^i), \quad (4.19)$$

for all $\rho_{s_1}^i \in \mathcal{B}(\mathcal{H}_{s_1}^i)$. However, this way of considering trace preservation is, again, not well-adapted to the description of open quantum system dynamics as we introduced it, and it also ignores the causal structure of $\mathcal{T}_{2:1}$. The better way to think about ‘trace preservation’ is the following: whenever $\mathcal{T}_{2:1}$ acts on a trace preserving map \mathcal{M} , the resulting output state is of unit trace. Indeed, using Eq. (4.12), we obtain

$$\mathrm{tr}(\mathcal{T}_{2:1}[\mathcal{M}]) = \mathcal{T}_{2:1} \star \mathcal{M} \star \mathbb{1}_{s_2}^o = \rho_s \star \mathbb{1}_{s_1}^i \star \mathcal{M} \star \mathbb{1}_{s_2}^o = \rho_s \star \mathbb{1}_{s_1}^o = \mathrm{tr}(\rho_s) = 1, \quad (4.20)$$

where we have used the fact that \mathcal{M} is trace preserving, *i.e.*, $\mathbb{1}_{s_1}^i \star \mathcal{M} = \mathbb{1}_{s_1}^o$. Finally, when acting on a (trace non-increasing) CP map $\mathcal{M}^{(j)}$, the trace of the output state corresponds to the implementation probability of $\mathcal{M}^{(j)}$, *i.e.*, $\mathbb{P}(j|\mathcal{J}_1) = \mathrm{tr}(\mathcal{T}_{2:1}[\mathcal{M}^{(j)}])$.

Before we continue to generalize the introduced concepts to the multi-step case, it is worth pausing for a moment and to take stock of the resolution of the problems of the description of open quantum system dynamics that we encountered so far in this chapter. A meaningful descriptor of open quantum system dynamics should be experimentally reconstructible and map all preparable inputs to the correct outputs. Both of these requirements are satisfied by the superchannel $\mathcal{T}_{2:1}$. Additionally, it is CP and – in a particular sense – trace preserving, crucial properties that can be (and have been) used to resolve apparent paradoxes in quantum information theory [113].

The superchannel is obtained by switching perspective, regarding what is considered the correct input to the process. This switch directly leads to a description in terms of higher order quantum maps – in this case, a one-slot comb⁸ – with a final output wire. It contains all information about the process, that can be probed by operations on the system only, and can be reconstructed in a finite number of experiments. In particular, it allows one to deduce if there are initial correlations, *i.e.*, if there are memory effects that play a non-negligible role for the dynamics.

⁸ To make the relation to open system dynamics clear, we will continue to label the corresponding higher-order maps by \mathcal{T} instead of \mathcal{C} .

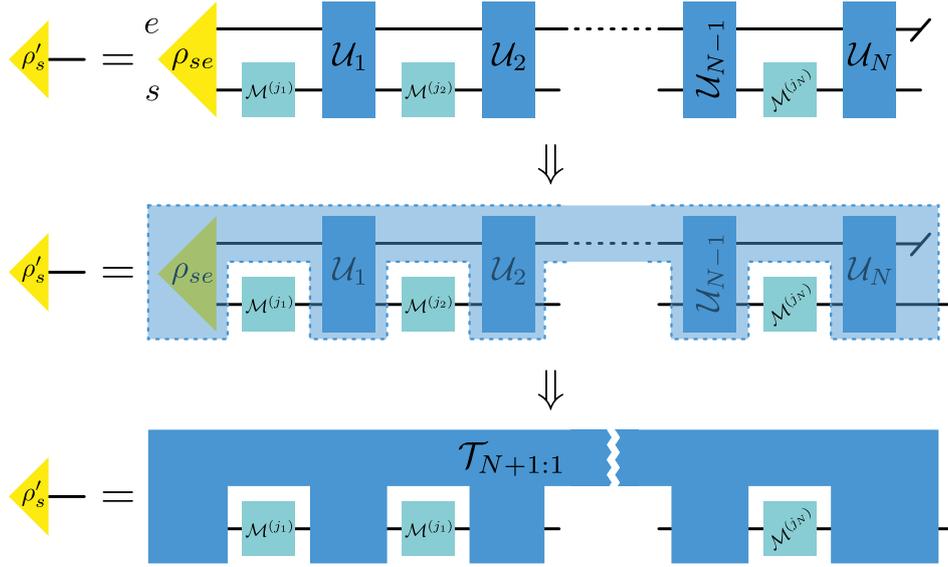


Figure 4.5: Illustration of a multi-time experiment. The experimenter performs operations $\mathcal{M}^{(j_\alpha)}$ at times $t_\alpha \in \{t_1, \dots, t_N\}$, and the resulting system state ρ'_s at t_{N+1} is obtained via the action of the process tensor $\mathcal{T}_{N+1:1}$ on the sequence of operations. Notably, the process tensor contains all the parts of the dynamics that cannot be accessed experimentally.

Finally, even though we phrased the results of this chapter mainly for the case where the input and output spaces of all involved control operations are isomorphic and coincide with the space the final state ρ'_s lives on, the generalization to distinct input and output spaces is straightforward. Now, it is time to extend our formalism to the multi-step case.

4.4 MULTI-STEP PROCESSES AND PROCESS TENSORS

So far, we have only considered experimental situations, where the experimenter performs an operation on the system at the beginning t_1 of the experiment and measures the corresponding outcome at t_2 . More generally, an experimenter could choose to perform operations at times t_1, \dots, t_N , and finally measure the corresponding output at t_{N+1} . For example, they could choose to do so, in order to determine the memory structure of the process [5, 6, 32, 105, 147, 148], or to steer the system of interest to a desired output state at t_{N+1} [149–152]. These control and probing operations could, *e.g.*, be unitary operations or sequential measurements; in the most general setting, any (trace non-increasing) CP maps. We will assume that they are implemented on a time scale much smaller than those typical of the system dynamics.

Following the reasoning of the previous section, the final state ρ'_s at t_{N+1} can be computed via

$$\rho'_s = \text{tr}_{e_{N+1}} \left(\mathcal{U}_N \left[(\mathcal{M}^{(j_N)} \otimes \mathcal{I}_e) \left[\dots \mathcal{U}_1 \left[(\mathcal{M}^{(j_1)} \otimes \mathcal{I}_e) [\rho_{se}] \right] \dots \right] \right] \right), \quad (4.21)$$

where $\mathcal{U}_\alpha : \mathcal{B}(\mathcal{H}_{s_\alpha}^i \otimes \mathcal{H}_{e_\alpha}^i) \rightarrow \mathcal{B}(\mathcal{H}_{s_{\alpha+1}}^o \otimes \mathcal{H}_{e_{\alpha+1}}^o)$ is the system-environment dynamics from t_α to $t_{\alpha+1}$, $\mathcal{M}^{(j_\alpha)} : \mathcal{B}(\mathcal{H}_{s_\alpha}^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_\alpha}^i)$ is the operation on the system that was performed at t_α , and $\text{tr}_{e_{N+1}}$ denotes the final trace over the environment at time t_{N+1} (see Fig. 4.5). Importantly, the unitary evolutions between times can differ. In terms of link products, Eq. (4.21) reads

$$\begin{aligned}
 \rho'_s &= \mathbb{1}_{e_{N+1}}^o \star \mathbf{U}_N \star \mathbf{M}^{(j_N)} \star \cdots \star \mathbf{U}_1 \star \mathbf{M}^{(j_1)} \star \rho_{se} \\
 &= (\mathbf{M}^{(j_N)} \otimes \cdots \otimes \mathbf{M}^{(j_1)}) \otimes (\mathbb{1}_{e_{N+1}}^o \star \mathbf{U}_N \star \cdots \star \mathbf{U}_1 \star \rho_{se}) \\
 &=: (\mathbf{M}^{(j_N)} \otimes \cdots \otimes \mathbf{M}^{(j_1)}) \star \mathbb{T}_{N+1:1},
 \end{aligned} \tag{4.22}$$

where we have defined the *process tensor*, $\mathbb{T}_{N+1:1} \in \mathcal{B}(\mathcal{H}_{s_{N+1}}^o \otimes \mathcal{H}_{s_N}^i \otimes \cdots \otimes \mathcal{H}_{s_1}^i \otimes \mathcal{H}_{s_1}^o)$, which – in clear analogy to the superchannel defined in the previous section – maps every sequence $\mathcal{M}^{(\vec{j})} := (\mathcal{M}^{(j_N)}, \dots, \mathcal{M}^{(j_1)})$ of control operations to the correct corresponding output state

$$\rho^{(\vec{j})'} = \mathbb{T}_{N+1:1}[\mathcal{M}^{(j_N)}, \dots, \mathcal{M}^{(j_1)}]. \tag{4.23}$$

The process tensor was introduced in [32] to describe open quantum system dynamics with memory. From Eq. (4.22), we see directly, that $\mathbb{T}_{N+1:1}$ is a higher order quantum map, which naturally satisfies the causality constraints of deterministic combs. Additionally, it is linear, CP, and trace preserving, where the interpretation of these properties is as in the previous section:

Complete positivity implies that, if the process tensor acts on a sequence of operations $(\widetilde{\mathcal{M}}^{(j_N)}, \dots, \widetilde{\mathcal{M}}^{(j_1)})$, where each of the operations $\widetilde{\mathcal{M}}^{(j_\alpha)} : \mathcal{B}(\mathcal{H}_{s_\alpha}^o \otimes \mathcal{H}_{a_\alpha}^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_\alpha}^i \otimes \mathcal{H}_{a_\alpha}^i)$ acts on the system and an ancilla, then the resulting

$$\mathcal{N} = (\mathbb{T}_{N+1:1} \otimes \mathcal{I})[\widetilde{\mathcal{M}}^{(j_N)}, \dots, \widetilde{\mathcal{M}}^{(j_1)}] \tag{4.24}$$

is a completely positive map $\mathcal{N} : \mathcal{B}(\mathcal{H}_{a_N}^o \otimes \cdots \otimes \mathcal{H}_{a_1}^o) \rightarrow \mathcal{B}(\mathcal{H}_{s_{N+1}}^o \otimes \mathcal{H}_{a_N}^i \otimes \cdots \otimes \mathcal{H}_{a_1}^i)$; trace preservation implies, that, if all the maps $\{\mathcal{M}^{(j_\alpha)}\}$ are CPTP, then the final output state is of unit trace.

More generally, even though we motivated the process tensor as acting on independent control operations, its action is – in the same way as the deterministic quantum combs of Ch. 2 – also well-defined on temporally correlated operations. For example, an experimenter at time t_α could condition their choice of instrument on an outcome at $t_{\alpha'} < t_\alpha$, which would lead to classically correlated control operations. More generally, the experimenter at $t_{\alpha'}$ could send forward the quantum system they used to implement their control operation $\mathcal{M}^{(j_{\alpha'})}$, and the experimenter at t_α could use the said system to implement their operation, which generally leads to control operations that temporally are correlated in a genuinely quantum way (see Fig. 4.6).

In the language of Ch. 2, this corresponds to the contraction of the comb $\mathbb{T}_{N+1:1}$ with a comb $\mathbf{M}^{(j_{N:1})}$. If $\mathbf{M}^{(j_{N:1})}$ is deterministic, *i.e.*, if it can be implemented with unit probability, then the trace preservation condition of $\mathbb{T}_{N+1:1}$ implies that the resulting output state is of unit trace. Otherwise the trace of the output state corresponds to the success probability of implementation of $\mathbf{M}^{(j_{N:1})}$. Temporally correlated control operations are, for example, used to optimally solve information theoretic tasks [47]. We will encounter them again in Sec. 4.9 when we discuss the Markov order of quantum processes. Despite being a little bit more technical, both complete positivity, as well as trace preservation for process tensors are the natural extension of the analogous properties for the superchannel case. In the same manner, the experimental reconstruction that we discussed for the case of superchannels directly translates to the multi-time case.

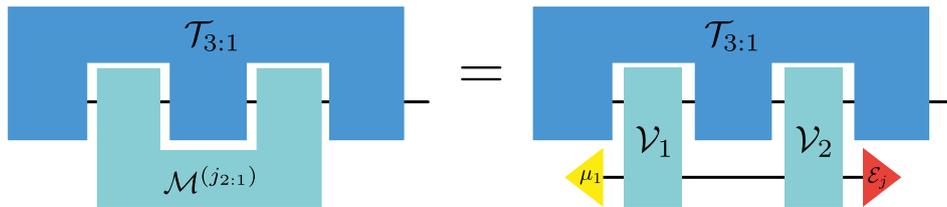


Figure 4.6: Example of a correlated control operation over two times. The experimenter at t_1 correlates the state μ_1 of an ancillary system they prepare with the system state via a unitary \mathcal{V}_1 and feeds the resulting ancilla state forward. At t_2 , the experimenter performs a generalized measurement on the system, by correlating it via the unitary \mathcal{V}_2 with the ancilla they receive from the experimenter, and performing a measurement (with outcome j) on the ancilla. If the corresponding effect \mathcal{E}_j is trace preserving (*i.e.*, the trace operation), then $\mathcal{M}^{(j_{2:1})}$ is deterministic, and the final system state is of unit trace. Otherwise, the trace of the final system state corresponds to the success probability of implementing $\mathcal{M}^{(j_{2:1})}$.

Due to its linearity, the process tensor can be experimentally reconstructed by measuring the output states $\{\rho^{(\vec{j})'}\}$ corresponding to a basis $\{\mathcal{M}^{(\vec{j})}\}$ of sequences of input operations. We have

$$\mathbb{T}_{N+1:1} = \sum_{\vec{j}} \rho^{(\vec{j})'} \otimes \mathfrak{m}^{(\vec{j})*}, \quad (4.25)$$

where $\{\mathfrak{m}^{(\vec{j})}\}$ is the dual set to the basis $\{\mathcal{M}^{(\vec{j})}\}$, *i.e.*, $\text{tr}(\mathfrak{m}^{(\vec{j})} \mathcal{M}^{(\vec{\ell})\dagger}) = \delta_{\vec{j}\vec{\ell}}$. As for the case of superchannels, the validity of Eq. (4.25) can be shown by direct insertion. At each time t_α , there are d_s^4 linearly independent control operations, and consequently, for N times, there are d_s^{4N} linearly independent sequences $\mathcal{M}^{(\vec{j})}$ of control operations. Notably, the basis of control operations could be sequences of the form $\mathcal{M}^{(j_N)} \otimes \dots \otimes \mathcal{M}^{(j_1)}$, but also temporally correlated operations. As long as they are linearly independent, any set of d_s^{4N} control operations will yield the same process tensor $\mathbb{T}_{N+1:1}$.

While the process tensor formalism was developed for the description of general open quantum processes, we see that, mathematically, it is a deterministic N -slot comb, and coincides with the descriptors of quantum stochastic processes we discussed in the previous chapter. Looked at it through this lens, all the physical elements that make up open quantum system dynamics now assume the role of the building blocks of quantum networks that we considered in Ch. 2.

Having discussed higher order quantum maps in Ch. 2, and general (quantum) stochastic processes in the previous chapter, it might appear somewhat excessive to have introduced process tensors in this section as the natural descriptor of open quantum system dynamics. However, *a priori*, the connection between these fields is unclear from our previous discussions. Higher order quantum maps and stochastic processes as we introduced them are rather abstract concepts, and their emergence from an underlying physical process is not directly obvious. Here, this connection is provided in an intuitive way.

On the other hand, considering open quantum system dynamics as a quantum network allows one to make clear-cut assertions about information flow, or, put in the language of open quantum systems, memory effects that play a non-negligible role, and to use the structure of Choi states for their quantification. Before we do so in Sec. 4.7, next, we shall complete the

discussion of the connections between higher order quantum maps and open quantum system dynamics.

4.5 PURIFICATIONS, DILATIONS, AND CHOI STATES

In the previous section, we have seen, that every open system dynamics, with initial state ρ_{se} and intermediary system-environment dynamics \mathcal{U}_α , where control operations can be applied at times t_N, \dots, t_1 , corresponds to a deterministic comb⁹ $\mathbb{T}_{N+1:1}$ with N slots and a final output wire. Somewhat unsurprisingly, the converse also holds. We have the following theorem, due to Chiribella *et al.* [17, 32]:

THEOREM 4.1: *Every deterministic N -slot comb $\mathbb{T}_{N+1:1}$ can be represented as a quantum network with initial state ρ_{se} , N unitaries \mathcal{U}_α and a final partial trace.*

Put differently, every comb can be represented as an open quantum system dynamics. While we do not provide the explicit proof here (it can, for example, be found in Ref. [17]), we shall embed it into the wider idea of *purification*, which is one of the fundamental principles in the axiomatization of quantum mechanics [11, 153–158].

In Ch. 2, we considered states, transformations and effects as the basic building blocks of quantum mechanics. Making the reasonable demand that any physical theory be fundamentally reversible [11, 153] dictates that each of these building blocks should arise from a combination of pure states, unitary dynamics, and a discarding of degrees of freedom. Irreversibility then arises from our ignorance of additional degrees of freedom, but is not a fundamental trait of nature. In this sense, any meaningful fundamental theory must be *purifiable* [11].

It is well-known that states, transformations and effects can indeed be purified. Any quantum state $\rho \in \mathcal{B}(\mathcal{H})$ can be extended to a pure state $|\Psi_\rho\rangle \in \mathcal{H} \otimes \mathcal{H}_a$, such that $\text{tr}_a(|\Psi_\rho\rangle\langle\Psi_\rho|) = \rho$. For example, we can choose

$$|\Psi_\rho\rangle = \sum_{i=1}^{d_R} \sqrt{\lambda_i} |v_i\rangle |a_i\rangle, \quad (4.26)$$

where $\{\lambda_i\}$ ($\{|v_i\rangle\}_{i=1}^{d_R}$ are the eigenvalues (eigenvectors) of ρ , d_R is the rank of ρ , and $\{|a_i\rangle \in \mathcal{H}_a\}_{i=1}^{d_R}$ are mutually orthogonal states.

Analogously, Stinespring's theorem [94] tells us that *every* CPTP map $\mathcal{L} : \mathcal{B}(\mathcal{H}_s^1) \rightarrow \mathcal{B}(\mathcal{H}_s^0)$ can be obtained from an initially uncorrelated system-environment state on $\mathcal{B}(\mathcal{H}_s^1 \otimes \mathcal{H}_a^1)$, a system-environment unitary $\mathcal{U} : \mathcal{B}(\mathcal{H}_s^1 \otimes \mathcal{H}_a^1) \rightarrow \mathcal{B}(\mathcal{H}_s^0 \otimes \mathcal{H}_a^0)$ and a final discarding of the excessive degrees of freedom:

$$\mathcal{L}[\rho] = \text{tr}_{a^0} (\mathcal{U}[\rho \otimes \eta_a]). \quad (4.27)$$

Then, since the only trace preserving effect is the trace operation, all the building blocks of a deterministic quantum network can be purified. Consequently, one could prove Thm. 4.1 by individually purifying each of the constituents of the deterministic comb $\mathbb{T}_{N+1:1}$. The resulting network would be a purification (or *dilation*) to a network that only contains a pure initial state, unitary transformations, and a final trace over excessive degrees of freedom.

⁹ Henceforth, whenever there is no risk of confusion, we shall use the terms comb and process tensor interchangeably.

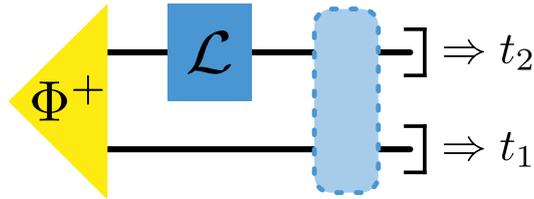


Figure 4.7: Experimental preparation of the Choi state of a channel \mathcal{L} . The resulting quantum state (represented by the blue box with dotted outline) contains the temporal correlations of the process in terms of its spatial correlations. The brackets denote the degrees of freedom of \mathcal{L} that correspond to the different times $\{t_1, t_2\}$ of the process.

Finally, trace non-preserving operations also possess a ‘purification’; every CP map (including effects) can be represented as an initially uncorrelated system-environment state, a unitary dynamics, a projective measurement on excessive degrees of freedom, and a final trace. For example, by Neumark’s theorem¹⁰ [159–161], a POVM element E_j^T can be implemented via

$$\text{tr}(\rho E_j^T) = \text{tr}(\mathcal{U}[\rho \otimes \eta] |j\rangle\langle j|). \quad (4.28)$$

Consequently, by dilating all trace non-preserving elements of a quantum network, and pulling all the wires with measurements to the far right [162], we can also see that probabilistic combs can be implemented by means of pure states, unitary dynamics, and *one* final projective measurement [15, 17].

Importantly, the fact that any deterministic comb can be dilated means that quantum networks and open system dynamics are two sides of the same coin, providing every comb with a physical realization. The final remaining question regarding this connection of quantum networks and open system dynamics concerns whether the Choi state of a map $\mathcal{T}_{N+1:1}$ corresponds to an actual physical object, or if it is merely a mathematical tool with nice properties that is used to describe open quantum evolution. It turns out that the former is the case.

We have already seen, that – up to normalization – the Choi state $\mathbb{L} \in \mathcal{B}(\mathcal{H}_{s_2}^o \otimes \mathcal{H}_{s_1}^i)$ of a map $\mathcal{L} : \mathcal{B}(\mathcal{H}_{s_1}^i) \rightarrow \mathcal{B}(\mathcal{H}_{s_2}^o)$ is obtained by letting \mathcal{L} act on one half of a maximally entangled state $|\Phi^+\rangle = \sum_{i=1} \frac{1}{\sqrt{d_s}} |i_s\rangle |i_s\rangle$, where $\{|i_s\rangle\}_{i=1}^{d_s}$ is the computational basis of $\mathcal{H}_{s_1}^i$, and for the remainder of this section, $|\Phi^+\rangle$ will denote the *normalized* maximally entangled state. We have

$$\mathbb{L} = d_s (\mathcal{L} \otimes \mathcal{I}_{s_1^i}) [|\Phi^+\rangle\langle\Phi^+|]. \quad (4.29)$$

The resulting quantum state contains the temporal properties of the process in terms of spatial correlations (see Fig. 4.7). From an experimental perspective, Eq. (4.29) means that – up to normalization – the matrix \mathbb{L} can be prepared without any knowledge of the underlying process.

In an analogous way, the Choi state $\mathbb{T}_{N+1:1}$ of a process tensor $\mathcal{T}_{N+1:1}$ can be prepared experimentally by feeding one half of a maximally entangled state into the process at every time t_α , *i.e.*, by an N -fold concatenation of Eq. (4.29) (see Fig. 4.8). Mathematically, we have

$$\mathbb{T}_{N+1:1} = d_s^N (\mathcal{T}_{N+1:1} \bigotimes_{\alpha=1}^N \mathcal{R}_{\Phi^+}^{(\alpha)}) [\mathbf{S}_{s\alpha}^{(\alpha)}], \quad (4.30)$$

¹⁰ In the original form, the space does not have to be extended by a tensor product, but merely a tensor sum. The version we provide lends itself more easily to experimental implementation and is sufficient for our purposes.

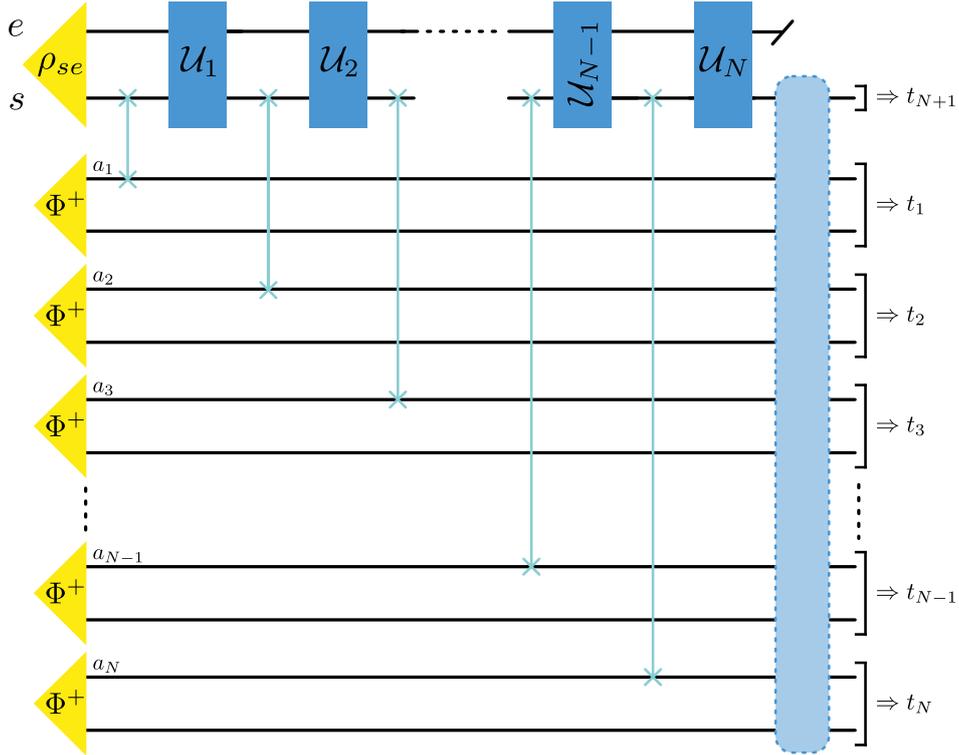


Figure 4.8: Experimental preparation of $\mathbb{T}_{N+1:1}$. At each time t_α , one half of a maximally entangled state is fed into the process by means of a swap operation (depicted by the light green vertical lines). The resulting many-body quantum state (after tracing over the degrees of freedom of the environment) is the Choi state $\mathbb{T}_{N+1:1}$ (denoted by the blue box with dotted outline). The brackets signify the degrees of freedom of $\mathbb{T}_{N+1:1}$ that correspond to the different times t_α of the process. Temporal correlations in the process will be expressed as spatial correlations in $\mathbb{T}_{N+1:1}$.

where $\mathcal{R}_{\Phi^+}^{(\alpha)}$ is the map that prepares the α^{th} maximally entangled state, $\mathbf{S}_{sa}^{(\alpha)}$ is the swap operation between the system s and one half a_α of the α^{th} maximally entangled state, and we have omitted the respective identity maps [8].

The validity of Eq. (4.30) has been shown explicitly in [32]. As for the case of channels, it implies, that $\mathbb{T}_{N+1:1}$ can be prepared experimentally – up to normalization – without any knowledge of the process. Moreover, via the Choi isomorphism, all temporal correlations of the process $\mathbb{T}_{N+1:1}$ – *i.e.*, its memory structure – are mapped onto spatial correlations of $\mathbb{T}_{N+1:1}$.

Indeed, we have already encountered this fact when we discussed the structure of superchannels for the case of an initially uncorrelated system-environment state in Sec. 4.2. There, we saw that the corresponding superchannel is of the form $\mathbb{T} = \mathbb{L} \otimes \rho_s$, which, as we will discuss below, implies that the future dynamics is independent of the initial state preparation. Throughout this chapter, all of our analysis of memory effects will rely heavily on this correspondence of temporal correlations of the process, and the spatial correlations of its descriptor $\mathbb{T}_{N+1:1}$.

Having established process tensors as the most general descriptors of open quantum system dynamics, as well as their equivalence to the theory of higher order quantum maps, we will now use this framework, to analyze the memory properties of general open quantum processes.

4.6 MEMORY IN CLASSICAL PROCESSES

As soon as the dynamics of a system is open, memory effects can play a non-negligible role. Intuitively speaking, through the interaction with the environment, due to back-action, the future dynamics at a time t_α can depend on earlier states of the system, and not just on its state at t_α . Before we discuss this concept in detail, we shall first introduce it in the classical realm.

In classical physics (without interventions), the definition of memory for a given process is unambiguous; let $\mathbb{P}(x_{N+1}, t_{N+1}; x_N, t_N; \dots; x_1, t_1)$ denote the joint probability distribution describing a process. Whenever there is no risk of confusion, for compactness, we will drop the time arguments, and the respective time is denoted by the corresponding subscript, *i.e.*, x_α is an outcome at t_α .¹¹ We consider the process to be *memoryless* or *Markovian*, if its conditional probabilities satisfy

$$\mathbb{P}(x_{N+1}|x_N, \dots; x_1) = \mathbb{P}(x_{N+1}|x_N) \quad \forall N. \quad (4.31)$$

Intuitively, for a Markovian process, the probability to find the particle in region x_{N+1} at time t_{N+1} only depends on the region it was found in at t_N , but not at any of its positions at earlier times.¹² Put differently, the future evolution of a Markovian process from time t_N onward only depends on its state at t_N , but not on its *trajectory* up to t_N .

Importantly, even though we shall continue to call Markovian processes memoryless, this memorylessness is in general only a *conditional* one. To see this more clearly, consider the case of Markovian Brownian motion. Evidently, the probability to find a particle in some region x_{N+1} at time t_{N+1} is not independent of where it was initialized at t_1 . However, once its position at time t_N is known, the information of where it was initialized at t_1 does not provide any new information about the whereabouts of the particle at t_{N+1} . As for a Markovian process the conditional probabilities $\mathbb{P}(x_{N+1}, t_{N+1}|x_N, t_N; \dots; x_1, t_1) = \mathbb{P}(x_{N+1}, t_{N+1}|x_N, t_N)$ depend on the last time step only, they are sometimes considered to be of *Markov order 1*. A completely memoryless process, like, for example, a coin flip, would then be a process of Markov order zero, *i.e.*, $\mathbb{P}(x_{N+1}, t_{N+1}|x_N, t_N; \dots; x_1, t_1) = \mathbb{P}(x_{N+1}, t_{N+1})$ for all N . We generally drop this explicit distinction, and unless stated otherwise, Markovian processes will be considered to be processes of Markov order one or zero, and will be called memoryless.

Memoryless processes are of particular interest, as they allow one to obtain all higher order joint probability distributions from transition probabilities and an initial probability distribution. In detail, for a Markovian process, we have

$$\begin{aligned} \mathbb{P}(x_{N+1}, \dots, x_1) &= \mathbb{P}(x_{N+1}|x_N, \dots, x_1)\mathbb{P}(x_N, \dots, x_1) \\ &= \mathbb{P}(x_{N+1}|x_N)\mathbb{P}(x_N, \dots, x_1) \\ &= \dots = \prod_{\alpha=2}^{N+1} \mathbb{P}(x_\alpha|x_{\alpha-1})\mathbb{P}(x_1), \end{aligned} \quad (4.32)$$

¹¹ Additionally, as we do not have to sum over outcomes in this section, we can strip down our subscripts a little, *i.e.*, we will denote outcomes at t_α simply as x_α , instead of x_{j_α} . We shall return to more detailed subscripts whenever necessary.

¹² Obviously, the measured observable does not have to be position. Position is simply used for illustrative purposes.

which implies a substantial simplification in modeling complexity. While generally, this complexity grows exponentially with the length of the memory [163–165], here, it is sufficient to know the transition probabilities $\mathbb{P}(x_\alpha|x_{\alpha-1})$ and the initial distribution $\mathbb{P}(x_1)$ to fully model the process.

In an analogous way, we can define processes of higher Markov order. We consider a process to be of Markov order ℓ , if we have

$$\mathbb{P}(x_{N+1}|x_N, \dots, x_1) = \mathbb{P}(x_{N+1}|x_N, \dots, x_{N-\ell+1}) \quad \forall N. \quad (4.33)$$

As for the case of Markovian processes, processes of Markov order ℓ allow one to build up all joint probability distributions from lower order ones; if the process is of Markov order ℓ , then we have

$$\begin{aligned} \mathbb{P}(x_{N+1}, \dots, x_1) &= \mathbb{P}(x_{N+1}|x_N, \dots, x_1)\mathbb{P}(x_N, \dots, x_1) \\ &= \mathbb{P}(x_{N+1}|x_N, \dots, x_{N-\ell+1})\mathbb{P}(x_N, \dots, x_1) \\ &= \dots = \prod_{\alpha=\ell+1}^{N+1} \mathbb{P}(x_\alpha|x_{\alpha-1}, \dots, x_{\alpha-\ell})\mathbb{P}(x_\ell, \dots, x_1), \end{aligned} \quad (4.34)$$

and again, the process is fully characterized once the transition probabilities $\mathbb{P}(x_\alpha|x_{\alpha-1}, \dots, x_{\alpha-\ell})$ and the ‘initial’ probability distribution $\mathbb{P}(x_\ell, \dots, x_1)$ are known. As already alluded to, understanding the memory structure of a process is of crucial importance when trying to simulate a process, as well as in order to make assertions about the underlying microscopic model. It is hence an important question, how to generalize the concept of memory to the quantum case.

4.7 MEMORY IN QUANTUM PROCESSES

While in classical physics it is clear – at least in principle – how to deduce the memory length of a process, the situation presents itself much murkier in quantum mechanics. Here, measurements change the state of the system, and it is *a priori* not obvious how to meaningfully extend the notion of Markovianity – which relies on conditional probability distributions – from the classical to the quantum case. Consequently, a plethora of different, inequivalent definitions and corresponding measures of Markovianity for quantum processes can be found. What is common to most of these attempts to define non-Markovianity, is that they give sufficient conditions, like, *e.g.*, the breakdown of the monotonicity of trace-distance distinguishability [166], a failure of the divisibility of dynamics [131, 167] (see also below), or the detectable presence of initial correlations [136, 137, 168, 169], to name but a few (for a review of definitions of non-Markovianity, see, *e.g.*, [65, 170]). However, they fail to provide necessary conditions, and, most importantly, none of the aforementioned definitions reduces to Eq. (4.32) in the classical limit (for a review of different approaches to non-Markovianity that circumvent the aforementioned problems, see, *e.g.*, [66]).

Fundamentally, the shortcoming of most previous approaches is that they are based on a description of open system dynamics in terms of quantum channels. As we have seen in the previous sections, channels are not, in general, the correct descriptor of a quantum stochastic

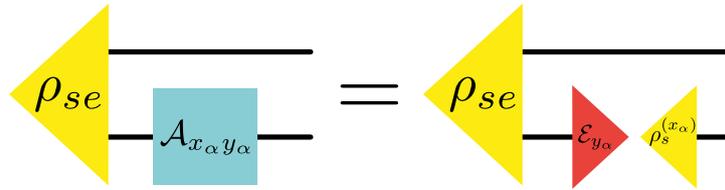


Figure 4.9: Action of a causal break $\mathcal{A}_{x_\alpha y_\alpha}$ on a system-environment state ρ_{se} . First, a measurement with corresponding effect \mathcal{E}_{y_α} is performed, followed by an independent preparation of a fresh state $\rho^{(y_\alpha)}$. The resulting system-environment state is of the form $\rho^{(x_\alpha)} \otimes \text{tr}_s(\rho_{se} \mathbf{E}_{y_\alpha}^T)$. Importantly, the causal break introduces a break of information flow on the level of the system – any information from the past that can be detected after the causal break must have ‘traveled’ through the environment.

process when memory effects play a role and when genuine multi-time correlations are of interest; channels can – by design – only capture two point correlations, and consequently, no definition of non-Markovianity in the quantum case that is based on the properties of quantum channels can fully capture all memory effects. We will analyze this problem in detail in Sec. 4.8 for the measure of CP divisibility, which is frequently employed to check if open system dynamics are Markovian. Additionally, many of the measures in use, while being mathematically well-defined, lack a proper operational underpinning. The process tensor approach, that we encountered in the previous section is tailored to capture genuine multi-time correlations, and allows one to remedy the problems encountered in properly defining the boundary between Markovian and non-Markovian processes in an operationally meaningful way which permits both the detection and quantification of memory effects.

Markovianity of a classical process is defined in terms of a conditional independence property; conditioned on the most recent state of the system, the future statistics is independent of the past. The problem of properly defining Markovianity in the quantum case lies in the invasiveness of measurements, which seems to prevent the investigation of conditional probabilities. However, as we have seen, the process tensor approach allows one to unambiguously define a process independent of the control operations that an experimenter performs on the system, and as such, it allows for the generalization of conditional probabilities to the quantum case [32].

Intuitively, the history-dependence of a process at time t_N can be probed by fixing its state at t_N and analyzing its future evolution for different pasts. This intuition can be made manifest as follows: Let the experimenter perform any admissible sequence of operations¹³ $\mathcal{M}_{N-1:1}$ – temporally correlated or not – at times t_{N-1}, \dots, t_1 . At time t_N they measure the system – with corresponding effect \mathcal{E}_{y_N} – and independently reprepare it in a known fresh state $\rho_s^{(x_N)}$. This operation has two purposes. On the one hand, the independent reparation breaks the flow of information on the level of the system (see Fig. 4.9). On the other hand, it leaves the system in a known state, which will allow us to ‘condition’ the future statistics on this state. We shall thus call the corresponding CP map a ‘causal break’ [105]. Mathematically, a causal break at t_N corresponds to a CP operation $\mathcal{A}_{x_N y_N}$, with its action on a system-environment state defined as

$$(\mathcal{A}_{x_N y_N} \otimes \mathcal{I}_e)[\rho_{se}] = (\mathbf{E}_{y_N} \otimes \rho_s^{(x_N)}) \star \rho_{se} = \rho_s^{(x_N)} \otimes \text{tr}_s[\rho_{se}(\mathbf{E}_{y_N}^T \otimes \mathbb{1}_e)] \quad (4.35)$$

¹³ For ease of future notation, from here on, we will label all operations by subscripts.

where ρ_{se} is the system-environment state at t_N , E_{y_N} is the POVM element corresponding to the performed measurement, and $\rho_s^{(y_N)}$ is the freshly prepared state [105]. From Eq. (4.35), we can see that the system-environment state after a causal break is of product form, and the system part only depends on the causal break, but not on ρ_{se} . However, the environment part can explicitly depend on ρ_{se} , and as such on all the operations $\mathcal{M}_{N-1:1}$, that were performed before the causal break. Consequently, if at any time step $t_{N'} > t_N$ one can discern two different sequences of past operations $\{\mathcal{M}_{N-1:1}, \mathcal{M}'_{N-1:1}\}$ and/or different measurements (with corresponding POVM elements $\{E_{y_N}, E'_{y_N}\}$), then memory must have passed through the environment.

In more detail, just like in the classical case, a process has memory if the state of the system at the next time step t_{N+1} does not only depend on the freshly prepared state $\rho_s^{(x_N)}$, but also on the measurement outcome at t_N and/or the sequence of operations $\mathcal{M}_{N-1:1}$ performed previously [105]. Expressed in terms of the process tensor, a process is Markovian iff for all N

$$\mathcal{T}_{N+1:1}[\mathcal{A}_{x_N y_N}, \mathcal{M}_{N-1:1}] \propto \mathcal{T}_{N+1:1}[\mathcal{A}_{x_N y'_N}, \mathcal{M}'_{N-1:1}], \quad (4.36)$$

$\forall x_N, y_N, y'_N, \mathcal{M}_{N-1:1}, \mathcal{M}'_{N-1:1}$, where the proportionality appears instead of an equality, as the operations $\mathcal{M}_{N-1:1}, \mathcal{M}'_{N-1:1}, \mathcal{A}_{x_N y_N}$ and $\mathcal{A}_{x_N y'_N}$ are not necessarily trace preserving [105]. If Eq. (4.36) is not satisfied, then the process is non-Markovian, since the only way past actions could influence the future evolution after a causal break is through some kind of memory that ‘traveled’ through the environment.

Put differently, if a quantum process is Markovian, once the state $\rho_s^{(x_N)}$ is known, then the process after t_N is independent of the past. This perfectly mirrors the conditional independence of classical Markovian processes of their history. Here, the sequence of operations $\mathcal{M}_{N-1:1}$ until t_{N-1} together with the measurement associated to E_{y_N} constitutes the ‘history’ or ‘trajectory’ of the system up until t_N , whereas $\rho_s^{(x_N)}$ is its state at t_N . The concept of operationally well-defined quantum trajectories¹⁴ has recently been used to define Markovianity in an equivalent way to Eq. (4.36) [7, 128].

Notably, in (4.36) the measurement made at t_N could also be temporally correlated with the previous operations $\mathcal{M}_{N-1:1}$ without making the definition of Markovianity more general (see below). Since $\mathcal{T}_{N+1:1}$ is a linear operator, Eq. (4.36) allows for the process to be checked for Markovianity with a finite number of experiments [32, 105]. At each time, there are d_s^2 linearly independent POVM elements, and d_s^2 linearly independent reparations, which means that there are d_s^4 linearly causal breaks at each time step. Consequently, in order to check if Eq. (4.36) is satisfied for all possible combinations of causal breaks and previous operations, one would have to perform full tomographic reconstruction of $\mathcal{T}_{N+1:1}$. On the other hand, finding two different pasts that lead to different future statistics is already sufficient to show the non-Markovianity of a process.

The definition of Markovianity in terms of causal breaks at once remedies the shortcomings that many traditional and experimentally employed approaches.¹⁵ On the one hand, it is a

¹⁴ Importantly, here we do not attach any ontological meaning to trajectories, but consider them as the sequence of operations corresponding to a sequence of outcomes.

¹⁵ An equivalent definition of quantum Markovianity can also be given in terms of a *Generalized Quantum Regression Formula* (see, e.g., [66]). As the definition in terms of causal breaks ties in more organically with the quantum comb framework, we shall use it exclusively throughout this thesis.

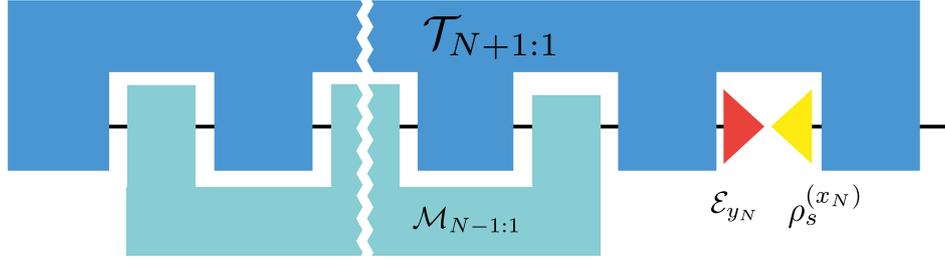


Figure 4.10: Process tensors and causal breaks. In order to check for Markovianity, one has to check for independence of the final state at t_{N+1} from earlier operations $\mathcal{M}_{N-1:1}$, and the measurement outcome at t_N (corresponding to the effect \mathcal{E}_{y_N}). If a dependence can be detected, then information/memory must have traveled through the environment, and the process is non-Markovian. Note that the white ‘break’ in the figure signifies that the respective elements extend over several time steps, not that there is a break in information flow. For better visibility, the labels of the effect and the reparation are placed under the corresponding graphical element.

logical statement of conditional independence, instead of a mathematical property, that might or might not be directly connected to Markovianity. On the other hand, the definition given in Eq. (4.36) coincides with the classical definition in the correct limit [105]; in detail, if we fix the (potentially different) instruments \mathcal{J}_α that can be used at each time, and each of them only consist of causal breaks, we can label the CP maps that make up the instrument \mathcal{J}_α by $\{\mathcal{A}_{z_\alpha}\}$. As the instruments are fixed, we obtain probability distributions $\mathbb{P}(z_{N+1}, z_N, \dots, z_1)$ from using these instruments to interrogate the process. Now, it is obvious to see that in this case, Eq. (4.36) implies that these probability distributions satisfy

$$\mathbb{P}(z_{N+1}|z_N, \dots, z_1) = \mathbb{P}(z_{N+1}|z_N) \quad \forall N, \quad (4.37)$$

which corresponds to the classical definition of Markovianity. In particular, this holds true, if all the causal breaks are projective measurements in the computational basis, *i.e.*, $\mathcal{A}_{z_\alpha} = |z_\alpha\rangle\langle z_\alpha| \otimes |z_\alpha\rangle\langle z_\alpha|$, which, arguably, most closely resembles to what one would consider a classical instrument. However, we already recover the original definition of Markovianity under the slightly weaker requirement that all employed instruments are fixed, consist of causal breaks only, and the process satisfies Eq. (4.36). However, this does not mean that such processes are indeed classical; the corresponding joint probability distributions would not necessarily have to satisfy Kolmogorov conditions.

Finally, even though the definition of Markovianity we introduced is intuitively plausible, it seems to rely on the concept of causal breaks. However, as we shall now see, Markovianity corresponds to a structural property of processes, that could be probed even without employing causal breaks. To see this, we rewrite Eq. (4.36) in terms of Choi states. A process is Markovian iff for all $\rho^{(x_N)}$, \mathcal{E}_{y_N} , \mathcal{E}'_{y_N} , $\mathcal{M}_{N-1:1}$, and $\mathcal{M}'_{N-1:1}$ we have

$$\mathcal{T}_{N+1:1} \star \rho^{(x_N)} \star \mathcal{E}_{y_N} \star \mathcal{M}_{N-1:1} \propto \mathcal{T}_{N+1:1} \star \rho^{(x_N)} \star \mathcal{E}'_{y_N} \star \mathcal{M}'_{N-1:1}. \quad (4.38)$$

By writing the link product out explicitly, this is equivalent to

$$\begin{aligned} \text{tr}_{N+1^0} \left[\left(\mathbb{1}_{s_{N+1}}^0 \otimes \rho^{(x_N)\text{T}} \otimes \mathcal{E}_{y_N}^{\text{T}} \otimes \mathcal{M}_{N-1:1}^{\text{T}} \right) \mathcal{T}_{N+1:1} \right] \\ \propto \text{tr}_{N+1^0} \left[\left(\mathbb{1}_{s_{N+1}}^0 \otimes \rho^{(x_N)\text{T}} \otimes \mathcal{E}'_{y_N}^{\text{T}} \otimes \mathcal{M}'_{N-1:1}^{\text{T}} \right) \mathcal{T}_{N+1:1} \right], \end{aligned} \quad (4.39)$$

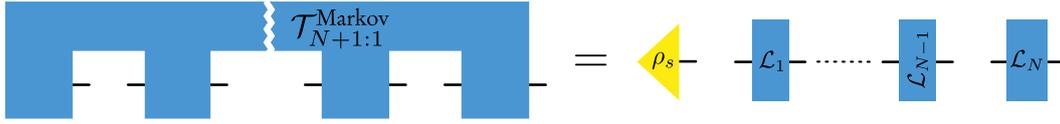


Figure 4.11: Structure of a Markovian process. Every Markovian process can be represented as a concatenation of independent CPTP maps. As the maps are not connected by environmental wires, no information can travel outside of the system, and the dynamics is Markovian.

where tr_{N+1^0} signifies the trace over all spaces but $\mathcal{H}_{s_{N+1}}^0$. In this form, it is easy to see that for this conditional independence to hold for *all* possible histories, the process tensor has to be of the form $\mathbb{T}_{N+1:1} = \mathbb{L}_N \otimes \mathbb{T}_{N:1}$, where $\mathbb{L}_N \in \mathcal{B}(\mathcal{H}_{s_{N+1}}^0 \otimes \mathcal{H}_{s_N}^1)$ is a CPTP map from t_N to t_{N+1} and $\mathbb{T}_{N:1} \in \mathcal{B}(\mathcal{H}_{s_N}^0 \otimes \mathcal{H}_{s_{N-1}}^1 \otimes \cdots \otimes \mathcal{H}_{s_1}^0)$ is a $(N-1)$ -slot process tensor on the times $\{t_N, \dots, t_1\}$. As this has to hold at all times, the process tensor $\mathbb{T}_{N+1:1}^{\text{Markov}}$ of a Markovian dynamics is of product form, *i.e.*,

$$\mathbb{T}_{N+1:1}^{\text{Markov}} = \mathbb{L}_N \otimes \mathbb{L}_{N-1} \otimes \cdots \otimes \mathbb{L}_2 \otimes \rho_s, \quad (4.40)$$

where $\rho_s \in \mathcal{B}(\mathcal{H}_{s_1}^0)$ is the initial state of the system. The process tensor of a Markovian process is simply a tensor product of independent CPTP maps [42, 50, 147] (see Fig. 4.11). This structural property also allows for the definition of operationally well-defined measures of Markovianity [105] as the deviation of a given process tensor from the product structure of Eq. (4.40). Importantly, the product form of Markovian processes does not coincide with CP divisibility [131, 167] (see below), and is a direct consequence of the operational condition in terms of causal breaks that we gave above, rather than the starting point of the discussion of Markovianity in the quantum case. Markovianity, as it is defined here, is a logical statement about conditional independence, and all mathematical properties follow from this requirement. It is important to emphasize this point, as it does not apply for the mathematically appealing, but not operationally motivated definitions of Markovianity in the quantum case, that exist in the literature.

Additionally, Eq. (4.40) implies that Markovian processes can be reconstructed with exponentially less resources than non-Markovian ones. Under the assumption that the experimenter is guaranteed that the underlying process is memoryless, it suffices for them to tomographically reconstruct each of the individual maps \mathbb{L}_α for a full characterization of the process. Each channel reconstruction requires the collection of d_s^4 conditional probabilities and the individual reconstructions can be carried out independently of each other. Concretely, this can, for example, be achieved by means of projective measurements $\{|z_{j\alpha}\rangle\langle z_{j\alpha}| \otimes |z_{j\alpha}\rangle\langle z_{j\alpha}|\}_{j\alpha=1}^{d_s^2}$, where $\{|z_{j\alpha}\rangle\langle z_{j\alpha}|\}_{j\alpha=1}^{d_s^2}$ constitutes a basis of $\mathcal{B}(\mathcal{H}_s)$. Consequently, the number of necessary measurements to reconstruct an N -step Markovian process scales as Nd_s^4 , while in general, in the presence of memory effects, it scales as d_s^{4N} .

The structure of Markovian processes given by Eq. (4.40) allows for another intuitive interpretation: each of the maps \mathbb{L}_α allows for a dilation in terms of an interaction (given by the a unitary map \mathcal{U}_α) of the system with an initially uncorrelated environment state η_α that is subsequently discarded (see Fig. 4.12). In this sense, Markovian dynamics can be considered as a ‘collision’ of the system with a fresh environment state at each time t_α [171–175]. Interpreted in more physical terms, the characteristic time for memory to persist is significantly smaller

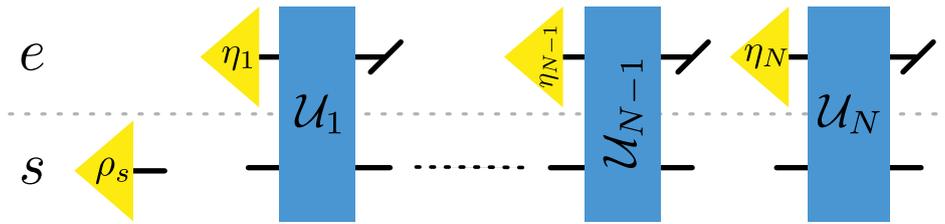


Figure 4.12: Possible dilation of a Markovian process. At each time t_α the system ‘collides’ with a fresh environment state η_α , which is discarded at the next time. Consequently, no memory can be transported via the environment.

than the difference of times $t_\alpha - t_{\alpha-1}$ at which the system is interrogated, and the environment relaxes to its equilibrium state in between control operations.

Below, we will discuss the extension of Markov order to the quantum case. However, first it is important to use the gained insights to make sense of existing witnesses of non-Markovianity, and examine their place in the wider framework of process tensors and the accompanying unambiguous definition of Markovianity. Here, we shall do so for CP divisibility, one of the predominantly used measures for Markovianity (or rather the absence thereof) in the literature. The results of this investigation are based on the findings in [1].

4.8 CP DIVISIBILITY AND MARKOVIANITY

The detection of Markovianity, and characterization of memory effects plays an important technological role. For example, successful error correction depends on a clear understanding of the underlying noise [176]. With the process framework introduced above, we can now furnish many existing witnesses of memory effects with a clear-cut operational meaning in terms of the temporal correlations that they can and cannot detect. Here, we shall do so for CP divisibility, the concepts that underlies most measures of non-Markovianity that are in use. To motivate the concept of CP divisibility, we approach it from its classical counterpart, and then expand the corresponding ideas to the quantum case.

We have seen that for classical Markovian processes, once the initial distribution $\mathbb{P}(x_{j_1})$ and the transition probabilities $\mathbb{P}(x_{j_{\alpha+1}}|x_{j_\alpha})$ are known, the process is fully described. As we discussed, in this case, all joint probability distributions can be built up via

$$\mathbb{P}(x_{j_{N+1}}, \dots, x_{j_1}) = \mathbb{P}(x_{j_{N+1}}|x_{j_N}) \cdots \mathbb{P}(x_{j_2}|x_{j_1})\mathbb{P}(x_{j_1}). \quad (4.41)$$

The values $\mathbb{P}(x_{j_\alpha})$ can be considered as the elements of a probability vector $\vec{\mathbb{P}}$ with $\vec{\mathbb{P}}_{j_\alpha} = \mathbb{P}(x_{j_\alpha})$. Analogously, we can regard the transition probabilities $\mathbb{P}(x_{j_{\alpha+1}}|x_{j_\alpha})$ to be elements of a stochastic transition matrix $Y_{\alpha+1|\alpha}$, with elements $Y_{j_{\alpha+1}j_\alpha} = \mathbb{P}(x_{j_{\alpha+1}}|x_{j_\alpha})$, where each of the transition matrices can vary for different time steps. As the matrices $Y_{\alpha+1|\alpha}$ are stochastic, *i.e.*, their entries are probabilities, and their columns sum to 1, they can be seen as a time evolution

of the probability vector $\vec{\mathbb{P}}$ between t_α and $t_{\alpha+1}$, and from Eq. 4.41 we obtain the probability vector at t_{N+1} , with entries $\mathbb{P}(x_{j_{N+1}})$, via

$$\begin{aligned} \mathbb{P}(x_{j_{N+1}}) &= \sum_{j_N, \dots, j_1} \mathbb{P}(x_{j_{N+1}}, \dots, x_{j_1}) \\ &= (\mathcal{Y}_{N+1|N} \cdot \mathcal{Y}_{N|N-1} \cdots \mathcal{Y}_{2|1} \cdot \vec{\mathbb{P}})_{j_{N+1}} := (\mathcal{Y}_{N+1|1} \cdot \vec{\mathbb{P}})_{j_{N+1}}. \end{aligned} \quad (4.42)$$

The natural counterpart in quantum mechanics of probability vectors and stochastic matrices are quantum states and CPTP maps, respectively. Consequently, by generalizing Eq. (4.42) to the quantum case, one obtains the requirement that for Markovian dynamics, the CPTP dynamics of the system state ρ_s from t_1 to t_{N+1} should be given by

$$\mathcal{L}_{N+1|1}[\rho_s] = (\mathcal{L}_{N+1|N} \circ \mathcal{L}_{N|N-1} \circ \cdots \circ \mathcal{L}_{2|1})[\rho_s], \quad (4.43)$$

where each of the CPTP maps $\mathcal{L}_{\alpha+1|\alpha}$ is a dynamics from t_α to $t_{\alpha+1}$, and $\mathcal{L}_{N+1|1}$ is the overall map from t_1 to t_{N+1} . We can use this idea for the definition of *CP-divisible* dynamics [177]:

DEFINITION 4.1: *The dynamics of a system is CP-divisible, if its time evolution can be divided into CPTP maps, i.e., for any three times $u < v < w$ we have¹⁶*

$$\mathcal{L}_{w|u} = \mathcal{L}_{w|v} \circ \mathcal{L}_{v|u}, \quad (4.44)$$

where $\mathcal{L}_{w|u}$, $\mathcal{L}_{w|v}$ and $\mathcal{L}_{v|u}$ are CPTP maps that describe time evolutions from u to w , v to w and u to v , respectively.

Notably, Eq. (4.44) expresses a semi-group property of the dynamics [114], which provides another motivation for a connection between CP divisibility and Markovianity, as dynamics that satisfy semi-group properties can be considered the solution of a GKSL equation with positive rates [116, 178], which are often used to model Markovian processes. Importantly, though, while mathematically meaningful and well-defined [129], the *operational* meaning of each of the maps $\mathcal{L}_{y|x}$ above is not entirely clear. That is, in an experimental setting, how does one construct such a family of maps? Consequently, in the absence of a clear operational understanding of the involved maps, it would be somewhat misleading to express Eq. (4.44) in terms of a link product $\mathcal{L}_{w|u} = \mathcal{L}_{w|v} \star \mathcal{L}_{v|u}$, as the link product suggest a structure of an underlying quantum network, which is not given by Eq. (4.44) without further qualification.

At first glance, Eq. (4.44) seems to imply that the dynamics at any point in time is fully described by a CPTP map that is independent of the respective history.¹⁷ Importantly though, CP divisibility does *not* imply a product structure (of the form of Eq. (4.40)) of the underlying process tensor, and as such, does not imply Markovian dynamics, not even in the classical case [179]. From our previous discussion of Markovianity, this fact is obvious: On the one hand, information about two-point correlations, as expressed by CPTP maps, is not enough to decide the Markovianity of a process, which is a statement about multi-time correlations. On the other

¹⁶ To be closer to the notation used in the respective literature, and to strip it down a little bit, here, we denote times by u , v and w . Conversion to the notational convention used throughout this thesis is always straightforwardly possible.

¹⁷ Throughout this section, we will consider CPTP maps to describe dynamics of the state of the system between two times, in contrast to a process tensor or comb, that constitute the full descriptor of the dynamics.

hand, the definition of CP divisibility – even when supplied with an additional operational meaning¹⁸ (see below) – is not a statement about conditional independence. Consequently, while its mathematical relation to Markovianity might be plausible, its logical connection is vague at best.

Nonetheless, many measures and witnesses for non-Markovianity are built upon the breakdown of CP divisibility (this holds true, for instance, for the measures introduced and used in [65, 131, 136, 166–168, 170]) and it is thus of importance to explicitly work out, what kinds of temporal conditions can exist despite the dynamics being CP-divisible. In turn, such a discussion will provide further intuitive understanding of the breakdown of CP divisibility in terms of information flow between the system and the environment [1, 65, 166, 181].

4.8.1 *iCP and oCP Divisibility*

As mentioned above, the constituents of the definition of CP divisibility do not possess a clear operational meaning, unless one specifies, how the maps $\mathcal{L}_{y|x}$ are to be experimentally reconstructed. While this lack of operational meaning is not a problem in terms of well-definedness of CP divisibility, it prevents its interpretation with respect to prevalent memory effects. An alternative definition of CP divisibility that furnishes the maps $\mathcal{L}_{y|x}$ with a clear-cut meaning is thus desirable. There are (at least) two non-equivalent ways to do so.

The first possibility to give the respective maps an experimental meaning is to tomographically reconstruct the maps $\mathcal{L}_{w|u}$ and $\mathcal{L}_{v|u}$, and obtain the third map by inversion. We shall call the resulting definition of CP divisibility *CP divisibility by inversion* (**iCP** divisibility). In detail, the underlying idea is, that at a fixed initial time u , the experimenter can prepare fresh states, and measure the respective output states at any later time. Consequently, they can reconstruct any map $\mathcal{L}_{x|u}$ that describes the dynamics of the system from time u to any later time x . Now, for $u < v < w$, if the map $\mathcal{L}_{v|u}$ is invertible, one can compute a map $\mathcal{K}_{w|v}$ via

$$\mathcal{K}_{w|v} = \mathcal{L}_{w|u} \circ \mathcal{L}_{v|u}^{-1}, \quad (4.45)$$

where we denote the inferred map obtained by inversion with the symbol \mathcal{K} to reserve \mathcal{L} for experimentally reconstructed maps (see Fig. 4.13a). With this, we obtain the definition of iCP divisibility [1]:

DEFINITION 4.2: *For a fixed initial time u , a dynamics is iCP-divisible if for all $w > v > u$ the map $\mathcal{K}_{w|v}$ computed according to Eq. (4.45) is CPTP.*

This definition of CP divisibility is advantageous from an experimental perspective, as it only necessitates the reconstruction of quantum maps from a fixed initial time to later times, but not the explicit construction of intermediate maps [65]. However, the operational meaning of $\mathcal{K}_{w|v}$, even when it is CPTP, is somewhat restricted.

In a particular sense, it is a fictitious map of the state of the system from time v to time w ; imagine an experimenter who has reconstructed the map $\mathcal{L}_{v|u}$, *i.e.*, for any system state that they

¹⁸ One possible operational interpretation of CP-divisible *dynamics* – but *not* the maps $\mathcal{L}_{y|x}$ themselves or the concept’s relation to Markovianity – has been considered elsewhere [180], but is unrelated to the subsequent discussion.

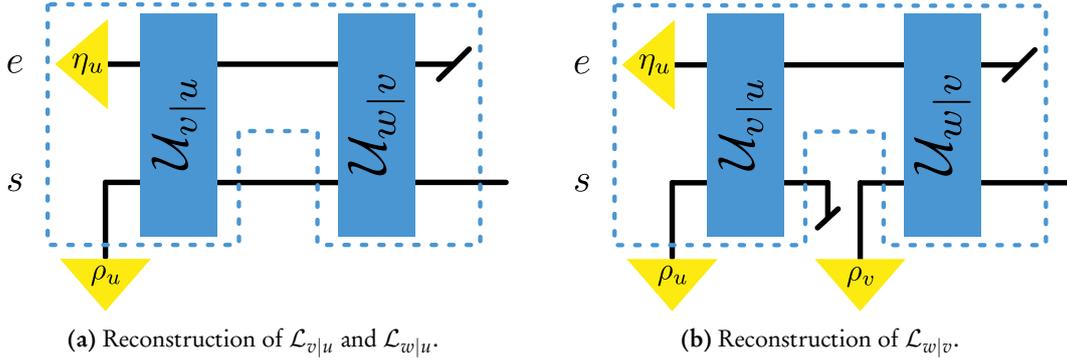


Figure 4.13: Experimental scenarios for the reconstruction of the maps $\mathcal{L}_{y|x}$. (a) For the reconstruction of $\mathcal{L}_{y|u}$, a basis of states is fed into the process at time u , and the corresponding output state is measured at time y . (b) For the reconstruction of $\mathcal{L}_{y|v}$, fresh states are fed in at time $v > u$ (thus decoupling system and environment at v), and the corresponding output state at time $y > v$ is measured. For the case considered in Ex. 4.1, both system and environment are qubits, the initial environment state η_u is maximally mixed, and $\mathcal{U}_{v|u}$ and $\mathcal{U}_{w|v}$ are the partial swap and a controlled-not gate, respectively. The (one-slot) process process tensor for both cases is depicted by the dotted blue line, with an open input and output wire. Both experimental procedures differ by the control operations, but not by the underlying process tensor.

feed into the process at time u (say ρ_u), they would know the corresponding output state ρ_v at time v .¹⁹ Letting the system, instead, evolve further to time w , the final state $\rho_w = \mathcal{L}_{w|u}[\rho_u]$ could be understood as the result of the action of $\mathcal{K}_{w|v}$ on the system state at ρ_v . However, as the experimenter did *not* intervene at v , this evolution is merely a mathematical crutch, rather than an operationally meaningful quantum channel. Nonetheless, evidently, if the underlying dynamics is Markovian, and the maps from u to v are invertible, then the dynamics also has to be iCP-divisible, which makes the breakdown of iCP divisibility (or the breakdown of properties derived therefrom) an experimentally accessible witness of non-Markovianity.

In the form defined above, the definition of CP divisibility does not tie in nicely with the process tensor framework, which presents us with a roadblock when trying to quantify the temporal correlations that can be present in a CP-divisible process. This problem can be circumvented by defining CP divisibility such that all involved maps have a clear operational meaning. While above we assumed that one has only direct access to the maps $\mathcal{L}_{v|u}$ from the initial time u to all later times v , and all intermediate maps $\mathcal{L}_{w|v}$ are obtained by inversion, one could also imagine the scenario where an experimenter has the ability to manipulate the system at any time $v > u$. Then, they can reconstruct any intermediate maps $\mathcal{L}_{w|v}$ by discarding the system state at time v , preparing a fresh state ρ_v that is uncorrelated with the environment, and measuring the corresponding output state at w . Consequently, the action of the map $\mathcal{L}_{w|v}$ reconstructed in this way is given by

$$\mathcal{L}_{w|v}[\rho_v] = \text{tr}_{e_v} (\mathcal{U}_{w|v}[\rho_v \otimes \eta_v]) , \quad (4.46)$$

where η_v is the state of the environment at v and $\mathcal{U}_{w|v}$ is the unitary dynamics between v and w . With this, we can define **oCP** divisibility:

¹⁹ Here, and in what follows, we shall use subscripts on states to denote the time at which they are prepared.

DEFINITION 4.3: A dynamics is operationally CP-divisible (oCP-divisible), if for any $u < v < w$

$$\mathcal{L}_{w|u} = \mathcal{L}_{w|v} \circ \mathcal{L}_{v|u} \quad (4.47)$$

holds, where the maps $\mathcal{L}_{y|x}$ are the maps defined in Eq. (4.46).

Importantly, complete positivity of the respective maps is guaranteed by design, as system-environment correlations are discarded for the reconstruction of $\mathcal{L}_{w|v}$. Here, it is the satisfaction of the composition law in Eq. (4.47) that has to be checked for oCP divisibility to hold, while in the case of iCP divisibility, the composition property is satisfied by design and complete positivity has to be checked for. Operational CP divisibility does – unlike iCP divisibility – not rely on the invertibility of $\mathcal{L}_{v|u}$, as all of the involved maps are reconstructed experimentally. Formally, Eq. (4.47) is the same as Eq. (4.44), but with the important distinction that here each map has a clear operational meaning in terms of a quantum process tomography procedure.

However, we still have left a level of ambiguity in the reconstruction process of the intermediate maps: Suppose at time u , we begin the experiment with either ρ_u or ρ'_u and then construct the dynamics from v to w via the procedure outlined above. In principle, the environment state at v could depend on the input state at u , and as such, the map $\mathcal{L}_{w|v}$ given by Eq. (4.47) could also depend on the input state at u . In order for oCP divisibility to be well-defined, the respective maps cannot depend on inputs at earlier times, which is an implicit assumption when formulating oCP divisibility. This requirement constitutes a conditional non-signaling condition [182–184], that we will use below to derive the structural properties of oCP-divisible processes. Here, it is important to note that oCP divisibility requires independence of the map $\mathcal{L}_{w|v}$ of the input state at time u .

This independence of the intermediate dynamics $\mathcal{L}_{w|v}$ from ρ_u is, for example, satisfied if the environment state is constant in time. Naturally, before we discuss the relation of CP divisibility and Markovianity, the question arises, if iCP divisibility and oCP divisibility are equivalent (that is, in the cases where iCP divisibility is well-defined). Somewhat surprisingly, this is not the case [1].

4.8.2 iCP Divisibility \neq oCP Divisibility

It is straightforward to see that oCP divisibility implies iCP divisibility: If $\mathcal{L}_{w|u} = \mathcal{L}_{w|v} \circ \mathcal{L}_{v|u}$ and $\mathcal{L}_{v|u}$ is invertible, then $\mathcal{K}_{w|v} = \mathcal{L}_{w|u} \circ \mathcal{L}_{v|u}^{-1} = \mathcal{L}_{w|v}$ is CPTP, as $\mathcal{L}_{w|v}$ is by construction CPTP (as can be seen from Eq. (4.46)). On the other hand, there are dynamics that are iCP-divisible, but not oCP-divisible. We show this directly by example.

EXAMPLE 4.1: To see that iCP divisibility \neq oCP divisibility, consider the two circuits in Fig. 4.13, where both the system and the environment are considered to be qubits, and let the initial environment state be maximally mixed, i.e., $\eta_u = \mathbb{1}_2/2$. The first unitary between u and v is a partial swap

$$U_{v|u} = \cos(\Omega\Delta t)\mathbb{1}_4 - i\sin(\Omega\Delta t)\mathbf{S}, \quad (4.48)$$

where $\mathbf{S} |ij\rangle = |ji\rangle$ is the swap operator between system and environment, $\mathbb{1}_4$ is the identity operator on both qubits, and $\Delta t := v - u$. The state of the system at v will be

$$\rho_v = \cos^2(\Omega\Delta t)\rho_u + \sin^2(\Omega\Delta t)\mathbb{1}_2/2, \quad (4.49)$$

and if $\Omega\Delta t \neq \pi/2$ then the corresponding dynamical map $\mathcal{L}_{v|u}$ is invertible. Next, let $U_{w|v}$ be a controlled-not gate with control on the environment, i.e.,

$$U_{w|v} |i_s j_e\rangle = |(i \oplus j)_s j_e\rangle, \quad (4.50)$$

where $\{|i_s\rangle\}_{i=0}^1$ and $\{|j_e\rangle\}_{j=0}^1$ are the computational basis of the system and the environment, respectively, and \oplus denotes addition modulo 2. In this case, it is easy to see, that without an intervention at time v , the final state of the system at time w is $\mathbb{1}_2/2$, independent of the input at u . Consequently, $\mathcal{K}_{w|v} = \mathcal{L}_{w|u} \circ \mathcal{L}_{v|u}^{-1}$ is the completely dephasing map, i.e., $\mathcal{K}_{w|v}[\rho_v] = \mathbb{1}_2/2$, which is a CPTP map, and the dynamics is thus iCP-divisible. However, if we discard the state of the system at v and insert a fresh state ρ_v , we will find that the corresponding state at time w will depend on ρ_u , the state that was inserted at time u . In other words we have conditional signaling, and therefore the process is not oCP-divisible.

Here, for better intuition, we have chosen an example for iCP-divisible dynamics that is not oCP-divisible, that is discrete in time. See App. B.2 for an example which is continuous in time.

While the requirements for oCP divisibility are potentially harder to check experimentally than those for iCP divisibility, as it necessitates intermediate interventions where fresh system states are prepared, oCP divisibility has the twofold advantage that the involved maps have an clear-cut operational meaning (unlike $\mathcal{K}_{w|v}$), and the property of oCP divisibility ties in effortlessly with the process tensor framework. This allows one to explicitly work out the memory effects that can and cannot be picked up by oCP divisibility (see below). Additionally, as it is a strictly stronger requirement than iCP divisibility, oCP divisibility (or rather the breakdown thereof) is a better witness of non-Markovian effects. However, as we shall see now, it does still not coincide with Markovianity.

4.8.3 Markovianity and oCP Divisibility

To discuss the relation of CP divisibility and Markovianity, we will restrict ourselves to three fixed times u , v , and w – a generalization to more times is always possible (albeit cumbersome) – and only discuss oCP divisibility, as it is the stricter requirement. Importantly, there is no *a priori* reason, why it should coincide with Markovianity.

To see this directly, we shall consider the following example presented in [1] and inspired by collision models [171–174, 185] that use initially correlated environment states to model memory effects [175, 186]: Let the initial system-environment state at time u be uncorrelated, and let the environment be in a correlated state $\eta_{e_u e_v}$. The *se* unitaries between times are such that they only lead to interactions between the system and *one* part of the environment (the part labeled by e_u between time u and v , and the part labeled by e_v between time v and w) which is discarded after the interaction (see Fig. 4.14).

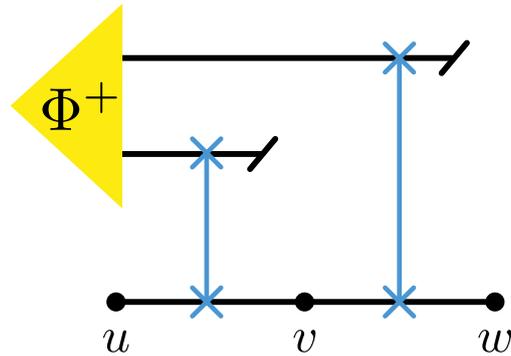


Figure 4.14: Non-Markovian oCP-divisible process. Independent of the input state at u , the system state at v is equal to the maximally mixed state. Analogously, if the system state is fed out at v , and a fresh system state is fed into the process, the corresponding system state at w is equal to the maximally mixed state, independent of the freshly prepared state. However, if the experimenter performs a measurement in the z -basis at time v , then the output state at w will be $|0\rangle$ ($|1\rangle$) for outcome 0 (outcome 1) independent of what state is fed forward. The future statistics does thus not only depend on the last state that was fed in, but earlier measurement outcomes, which implies that the process is non-Markovian.

It is easy to see, that the map $\mathcal{L}_{w|v}$ is well-defined, *i.e.*, it is independent of the system state that is fed into the process at time u , which implies that there is no signaling between times if we discard and freshly reprepare the system at time v . Now, if we choose the unitaries $\mathcal{U}_{v|u}$ and $\mathcal{U}_{w|v}$ to be swap operations, and the initial environment to be maximally entangled, then each of the operations $\mathcal{L}_{w|u}$, $\mathcal{L}_{w|v}$ and $\mathcal{L}_{v|u}$ coincides with the completely dephasing map, *i.e.*, $\mathbb{L}_{y|x} = \mathbb{1}_y^0 \otimes \mathbb{1}_x^1$, and we have $\mathcal{L}_{w|u} = \mathcal{L}_{w|v} \circ \mathcal{L}_{v|u}$, which means that the dynamics is oCP-divisible. It is not Markovian, though.

For example, if the system is a qubit, and the environment is initially in a maximally entangled two-qubit state, then performing a measurement in, say, the z -basis at time v that yields outcome $|0\rangle$ ($|1\rangle$) will leave the environment in state $|0\rangle$ ($|1\rangle$). Now, say the experimenter feeds forward a state ρ_v for either of these outcomes, then at time w they will obtain a system state $|0\rangle$ if the measurement outcome at v was 0, and $|1\rangle$ if the measurement outcome at v was 1. This dependence on past measurement outcomes implies that there are two different causal breaks with the same state that is fed forward, that lead to different future statistics. The process is thus non-Markovian (see Fig. 4.14 for more details).

In light of this example, we can give oCP divisibility a clear operational meaning: An oCP-divisible process is Markovian *on average*. If we *forget* about the measurement outcomes, then no memory effects can be detected. More generally, consider a multi-time process where an experimenter measures the system at each time, before independently preparing it in a new state. What oCP divisibility implies is that, if all past measurement outcomes are forgotten, then the future statistics only depend on the current preparation. Forgetting the measurement outcomes amounts to averaging over them, and is equivalent to discarding the system state before repreparation. A quantum Markov process, in contrast, requires that the future statistics only depend on the current preparation for *any* sequence of measurement outcomes [42, 50, 105,

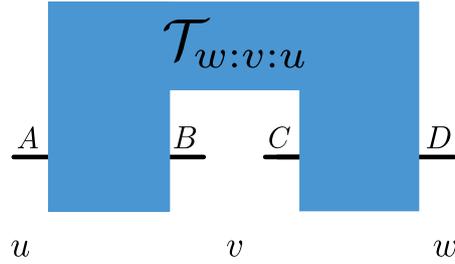


Figure 4.15: Descriptor $\mathcal{T}_{w:v:u}$ of a process defined on three times u , v , and w . For compact notation, the wires are denoted by A, B, C , and D .

147]. For a graphical representation of the relation of Markovianity and iCP- and oCP-divisible processes,²⁰ see Fig. 4.16.

4.8.4 Temporal Correlations in oCP-Divisible Processes

Unlike for iCP divisibility, the memory effects that persist when the dynamics is oCP-divisible can be quantified explicitly. To do so, we consider the one-slot process tensor depicted in Fig. 4.15 with an open input wire (corresponding to time u), and an open output wire (corresponding to time w) and the middle slot corresponding to time v . This process tensor $\mathcal{T}_{w:v:u}$ – where we choose the subscripts to emphasize what times it is defined on – could be the descriptor of a stochastic process that is interrogated at three times. For compact notation, we denote the wires by A (corresponding to time u), B, C (corresponding to time v , and D (corresponding to time w), *i.e.*, $\mathcal{T}_{w:v:u} \in \mathcal{B}(\mathcal{H}_D \otimes \mathcal{H}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A)$. If the process was Markovian, then we would have

$$\mathbb{T}_{w:v:u}^{\text{Markov}} = \mathbb{L}_{w|v} \otimes \mathbb{L}_{v|u}, \quad (4.51)$$

where $\mathbb{L}_{w|v} = \frac{1}{d_A} \text{tr}_{AB}(\mathbb{T}_{w:v:u}^{\text{Markov}})$ and $\mathbb{L}_{v|u} = \frac{1}{d_C} \text{tr}_{CD}(\mathbb{T}_{w:v:u}^{\text{Markov}})$ are CPTP maps between times v and w and between times u and v , respectively. They are deliberately denoted by the letter \mathbb{L} , as they correspond exactly to maps that one would obtain from the reconstruction procedure outlined for the definition of oCP divisibility. It is important, though, to emphasize the difference between Eq. (4.51) and Eq. (4.47); the former pins down a structural requirement for $\mathbb{T}_{w:v:u}$, the complete descriptor of the dynamics on times $\{u, v, w\}$, while the latter constitutes a weaker requirement on particular two-time descriptors of the process. Consequently, there are – as we have already seen – processes that satisfy Eq. (4.47) but are not Markovian, and do thus not satisfy Eq. (4.51).

To express the property of oCP divisibility in terms of process tensors, we first phrase the conditional non-signaling condition we encountered above in this language. Here, non-signaling means, that the output state that is measured at time w cannot depend on the state that was fed into the process at time u , if the system was discarded and freshly prepared at time v . Superficially, this non-signaling condition is reminiscent of the *no information back-flow* from

²⁰ An in-depth discussion and visualization of the *full*, multifaceted hierarchy of notions pertaining to (quantum) non-Markovianity can be found in Ref. [66]. Where exactly iCP and oCP divisibility fit in this general hierarchy is, *a priori*, unclear and left as an open question for future work.

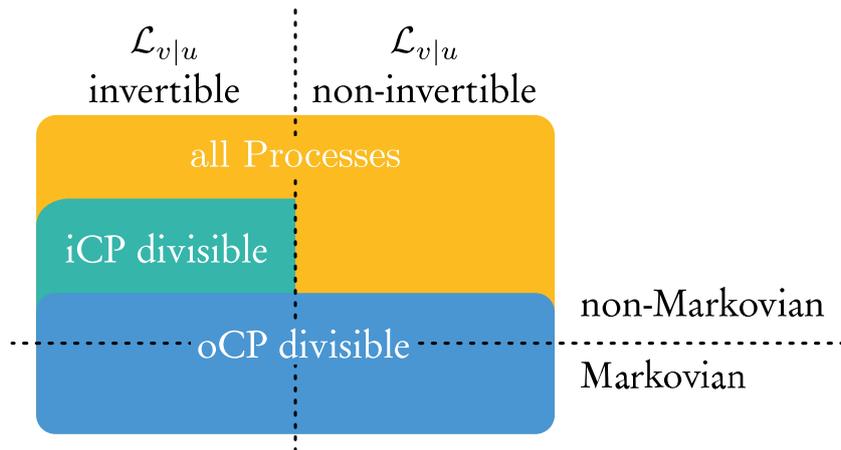


Figure 4.16: Hierarchy of processes. All Markovian processes are iCP and oCP-divisible, while the set of iCP-divisible processes – in the case where it is well-defined – contains both the set of oCP-divisible as well as that of Markovian processes. The inclusion is strict, *i.e.*, there are, for example, iCP processes that are not oCP, and none of the sets coincide. Notably, here, only the properties discussed in the text are shown. In general, the space of non-Markovian processes presents itself as much more layered (see [66] and, in particular, Fig. 1 therein for a more complete picture.).

the environment to the system that is attributed to CP-divisible processes [166]. However, in our case, signaling is a genuinely three-time statement that quantifies how much information about a preparation at time u is retained at time w given that the system state was discarded at time v . On the other hand, the increase of trace distance between to states over some period of time that is interpreted as an information backflow is a genuine two-time statement.

Expressed in terms of Choi states, the conditional non-signaling requirement for a process tensor $\mathbb{T}_{w:v:u}$ relevant for oCP divisibility reads

$$\mathrm{tr}_{ABC} \left[\left(\mathbb{1}_{DB} \otimes \rho_u^T \otimes \rho_v^T \right) \mathbb{T}_{w:v:u} \right] = \mathrm{tr}_{ABC} \left[\left(\mathbb{1}_{DB} \otimes \rho_u^T \otimes \rho_v^T \right) \mathbb{T}_{w:v:u} \right] \quad (4.52)$$

for all $\rho_u, \rho'_u \in \mathcal{B}(\mathcal{H}_A)$ and all $\rho_v \in \mathcal{B}(\mathcal{H}_C)$, where we have used the fact that discarding the system at time v corresponds to the effect $\mathbb{1}_B$, and the labeling of Hilbert spaces follows Fig. 4.17. In order for this equation to be satisfied for all possible input states at time u , we must have

$$\mathrm{tr}_B (\mathbb{T}_{w:v:u}) = \mathbb{1}_A \otimes \mathbb{L}_{w|v}, \quad (4.53)$$

which, unsurprisingly, is reminiscent of the causality constraints for quantum combs we encountered in Ch. 2 (see Fig. 4.17 for a graphical illustration). There, the corresponding trace constraints guaranteed that overall, no signaling from the future to the past was possible. Here, Eq. (4.53) ensures, that there is no signaling from time u to time w given that the system was discarded at time v . In this sense, as we alluded to above, a dynamics that satisfies Eq. (4.53) and is – in addition to that²¹– oCP-divisible, can be considered as a process that is Markovian *on average*; while there can still be individual causal breaks that allow one to detect memory effects, ‘averaged out’ causal breaks of the form $\rho_v \otimes \mathbb{1}_B$ do not allow one to detect the non-Markovianity of the process; put differently, the operations $\rho_v \otimes \mathbb{1}_B$ can fail as witnesses for the non-product form of $\mathbb{T}_{w:v:u}$.

²¹ For an example of dynamics that satisfy conditional non-signaling, but are not oCP-divisible, see App. B.3.

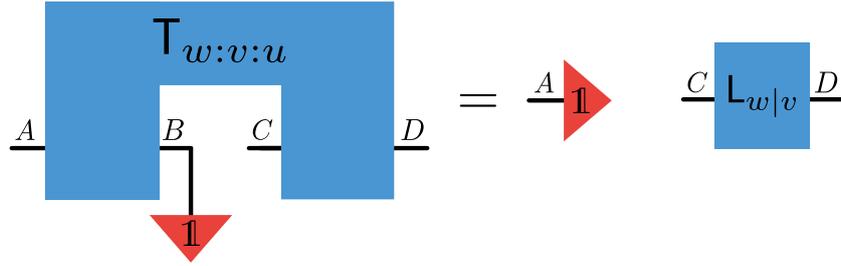


Figure 4.17: Non-signaling requirement for oCP divisible dynamics. Discarding the system at time v (corresponding to wires B and C) disconnects the output at time w (corresponding to wire D) from possible inputs at time u (corresponding to wire A).

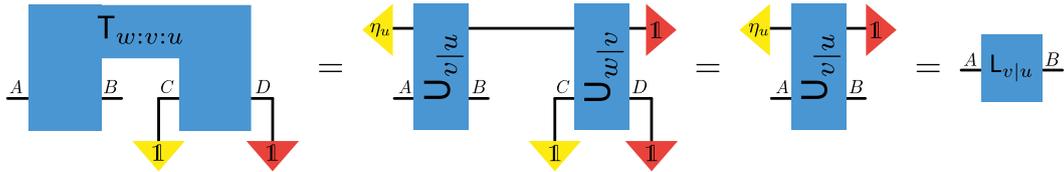


Figure 4.18: Graphical representation of $L_{v|u}$. Tracing out the final state at time w and feeding a unit trace state (here, the maximally mixed state, represented by $\mathbb{1}$) at time v leads to a trace on the environment, due to the trace condition of the unitary map $U_{w|v}$. The channel from u to v exactly corresponds to $\mathbb{1}_e \star U_{v|u} \star \eta_u$.

Finally, we can express oCP divisibility as a constraint on $T_{w:v:u}$. Naturally, we obtain the mapping $\mathcal{L}_{w|u}$ by ‘implementing’ a do-nothing channel at time v , *i.e.*,

$$L_{w|u} = \text{tr}_{BC} \left[\left(\mathbb{1}_{AD} \otimes \Phi_{BC}^{+T} \right) T_{w:v:u} \right] = \langle \Phi_{BC}^+ | T_{w:v:u} | \Phi_{BC}^+ \rangle. \quad (4.54)$$

On the other hand, using Eq. (4.53), we can compute $L_{w|v}$ as

$$L_{w|v} = \frac{1}{d_A} \text{tr}_{AB} (T_{w:v:u}). \quad (4.55)$$

Importantly, due to signaling effects, for general processes, one could not simply obtain intermediate channels by tracing out over subsystems that correspond to earlier time steps, and the channel $L_{w|v}$ is only well-defined and independent of past preparations due to the non-signaling condition (4.53). In a similar way, we can obtain $L_{v|u}$ via

$$L_{v|u} = \frac{1}{d_C} \text{tr}_{CD} (T_{w:v:u}), \quad (4.56)$$

where $L_{v|u}$ would be well-defined even in the presence of signaling, due to the causality constraints of $T_{w:v:u}$ (see Fig. 4.18). With this, we see that a process is oCP-divisible on times $\{u, v, w\}$, iff its process tensor $T_{w:v:u}$ satisfies Eq. (4.53) and

$$\langle \Phi_{BC}^+ | T_{w:v:u} | \Phi_{BC}^+ \rangle = \frac{1}{d_C} \text{tr}_{CD} (T_{w:v:u}) \otimes \frac{1}{d_A} \text{tr}_{AB} (T_{w:v:u}) \quad (4.57)$$

Evidently, this requirement does not force $T_{w:v:u}$ to be of product form. We can see this explicitly by representing $T_{w:v:u}$ as

$$T_{w:v:u} = \frac{1}{d_C} \text{tr}_{CD} (T_{w:v:u}) \otimes \frac{1}{d_A} \text{tr}_{AB} (T_{w:v:u}) + \chi = L_{w|v} \otimes L_{v|u} + \chi, \quad (4.58)$$

where $\chi \in \mathcal{B}(\mathcal{H}_D \otimes \mathcal{H}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A)$ contains all the temporal correlations of $\mathsf{T}_{w:v:u}$ that corresponds to deviations from Markovian behavior. It satisfies $\text{tr}_{CD}(\chi) = \text{tr}_{AB}(\chi) = 0$. With this, we see that non-Markovian correlations that can be present in a oCP-divisible process satisfy

$$\langle \Phi_{BC}^+ | \chi | \Phi_{BC}^+ \rangle = 0. \quad (4.59)$$

As this does *not* imply $\chi = 0$, this relation comprehensively quantifies the set of temporal correlations that cannot be detected by the criterion of oCP divisibility, and allows one to check if a given processes is oCP-divisible. For instance, for the example of Fig. 4.14 discussed above, it is easy to see, that the corresponding process tensor $\mathsf{T}_{w:v:u}$ is given by

$$\mathsf{T}_{w:v:u} = \frac{1}{d_B} \mathbb{1}_{AC} \otimes \Phi_{BD}^+, \quad (4.60)$$

where Φ_{BD}^+ is the unnormalized maximally entangled two-qubit state. While this process tensor is not Markovian, as it does not display a product structure between AB and CD , it satisfies Eq. (4.59); we have

$$\mathsf{T}_{w:v:u} = \frac{1}{2} \mathbb{1}_{AB} \otimes \frac{1}{2} \mathbb{1}_{CD} + \frac{1}{4} \mathbb{1}_{AC} \otimes (\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z), \quad (4.61)$$

where the correlation term $\chi = \frac{1}{4} \mathbb{1}_{AC} \otimes (\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z)$ satisfies $\text{tr}_{CD}(\chi) = \text{tr}_{AB}(\chi) = 0$ and $\langle \Phi_{BC}^+ | \chi | \Phi_{BC}^+ \rangle = 0$, but is not equal to zero.

Defining CP divisibility such that its constituents possess an operationally clear-cut interpretation thus allows one to quantify how well it works as a witness for non-Markovianity. If no non-Markovianity can be detected via measurements that check for oCP divisibility, only deviations from Markovianity of the form (4.59) can exist. In principle, one could carry out this investigation in a more fine tuned manner by decomposing χ in terms of two-time, three-time, and four-time correlations (see, for example [8]). Analogously, the above statements can be extended to more times by defining a χ term that encapsulates the deviation of a multi-step process $\mathcal{T}_{z:y:x:\dots}$ from a Markovian process. Instead of one Eq. (4.59) one would then obtain a set of equations for different combinations of time steps, where identity operations are inserted.

However, this discussion would prove to be quite cumbersome, without the benefit of providing further intuitive insight into the nature of oCP divisible processes. Rather, we shall now round off our discussion of the memory structure of general open processes, and briefly touch upon the generalization of Markov order to the quantum case.

4.9 QUANTUM MARKOV ORDER

Up to this point, we have discussed the generalization of the definition of Markovian processes to the quantum regime. Quite naturally, such a generalization also exists for the concept of *Markov order*, which we shall briefly discuss in this section. For a thorough introduction to, as well as a comprehensive taxonomy of quantum processes with finite Markov order, see Refs. [5, 6, 187].

As already mentioned in Sec. 4.6, a classical process with Markov order ℓ is a process that is conditionally independent of the past, given the last ℓ outcomes, *i.e.*

$$\mathbb{P}(x_{N+1}|x_N, \dots, x_1) = \mathbb{P}(x_{N+1}|x_N, \dots, x_{N-\ell+1}) \quad \forall N, \quad (4.62)$$

and all multi-time joint probability distributions can be built up, once $\mathbb{P}(x_\ell, \dots, x_1)$ and the transition probabilities $\mathbb{P}(x_\alpha|x_{\alpha-1}, \dots, x_{\alpha-\ell})$ are known. To simplify notation, following Ref. [6], we will group the times the process is defined on into future $F = \{t_{N+1}, \dots, t_{\alpha+1}\}$, memory $M = \{t_\alpha, \dots, t_{\alpha-\ell}\}$, and history $H = \{t_{\alpha-\ell-1}, \dots, t_1\}$, with the corresponding sequences of outcomes and joint probability distribution denoted by x_F, x_M , and x_H , and \mathbb{P}_{FMH} , respectively. With this convention, Eq. (4.62) reads

$$\mathbb{P}_{FMH}(x_F|x_M, x_H) = \mathbb{P}_{FMH}(x_F|x_M). \quad (4.63)$$

Importantly, as for the case of Markovian dynamics, the independence between the future and the history is a *conditional* one, *i.e.*, for a process with Markov order ℓ , once the last ℓ outcomes are revealed, knowledge of any earlier outcomes does not provide any additional advantage when predicting future statistics. However, if one did not know the last ℓ outcomes, knowledge about the prior past would, in general, improve predictions about future outcomes.

Now, with the framework of process tensors at hand, a natural generalization of Eq. (4.63) suggests itself: We could imagine an experimenter that uses (possibly temporally correlated) instruments $\mathcal{J}_F, \mathcal{J}_M$, and \mathcal{J}_H to probe the system of interest at times in F, M , and H , respectively. They would consider the process to be of (quantum) Markov order $\ell = |M|$ if, for any choice of \mathcal{J}_F and \mathcal{J}_H , the statistics at F were independent of the statistics at H , *given* the outcomes at M . We thus arrive at the definition of (quantum) Markov order [6]:

DEFINITION 4.4 ((QUANTUM) MARKOV ORDER): *A process is said to be of (quantum) Markov order $|M| = \ell$ with respect to an instrument \mathcal{J}_M , if for all possible instruments $\{\mathcal{J}_F, \mathcal{J}_H\}$ the relation*

$$\mathbb{P}_{FMH}(x_F, \mathcal{J}_F|x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) = \mathbb{P}_{FMH}(x_F, \mathcal{J}_F|x_M, \mathcal{J}_M) \quad (4.64)$$

is satisfied, where, $\mathbb{P}_{FMH}(x_F, \mathcal{J}_F|x_M, \mathcal{J}_M; x_H, \mathcal{J}_H)$ denotes the probability to measure outcome x_F , given that the outcomes x_M and x_H were obtained at M and H , and the instruments $\mathcal{J}_F, \mathcal{J}_M$, and \mathcal{J}_H were used to probe the system.

Importantly, the above definition is with respect to a *fixed* instrument \mathcal{J}_M . Demanding for it to hold for all possible instruments at M , is too strict a requirement [6]. This fact notwithstanding, a process could also have finite quantum Markov order with respect to a whole family of instruments [6].

Before we analyze the basic structural properties of process tensors that satisfy Eq. (4.64), a plausibility check of this definition is in order. Firstly, Markovian processes as defined in the previous section, are processes of Markov order $|M| = 1$, where the instrument \mathcal{J}_M consists of causal breaks only. In the Markovian case, *any* instrument that only consists of causal breaks, would satisfy Def. 4.4, and as such, a Markovian process has Markov order 1 with respect to a whole family of instruments. Secondly, if all instruments correspond to sharp (*i.e.*, projective) measurements (or sequences thereof) in a fixed classical basis, then it is easy to see that Def. 4.4

coincides with the classical definition [6]. Consequently, this definition of (quantum) Markov order fits in nicely with existing classical definitions, as well as the definition of Markovianity in the quantum case that we introduced above.

Evidently, Eq. (4.64) imposes structural restrictions on a process tensor of Markov order ℓ with respect to the instrument \mathcal{J}_M . To see this, let the corresponding process tensor be T_{FMH} , and let the instruments be $\mathcal{J}_X = \{\mathsf{M}_{x_X}\}$. Then, the outcome probabilities can be written as

$$\mathbb{P}_{FMH}(x_F, \mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) = \mathsf{T}_{FMH} \star \mathsf{M}_{x_F} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}, \quad (4.65)$$

while the respective conditional probabilities are given by

$$\begin{aligned} & \mathbb{P}_{FMH}(x_F, \mathcal{J}_F | x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) \\ &= (\mathsf{T}_{FMH} \star \mathsf{M}_{x_F} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}) / (\mathsf{T}_{MH} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}), \end{aligned} \quad (4.66)$$

where $\mathsf{T}_{MH} = \frac{1}{d_{F^1}} \text{tr}_F(\mathsf{T}_{FMH})$ and d_{F^1} is the dimension of all the input spaces that correspond to times in F .²² In order for Eq. (4.64) to be satisfied, we must have

$$\begin{aligned} & \frac{\mathsf{T}_{FMH} \star \mathsf{M}_{x_F} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}}{\mathsf{T}_{MH} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}} \\ &= \frac{\sum_{x_H} \mathsf{T}_{FMH} \star \mathsf{M}_{x_F} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}}{\sum_{x_H} \mathsf{T}_{MH} \star \mathsf{M}_{x_M} \star \mathsf{M}_{x_H}} := \frac{\mathsf{T}_{FMH} \star \mathsf{M}_{x_F} \star \mathsf{M}_{x_M} \star \mathsf{M}_H}{\mathsf{T}_{MH} \star \mathsf{M}_{x_M} \star \mathsf{M}_H}, \end{aligned} \quad (4.67)$$

where we have defined the overall CPTP map²³ M_H of the instrument \mathcal{J}_H . Evidently, for Eq. (4.67) to hold for *all* instruments $\{\mathcal{J}_F, \mathcal{J}_H\}$, and all CP maps M_{x_M} belonging to the fixed instrument \mathcal{J}_M , it is sufficient that the action of M_{x_M} leaves the process tensor in a product, *i.e.*,

$$\mathsf{T}_{FMH} \star \mathsf{M}_{x_M} = \mathsf{T}_F^{(x_M)} \otimes \tilde{\mathsf{T}}_H^{(x_M)}, \quad (4.68)$$

where $\mathsf{T}_F^{(x_M)}$ is a causally ordered process tensor on F , and the set $\{\tilde{\mathsf{T}}_H^{(x_M)}\}$ forms a tester on H , *i.e.*, $\tilde{\mathsf{T}}_H^{(x_M)} \geq 0$ and $\mathsf{T}_H^{(\mathcal{J}_M)} = \sum_{x_M} \tilde{\mathsf{T}}_H^{(x_M)}$ is a deterministic comb. Indeed, plugging Eq. (4.68) into Eq. (4.67), we obtain (see also App. B.4 for more details)

$$\mathbb{P}(x_F, \mathcal{J}_{x_F} | x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) = \mathsf{T}_F^{(x_M)} \star \mathsf{M}_{x_F} = \mathbb{P}(x_F, \mathcal{J}_{x_F} | x_M, \mathcal{J}_M) \quad (4.69)$$

Intuitively, the structure of Eq. (4.68) exactly mirrors the classical definition of Markov order; once the outcomes on the memory block M are given, the future and past are independent, which is reflected by the product structure of the RHS of Eq. (4.68) (see also Fig. 4.19). The fact that the resulting ‘process’ $\tilde{\mathsf{T}}_H^{(x_M)}$ is not necessarily a causally ordered process, but only a tester element, might seem counterintuitive at first sight. However, it arises from conditioning on an outcome on the memory block, and as we have seen in Ch. 2, conditioning can lead

²² Due to the causal ordering of T_{FMH} , contraction with *any* deterministic comb yields the *same* process on M and F , *i.e.*, $\mathsf{T}_{FMH} \star \mathsf{M}_F = \mathsf{T}_{MH}$ for every deterministic comb M_F . The definition $\mathsf{T}_{MH} = \frac{1}{d_{F^1}} \text{tr}_F(\mathsf{T}_{FMH})$ then follows from the choice $\mathsf{M}_F = \mathbb{1}_F / d_{F^1}$.

²³ If the respective instrument extends over several times, then M_H is a deterministic comb, and its elements are tester elements. As this distinction does not really affect the following results, we shall often continue to call M_H a CPTP map, and \mathcal{J}_X an instrument (instead of a tester).

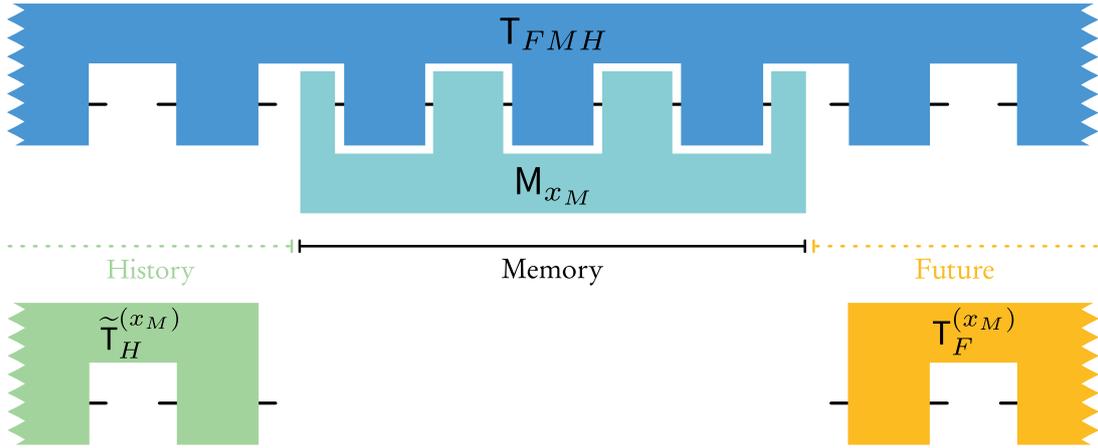


Figure 4.19: Process with finite Markov order. Every outcome x_M of the fixed instrument \mathcal{J}_M ‘cuts’ the process tensor into two independent parts. Consequently, the future is conditionally independent on the history, given that instrument \mathcal{J}_M was used on the memory block.

to processes that are non-deterministic. We will return to this point in Ch. 6, where we use conditioning to simulate causally disordered processes.

In Ref. [6], it has been shown, that requirement (4.68) is not just sufficient but also necessary for a process to be of Markov order $\ell = |M|$, which leads to the following theorem:

THEOREM 4.2: *A process has Markov order $\ell = |M|$ with respect to an instrument $\mathcal{J}_M = \{M_{x_M}\}$ iff its process tensor satisfies*

$$T_{FMH} \star M_{x_M} = T_F^{(x_M)} \otimes \tilde{T}_H^{(x_M)}, \quad (4.70)$$

where all the terms of the equality are as defined above.

For comprehensiveness, the proof of this structural statement is reproduced in App. B.4. Its graphical representation can be found in Fig. 4.19. Importantly, even though each of the outcomes x_M occurs *probabilistically*, the instrument \mathcal{J}_M blocks the memory *deterministically*; no matter what outcome the experimenter receives at M (when using the instrument \mathcal{J}_M), the corresponding statistics of future and history are independent of each other. Put differently, in *every* run of the experiment, the future and the past are independent if the memory blocking instrument \mathcal{J}_M is used at times in M .

Using Thm. 4.2, one can straightforwardly deduce the explicit structure of processes that have finite Markov order with respect to an instrument, and investigate the wealth of different memory structures that exist for such processes. Here, we shall refrain from providing this (somewhat technical) analysis, and refer the reader to the comprehensive investigation in Refs. [5, 6, 187].

Importantly, the process tensor framework, and the operationally clear-cut understanding of memory effects it provides, allows one to generalize key concepts from the study of classical stochastic processes, and enables the comprehensive study of temporal correlations in both classical and quantum physics. Indeed, even the classical definition of Markov order is an *ante litteram* example of an instrument dependent definition of Markov order, as it is phrased with the implicit assumption that all of the employed instruments only consist of projective measurements in the classical basis. Allowing for active interventions – as one can in classical

physics, and must in quantum mechanics – makes the instrument dependence of Markov order obvious, and provides a rich tapestry of classical and quantum processes with finite Markov order [5].

4.10 OPEN QUANTUM PROCESSES – SUMMARY

In this chapter, we have shed light on quantum stochastic processes from a physically motivated direction, that of open quantum system dynamics, where the system an experimenter can manipulate, is coupled in an uncontrollable way to an unknown environment. Trying to overcome the problem of initial system-environment correlations, and implementing multi-time statistics into the framework, one is forced to switch to an operational perspective. Here, instead of system states, as in the traditional approach, experimentally controllable parameters are considered input to the process, while uncontrollable parts of the dynamics are considered the process itself.

In a sense, this is a consequent continuation of the program of quantum channels, that arise from encapsulating all uncontrollable parts of the dynamics in a map that maps the inputs (quantum states of the system) to outputs (quantum states of the system). Following this path, higher order quantum maps arise naturally as the descriptor of open quantum system dynamics, and provide a powerful tool for the investigation of memory properties. In a sense, the seemingly long detour we took when discussing higher order maps, investigating their properties, and laying out their use for the description of general quantum stochastic processes, led us right to the heart of the theory of open quantum system dynamics, and the concepts we worked out enable the characterization of different kinds of dynamics and their complexity. As such, we have arrived at a comprehensive and cohesive picture of higher order quantum maps, quantum stochastic processes, and open quantum processes.

Having a handle on the memory structures of open processes, as provided by the process tensor approach, is not simply of theoretical appeal; building near-term quantum technologies will require effective methods for addressing non-Markovian noise [176]. Here, we discussed the memory properties of processes, in particular the structure of Markovian processes and processes with finite Markov order. The concepts of Markovianity and Markov order are first and foremost logical statements about conditional independence of the future from the past, given a set of previous outcomes. Using the operational approach, we extended these logical concepts to definitions in the quantum case, that can be expressed as structural requirements of the corresponding process tensor, and can be probed for experimentally. In turn, these structural consequences of the logical definitions we provided, allow for a clear delineation of the memory effects that can and cannot be detected by witnesses of non-Markovian effects.

A natural check for non-Markovianity is to see if a process is indivisible. We have shed light on CP divisibility from an operational point of view, which helps us to understand and identify the classes of temporal correlations that may evade such a check. However, there are trade-offs between uncovering temporal correlations and the requisite number of experiments that must be performed. Our results enable experimenters to make informed decisions about investing resources in classifying the non-Markovian noise at hand. Furthermore, they allow for quanti-

tative and tangible assertions about the amount and the type of temporal correlations present in an experiment. As such, they provide the starting point for an operational understanding of existing witnesses of non-Markovianity, that are predominantly based on the breakdown of CP divisibility.

Given all of these advantages of the process tensor description, it would be ideal to have the experimental means to fully reconstruct them in the laboratory. However, there are limitations to what an experimenter can do in practice. On the one hand, as we discussed, the number of necessary experiments for reconstruction grows exponentially with the number of time steps. While process tensors with one slot have been reconstructed in the laboratory [145] reconstruction for more times proves prohibitively cumbersome. On the other hand, even if the number of necessary experiments was manageable, the reconstruction of a process tensor would still require for the experimenter to have full control over the system of interest. In many experimental situations, such a level of control is out of reach. Both of these problems pose limits on the usability of the process tensor approach in practice. While efficient tomography is an interesting subject in its own right [188, 189] we will put this question aside in the following chapter, and focus on the second problem of process tensor reconstruction by means of limited resources.

PROCESS TENSORS AND LIMITED RESOURCES

The process tensor framework overcomes the problems that traditional approaches to the description of open quantum system dynamics are plagued by. However, in order to reconstruct it experimentally, *i.e.*, in order to collect all accessible information about the underlying dynamics, it is necessary to be able to perform manipulations of the system that constitute a basis of all physical operations; that is, a basis of the set of (trace non-increasing) CP maps acting on the system of interest. This requirement of full local control is not met in most real-world experiments.

Generally, experimenters are only interested in the effects of a restricted set of manipulations, or the set of available manipulations is fundamentally limited by the experimental setup. Well-known examples are ubiquitous in the field of quantum control [149], where, *e.g.*, shaped laser pulses are used to control the dynamics of molecules [150] or partial measurements are employed to steer the system of interest to a desired final state [151]. Other prominent examples include dynamical decoupling experiments [152], and experimental setups in quantum optics, where beam splitters and phase shifters can be used to implement arbitrary (single-photon) unitary gates [190], but no non-unitary operations.

A very recent concrete and important example of such a scenario is provided by the publicly available quantum computer by IBM [191]. This computer only allows for a sequence of one and two-qubit unitary gates followed by a final measurement on each qubit; *a priori*, it is not obvious how to experimentally reconstruct a process tensor for this scenario, *i.e.*, a map that yields the correct output states for any sequence of unitary input operations. Neither is it clear, what the operational meaning and the mathematical properties of such an object would be.

Understanding processes when only limited control is available plays, for instance, an important role in building a model of (non-Markovian) errors and error correction in the non-Markovian regime [192–194]. Intuitively, measuring the output states corresponding to only a limited set of sequences of control operations still provides some information about the dynamics of the underlying dynamics. Naturally, given their experimental relevance, we aim to answer two questions in this chapter: How can we reconstruct meaningful process tensors for *any* level of experimental control, and what can we learn from the resulting reconstructed objects?

All of the experimental procedures listed above can readily and meaningfully be cast in the language of the process tensor formalism as mappings from sequences of experimental operations to a final state when the experiment has concluded [2]. However, neither the influence of laser pulses, nor partial measurements, nor unitary operations, when taken as

distinct sets, constitute a basis of the set of all possible manipulations of the system of interest. To make matters worse, these operations do not even constitute a convex space, let alone a linear vector space. It is therefore unclear if a meaningful process tensor can be experimentally reconstructed from this limited set of accessible manipulations.

Moreover, on top of the mere experimental reconstruction, we can ask what can be inferred about correlations and memory effects based on the resulting *restricted* process tensors. As we have seen, the complete process tensor allows one to determine whether there are detectable correlations between the system and its environment, and/or memory effects in the dynamics. At first sight, it is not obvious to what extent similar assertions can be made based on only a limited set of available experimental operations.

Slightly rephrased, the overarching theme of this chapter is to investigate: what is the maximal information that can be inferred about a process based on a limited set of experimental resources? Thus, we shall bring the process tensor framework closer to experimental realities in this chapter. The corresponding discussion and results are based on Ref. [2].

Similar questions considering the reconstruction of quantum channels based on incomplete information have been investigated (and answered) in [195, 196], while Ref. [144] examined the superchannel description of a process, when only projective measurements on the system can be performed. The corresponding descriptor is well-defined defined on the set of projective preparations and we will see that it is a particular case of the restricted process tensors we introduce in this chapter.

Finally, besides the experimental appeal, employing the concepts introduced throughout this chapter, we will be able to answer a question we encountered in Ch. 3: How can we characterize the set of processes that appear classical to an experimenter with only classical means of measurement? To do so, we shall use the breakdown of Kolmogorov conditions as the defining signature of non-classicality and use the framework of restricted process tensors to provide a full characterization of classical multi-step processes. Thus, we extend the corresponding characterization of Markovian dynamics given in Ref. [80] to the general, non-Markovian case, where memory effects play a non-negligible role.

We shall begin this chapter by tailoring the experimental reconstruction of process tensors to the case of limited control, and furnishing the resulting experimental procedure with concrete examples.

5.1 RESTRICTED TOMOGRAPHIC RECONSTRUCTION

The process tensor $\mathbb{T}_{N+1:1}$, with its action defined on control operations at times $\{t_1, \dots, t_N\}$ provides the full description of an open system's dynamics, *i.e.*, it maps every possible sequence of operations to the correct output state at t_{N+1} . We briefly reiterate its experimental reconstruction to motivate its extension to the case of limited control.

In Ch. 4 we have seen that it can be reconstructed by measuring the output states $\rho^{(\vec{j})'}$ at t_{N+1} corresponding to a basis $\{\mathbb{M}_{\vec{j}}\}$ of sequences of operations at times $\{t_1, \dots, t_N\}$:

$$\mathbb{T}_{N+1:1} = \sum_{\vec{j}} \rho^{(\vec{j})'} \otimes \mathbb{m}_{\vec{j}}^*, \quad (5.1)$$

where $\{\mathfrak{m}_{\vec{\gamma}}\}$ is the dual set to $\{\mathfrak{M}_{\vec{\gamma}}\}$. If the basis elements $\{\mathfrak{M}_{\vec{\gamma}}\}$ consist of independent operations, then we have

$$\mathfrak{M}_{\vec{\gamma}} = \mathfrak{M}_{j_N} \otimes \cdots \otimes \mathfrak{M}_{j_1}, \quad (5.2)$$

and, analogously, the corresponding dual set is of the form

$$\mathfrak{m}_{\vec{\gamma}} = \mathfrak{m}_{j_N} \otimes \cdots \otimes \mathfrak{m}_{j_1}. \quad (5.3)$$

In what follows, we will consider the operations that can be implemented at different times to be independent. A generalization to temporally correlated controls is always straightforwardly possible.

Evidently, Eq. (5.1) only yields the full process tensor (with all of its appealing mathematical properties), if the output states $\rho^{(\vec{j})}$ at t_{N+1} are recorded for a complete basis of sequences of control operations. This fact notwithstanding, by assuming a distinctly operational stance, meaningful statements can be made about a process based on limited control: An experiment is fully described if one can predict the correct output state for any experimentally performable sequence of operations. We shall employ this point of view in conjunction with Eq. (5.1), to derive operationally meaningful descriptors of situations with restricted resources.

To do so, let us denote the set of experimentally available operations that can be performed by the experimenter by R . For example, R could be the set of unitary operations acting on the system, or a set of operations generated by a finite control algebra [152]. Now, this restricted set of operations spans a vector space that we denote by $W := \text{Span}(R)$, *i.e.*, the vector space W consists of linear combinations of elements of R . As we alluded to above, a proper descriptor of an experimental procedure where at each time t_α only the set R of operations can be performed, must map every available input operation sequence to the correct output state. The reconstruction of this descriptor, which we will henceforth call a *restricted process tensor* follows directly from Eq. (5.1): It can be reconstructed by measuring the output states corresponding to a *basis* of performable operations, which, in this case, is the basis of the space W that R spans.¹

In detail, there exists a set $\{\mathfrak{R}_{\vec{\gamma}}\}_{\vec{\gamma}=1}^{d_w} \subset R$ that constitutes a basis of W , where d_w is the dimension of W . Following Eq. (5.1), the restricted process tensor $\mathbb{T}_{N+1:1}^R$ is then given by

$$\mathbb{T}_{N+1:1}^R = \sum_{\vec{\gamma}} \rho'_{\vec{\gamma}} \otimes r_{\vec{\gamma}}^*, \quad (5.4)$$

where $\rho'_{\vec{\gamma}}$ is the output state corresponding to the sequence of operations $\mathfrak{R}_{\vec{\gamma}} = \mathfrak{R}_{\gamma_N} \otimes \cdots \otimes \mathfrak{R}_{\gamma_1}$ and $\{r_{\vec{\gamma}}\}$ is the dual set to $\{\mathfrak{R}_{\vec{\gamma}}\}$ [2]. Due to linearity, the resulting restricted process tensor $\mathbb{T}_{N+1:1}^R$ yields the correct output state for any (possibly temporally correlated) admissible sequence $\mathfrak{R}_{N:1}$ that lies in $W^{\otimes N}$; any such sequence can be written as a real linear combination $\mathfrak{R}_{N:1} = \sum_{\vec{\gamma}} c_{\vec{\gamma}} \mathfrak{R}_{\vec{\gamma}}$ of basis elements $\mathfrak{R}_{\vec{\gamma}}$, with $c_{\vec{\gamma}} \in \mathbb{R}$, and hence we have

$$\mathcal{T}_{N+1:1}^R[\mathfrak{R}_{N:1}] = \mathbb{T}_{N+1:1}^R \star \mathfrak{R}_{N:1} = \sum_{\vec{\gamma}, \vec{\omega}} c_{\vec{\gamma}} \rho'_{\vec{\omega}} \text{tr}(r_{\vec{\omega}}^\dagger \mathfrak{R}_{\vec{\gamma}}) = \sum_{\vec{\gamma}} c_{\vec{\gamma}} \rho'_{\vec{\gamma}},$$

¹ We shall assume throughout this chapter that, at the end of the experiment, full state tomography can be performed. However, the respective arguments would not change if at t_{N+1} only a set of measurements that is not IC can be performed. In this case, the respective outputs that we use in the derivation of restricted process tensors have to be adjusted accordingly.

which means that the operator $\mathbb{T}_{N+1:1}^R$ reconstructed via Eq. (5.4) maps every experimentally feasible operation in $W^{\otimes N}$ to the correct output state. Importantly, in general, W contains a larger set of CP maps than R . For example, the space spanned by the set of unitary maps contains all unital maps (see Sec. 5.2.1 below), and a restricted process tensor reconstructed from a set of unitary operations can predict the output state for any sequence of unital maps.

Eq. (5.4) establishes how a process tensor for an N -step process can be reconstructed based on a restricted set R of local operations; given R , one derives the dimension d_w of the vector space W that R spans, determines a set $\{\mathbb{R}_\gamma\}_{\gamma=1}^{d_w} \subset R$ of d_w linearly independent operations, and measures the output state for each of the d_w^N possible $\mathbb{R}_{\vec{\gamma}}$. The reconstructed process tensor yields the correct output state for any admissible sequence of operations $\mathbb{R}_{N:1} \in W^{\otimes N}$.

Unsurprisingly, outside the space W , the restricted process tensor does not yield meaningful results; the vector space of all possible sequences of local operations at N times can be decomposed as $[W \oplus W^\perp]^{\otimes N}$, where W^\perp is the orthogonal complement of W , *i.e.*, $\text{tr}(w^\dagger w') = 0$ for all $w \in W, w' \in W^\perp$. Denoting a basis of W^\perp (W) by $\{w'_\beta\}_{\beta=1}^{d-d_w}$ ($\{\mathbb{R}_\gamma\}_{\gamma=1}^{d_w}$), where d is the dimension of the system of interest, the set of all possible N -fold tensor products of elements of $\{\mathbb{R}_\gamma\}$ and $\{w'_\beta\}$ forms a basis of $W^{\otimes N} = [W \oplus W^\perp]^{\otimes N}$, and it is easy to see that every basis element that does not exclusively contain elements of $\{\mathbb{R}_\gamma\}_{\gamma=1}^{d_w}$ lies in the kernel of $\mathbb{T}_{N+1:1}^R$. Consequently, the action of $\mathbb{T}_{N+1:1}^R$ and the corresponding full process tensor $\mathcal{T}_{N+1:1}$ coincide on $W^{\otimes N}$, but the restricted process tensor does not allow for meaningful predictions of the output state for any sequence of operations $\mathbb{B}_{N:1} \notin W^{\otimes N}$.

By how much the full and the restricted process tensor differ depends both on the number of basis elements that get mapped to zero by $\mathbb{T}_{N+1:1}^R$, *i.e.*, the dimension of its kernel, as well as the action of the full process tensor on said basis elements. The number is given by $d^{4N} - d_w^N$, while the action of $\mathcal{T}^{N+1:1}$ outside of $W^{\otimes N}$ depends on the system-environment unitary maps that govern the evolution in between the times $\{t_\alpha\}$. Importantly, once the restricted process tensor has been reconstructed, there is no further information about the process, that can be gathered by means of operations in the set R , and as such, $\mathbb{T}_{N+1:1}^R$ is the maximal descriptor of the process for the given experimental situation.

5.2 PROJECTIVE MEASUREMENTS AND UNITARY CONTROL

Before we examine the mathematical properties of restricted process tensors, it is instructive to illustrate their explicit reconstruction for two concrete extremal examples of experimental control – the case where only unitary operations are available, and the case where only projective measurements can be performed. These two cases are extremal in the sense that unitary operations do not allow one to infer any information about the state of the system, while projective measurements provide direct access to information, but lead to a collapse into a definite (pure) state, and destroy system-environment correlations. As they decouple the system from its environment, projective measurements enable the construction of direct witnesses of correlations (see Sec. 5.4). On the mathematical side, these two sets of operations are tractable enough to derive the dimensions of their respective spans.

Additionally, unitary gates play a fundamental role in quantum computation [19, 191], while the scenarios where only projective measurements would, for example, suffice to fully reconstruct Markovian processes. Together, unitary maps and projective measurements are exhaustive in the sense that being able to perform both unitary operations as well as projective measurements is tantamount to having full experimental control over the system of interest. For example, every causal break can be considered a projective measurement followed by a unitary map that is independent of the measurement outcome.

5.2.1 Unitary Operations

A unitary map $\mathcal{V} : \mathcal{B}(\mathcal{H}_\alpha^0) \rightarrow \mathcal{B}(\mathcal{H}_\alpha^1)$ acting on a quantum state $\rho \in \mathcal{B}(\mathcal{H}_\alpha^0)$ at time t_α is of the form $\mathcal{V}[\rho] = V\rho V^\dagger$, where $V \in \text{SU}(d)$ is a unitary matrix. It is straightforward to see that, up to normalization, a unitary map \mathcal{V} corresponds to a pure, (unnormalized) maximally entangled state $V \in \mathcal{B}(\mathcal{H}_\alpha^1 \otimes \mathcal{H}_\alpha^0)$. We denote the span of the set of unitary maps by W_U . As $\text{tr}_{\alpha^1} V = \mathbb{1}_\alpha^0$ and $\text{tr}_{\alpha^0} V = \mathbb{1}_\alpha^1$, any element $Y \in W_U$ can be written as a real linear combination

$$Y = \sum_{\mu} b_{\mu} \left(\mathbb{1}_\alpha^1 \otimes \mathbb{1}_\alpha^0 + \sum_{k,l=1}^{d^2-1} c_{kl}^{(\mu)} \sigma_k^1 \otimes \sigma_l^0 \right), \quad (5.5)$$

where $\{\sigma_k^x\}$ are the $d^2 - 1$ traceless generators of $\text{SU}(d)$. The converse also holds; any map that can be expressed in the form of Eq. (5.5) lies in W_U [197]. The set of completely positive maps contained in W_U coincides with the set of unital maps (those that leave the completely mixed state invariant) [198].

From the fact that the operators $\{\mathbb{1}_\alpha^1 \otimes \mathbb{1}_\alpha^0, \sigma_k^1 \otimes \sigma_l^0\}_{k,l=1}^{d^2-1}$ are linearly independent, we deduce that W_U is $d_u = (d^2 - 1)^2 + 1$ dimensional (whereas the space of all possible local operations is d^4 -dimensional). For example, in the qubit case we have $d_u = 10$ and a basis of W_U that consists of unitary maps can be readily constructed [2] (see App. C.1). In the general, higher dimensional case, it is sufficient to randomly choose a set of d_u linearly independent unitary maps for the construction of the restricted process tensor.

A process tensor $\mathcal{T}_{N+1:1}^U$ constructed based on the set of unitary local operations can be meaningfully applied to any sequence $Y_{N:1} \in W_U^{\otimes N}$ of (possibly temporally correlated) unital maps. This means that, by measuring the output states for $[(d^2 - 1)^2 + 1]^N$ sequences of independent unitary operations, the output state for any sequence of unital maps can be predicted, and in a laboratory with only unitary operations at their disposal, the experimenter has maximal deducible information about the process at hand.

5.2.2 Projective Operations

If the experimental setup only allows for projective measurements (in arbitrary bases) of the system of interest, the set R_P of available operations coincides with rank-1 projective maps $P_\gamma \in \mathcal{B}(\mathcal{H}_\alpha^0 \otimes \mathcal{H}_\alpha^0)$. We denote the span of R_P by W_P . The action of a projective map \mathcal{P}_γ on a quantum state $\rho \in \mathcal{B}(\mathcal{H}_\alpha^0)$ is given by

$$\rho'_\gamma = \mathcal{P}_\gamma[\rho] = \Pi_\gamma \rho \Pi_\gamma = \Pi_\gamma \text{tr}(\Pi_\gamma \rho), \quad (5.6)$$

where Π_γ is a projector on a pure state, and we have $P_\gamma = \Pi_\gamma \otimes \Pi_\gamma^\top$. The (unnormalized) state ρ'_γ after the action of \mathcal{P}_γ is given by Π_γ , and $\text{tr}(\rho'_\gamma) = \text{tr}(\Pi_\gamma \rho)$ yields the probability to measure the outcome corresponding to Π_γ . Unlike unitaries, projective measurements destroy any correlations between the system and its environment, as $\mathcal{P}_\gamma \otimes \mathcal{I}_e[\rho_{se}] = \Pi_\gamma \otimes \text{tr}_s[(\Pi_\gamma \otimes \mathbb{1}_e)\rho_{se}]$, where ρ_{se} is a system-environment state. Consequently, they can be used to construct witnesses for system-environment correlations (see Sec. 5.4 and Ref. [144]).

As every projective map is of the form $P_\gamma = \Pi_\gamma \otimes \Pi_\gamma^\top$, any map $N \in W_P$ can be written as

$$N = \sum_\nu b_\nu \Pi_\nu \otimes \Pi_\nu^\top, \quad (5.7)$$

where Π_ν are pure states and $b_\nu \in \mathbb{R}$. W_P is at most $\frac{1}{4}d^2(d+1)^2$ dimensional [2] (see App. C.2). For the qubit case, a set $\{\Pi_\gamma\}$ of $\frac{1}{4}d^2(d+1)^2 = 9$ pure states that correspond to linearly independent maps $\{\mathcal{P}_\gamma\}$ has been constructed in Ref. [144]. For completeness, this set is reproduced in App. C.2.

For example, a one-slot process tensor $\mathcal{T}_{2:1}^P$ (i.e., a superchannel) constructed based on projections alone can be obtained by measuring the output states at t_2 for $\frac{1}{4}d^2(d+1)^2$ linearly independent projections $P_\gamma \in R_P$ that were performed at t_1 . From examination of Eq. (5.7), we see that it can be meaningfully applied to any CP map N that satisfies $\text{tr}_{1^o}(N) \cong (\text{tr}_{1^i}(N))^\top$, where we use ‘ \cong ’ instead of ‘ $=$ ’ as both elements live on different – albeit isomorphic – spaces. Analogously, an N -step restricted process tensor $\mathcal{T}_{N+1:1}^P$ can be reconstructed by measuring the output states for $\frac{1}{4}[d^2(d+1)^2]^N$ linearly independent sequences of projections; it can be applied to any physically admissible sequence of operations in $W_P^{\otimes N}$.

5.3 PROPERTIES OF RESTRICTED PROCESS TENSORS

The reconstruction scheme presented above shows how to obtain the maximal amount of information that can be deduced about the process at hand by means of limited resources. Naturally, besides the mere reconstruction, we are also interested in the structural properties of the reconstructed restricted objects.

The *full* descriptor $\mathbb{T}_{N+1:1}$ of a process satisfies several important mathematical requirements, each of them corresponding to a physical property of open quantum system dynamics. Here, we shall discuss which of these properties still hold for restricted process tensors, and which of them break down.

First, $\mathbb{T}_{N+1:1}$ acts linearly on the set of experimental interventions. This property is, by construction, also satisfied by restricted process tensors $\mathbb{T}_{N+1:1}^R$, and the action of $\mathbb{T}_{N+1:1}$ and $\mathbb{T}_{N+1:1}^R$ is identical on $W = \text{Span}(R)$. Secondly, $\mathbb{T}_{N+1:1}$ displays a ‘containment property’ [32] that we have already encountered when discussing the generalized Kolmogorov conditions that apply for process tensors. Given $\mathbb{T}_{N+1:1}$, defined on times $\{t_1, \dots, t_{N+1}\}$, all process tensors defined on subsets $\Lambda_k \subset \{t_1, \dots, t_{N+1}\}$ can be obtained from $\mathbb{T}_{N+1:1}$. *A priori*, it is unclear if this property also holds for restricted process tensors, and we will see shortly that in general, restricted process tensors only possess a partial containment property.

Finally, process tensors are completely positive, and, in the correct sense, trace preserving. As it turns out, neither of these properties is meaningful for restricted process tensors, as they

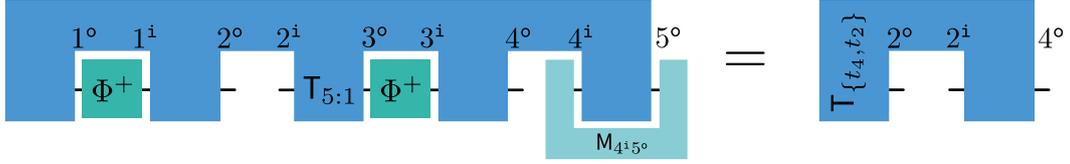


Figure 5.1: Containment property of process tensors. Intermediate slots can be ‘removed’ by contracting them with an identity channel, represented by Φ^+ , while final slots can be ‘removed’ by contraction with a deterministic comb. Here, the process tensor $\mathbb{T}_{\{t_5, t_2\}}$ defined on times $\{t_5, t_2\}$ is obtained from $\mathbb{T}_{5:1}$ defined on $\{t_5, \dots, t_1\}$ by applying an identity map at t_1 and t_3 and a deterministic comb $M_{5^o 4^i}$ over the final time step. For improved visibility, the labels of the wires are placed above them.

do not even necessarily output proper states when acting on operations that lay outside W , and in general their Choi states do not have to be positive or satisfy trace constraints. We will discuss the breakdown of complete positivity and trace preservation for restricted process tensors in detail after deriving their partial containment properties. Analogously, we shall investigate trace preservation of restricted process tensors, and see, that is is not well-defined, either.

5.3.1 Partial Containment Properties

As we have discussed in detail in Ch. 3, descriptors of a quantum stochastic process for different sets of times satisfy a generalized containment property. Given the descriptor \mathbb{T}_{Λ_K} of a process defined on times Λ_K the correct descriptor \mathbb{T}_{Λ_k} defined on the times $\Lambda_k \subseteq \Lambda_K$ is obtained by letting \mathbb{T}_{Λ_K} act on identity maps at the excessive times, *i.e.*,

$$\mathbb{T}_{\Lambda_k} = \mathbb{T}_{\Lambda_K} \star_{\alpha \in \Lambda_K \setminus \Lambda_k} \Phi_{\alpha}^+. \quad (5.8)$$

Here, the situation is slightly different, as the process tensors we consider have – in contrast to the scenario discussed in Ch. 3 – a final open output line. However, the reasoning is still fundamentally the same, with one slight difference.

To see this, consider a process tensor $\mathbb{T}_{5:1}$ with slots at times $\{t_4, t_3, t_2, t_1\}$ and a final output wire at t_5 (see Fig. 5.1). If we want to obtain the correct process tensor $\mathbb{T}_{\{t_4, t_2\}}$ that has one slot at time t_2 and a final open wire at t_4 , we have to contract the slots at t_1 and t_3 with identity maps. The remaining wires that are not part of $\mathbb{T}_{\{t_4, t_2\}}$, *i.e.*, the wires labeled by 4^i and 5^o in Fig 5.1, cannot be contracted with an identity map, as they do not form a slot. However, due to the causality constraint that $\mathbb{T}_{5:1}$ satisfies, the remaining wires (corresponding to the spaces $\mathcal{B}(\mathcal{H}_4^i)$ and $\mathcal{B}(\mathcal{H}_5^o)$) can be contracted with *any* deterministic comb to yield the correct ‘marginal’ (see Fig. 5.1). This slight difference of marginalization process is entirely technical in nature, and does not represent a fundamental difference in how process tensors and quantum stochastic processes are marginalized.

Now, to see in what sense restricted process tensors satisfy containment properties, let us consider the analogous example of a restricted process tensor $\mathbb{T}_{5:1}^R$ defined on times $\{t_5, \dots, t_1\}$, and we wish to deduce the correct restricted process tensor $\mathbb{T}_{\{t_4, t_2\}}^R$ defined on times $\{t_4, t_2\}$, with no initial input wire and a final output wire at t_4 (see Fig. 5.1 for orientation). If the

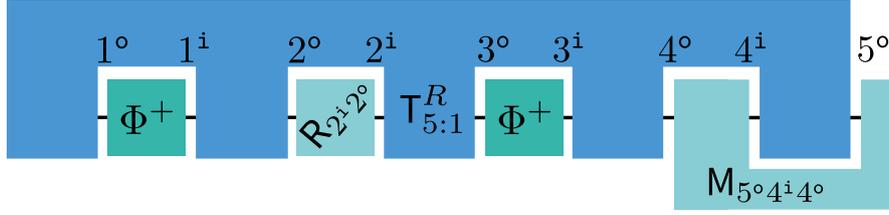


Figure 5.2: Partial containment property. While the state of the system at t_4 cannot generally be deduced from $T_{5:1}^R$, probabilities for intermediate operations (or sequences thereof) that lie in R can be obtained. In the figure, the probability for the implementation of the map $R_{2^i 3^o} \in R$ is represented.

experimental setup allows one to do nothing at a time step, *i.e.*, $\Phi^+ \in R$, then we can contract $T_{5:1}^R$ with identity maps at times $\{t_1, t_3\}$ to obtain

$$T_{\{t_5, t_4, t_2\}}^R = T_{5:1}^R \star \Phi_3^+ \star \Phi_1^+, \quad (5.9)$$

the correct restricted descriptor on times $\{t_2, t_4, t_5\}$. If the experimental setup does not allow one to do nothing, we can replace the identity map with the respective default map² for the process in the above equation, and, again, obtain the corresponding correct descriptor on $\{t_5, t_4, t_2\}$.

Now, in order to obtain the restricted process tensor that ends at t_4 instead of t_5 , we would have to be able compute the state of the system at t_4 for any earlier sequence of implementable operations. In general, the set R of available operations does *not* allow one to do so. For instance, if the operations that the experimenter can implement are unitary (with measurements only available at the final time t_5), no information about the system is obtained when the respective operations $\mathcal{V}_1, \dots, \mathcal{V}_4$ are performed, and a restricted process tensor that was reconstructed for sequences of unitary operations would not allow one to infer the state of the system at any of the times $t_\alpha \neq t_5$. However, the restricted process tensor $T_{\{t_5, t_4, t_2\}}^R$ computed according to Eq. (5.9) would enable the correct prediction of the *probability* of implementation of a CP map $R_{2^i 2^o} \in R$ at time t_2 , where we label maps by the respective Hilbert spaces they are defined on, to simplify orientation. To compute it, we have to contract $T_{\{t_5, t_4, t_2\}}^R$ with $R_{2^i 3^o}$ and *any* experimentally performable deterministic comb $M_{5^o 4^i 4^o} = \mathbb{1}_5^o \otimes M_{4^i 4^o}$, where $M_{4^i 4^o} \in R$, *i.e.*, this probability is equal to

$$T_{\{t_5, t_4, t_2\}}^R \star R_{2^i 3^o} \star \mathbb{1}_5^o \otimes M_{4^i 4^o}. \quad (5.10)$$

See Fig 5.2 for a graphical representation. With this, we can define a restricted process tensor

$$T_{\{t_2\}}^R = T_{\{t_5, t_4, t_2\}}^R \star (\mathbb{1}_5^o \otimes M_{4^i 4^o}), \quad (5.11)$$

that has one open slot at t_2 and lets one predict the probability of implementation of any CP map in R that is performed at t_2 , but not the corresponding output state at t_4 . In a sense, it corresponds to the correct restricted process tensor $T_{\{t_2\}}^R$, but with the final output wire cauterized. Consequently, while from $T_{\Lambda_X}^R$ we cannot in general obtain a restricted process

² This could, for example, happen if the respective experimental setup is immobile, and different parts of it cannot be turned off or removed independent of the rest. The default map at a time t_α would then be the overall CPTP map corresponding to the respective element of the setup.

tensor $\mathbb{T}_{\Lambda_k}^R$ for fewer times $\Lambda_k \subset \Lambda_K$, that yields the correct output states at the final time $t_{\max} = \max(\Lambda_k)$, we can always obtain a process tensor that yields the correct *probabilities* on operations on any subset of times [2]. This process tensor is obtained by contracting the given restricted process tensor $\mathbb{T}_{\Lambda_K}^R$ with identity maps at times $t_\alpha \in \Lambda_K \setminus \Lambda_k$ that satisfy $t_\alpha < t_{\max} = \max(\Lambda_k)$, and with a performable deterministic comb on times $t_\alpha \in \Lambda_K \setminus \Lambda_k$ that satisfy $t_\alpha \geq t_{\max} = \max(\Lambda_k)$.

The resulting object might, in some cases, be trivial. For example, when the performable operations are unitaries, all performable operations are CPTP, and the probability of implementation is (trivially) equal to 1. However, in more general cases, it yields the correct probability to perform any non-deterministic sequence of CP operations that the restricted process tensor can meaningfully act upon, and is thus a meaningful descriptor for the respective subset of times. In this well-defined sense, restricted process tensors possess a partial containment property.

A particular case of experimental control is the one, in which the set of operations R that is available to the experimenter allows them to perform an informationally complete measurement. Given this experimental control, it is possible to derive ‘actual’ intermediate restricted process tensors, that allow one to not only obtain the correct outcome probabilities, but also the final state at any time $t_\alpha \in \Lambda_K$. Availability of an informationally complete measurement implies that the state of the system at each time can be inferred by simple post-processing of the data contained in \mathbb{T}_{Λ_K} , and consequently all intermediate restricted process tensors can be constructed [2] (see App. C.3 for details). This is the case when the set of performable operations coincides with the set of projective measurements, and is the reason, why classical processes display containment properties. Here, the set of projective measurements in the computational basis is informationally complete, and hence the *a priori* restricted information given by joint probability distributions allows one to derive all descriptors for fewer times by simple marginalization.

5.3.2 Complete Positivity and Trace Preservation

In Ch. 4 we discussed the meaning of complete positivity for process tensors. Basically, this property ensures that *every* sequence of CP maps that acts on the system of interest and potentially some ancillary degrees of freedom gets mapped to a completely positive map.

For restricted process tensors $\mathbb{T}_{N+1:1}^R$, complete positivity is not a well-defined property. By construction, they can only be meaningfully applied to a *subset* of control operations, while they yield physically nonsensical results for operations outside this subset. Put differently, there are control operations $M_{N:1}^{(sa)}$ defined on the system of interest and an ancilla, whose restriction to the system degrees of freedom lies outside the set of operations that $\mathbb{T}_{N+1:1}^R$ is defined on. In turn, this also implies that $\mathbb{T}_{N+1:1}^R$ is not necessarily a positive matrix, but can always be extended to one (that is, its unrestricted counterpart $\mathbb{T}_{N+1:1}$). We should emphasize, though, that even in the special case where the Choi state of a restricted process tensor is positive, complete positivity is still not a well-defined property; a positive matrix $\mathbb{T}_{N+1:1}^R$ would indeed map *every* CP map $M_{N:1}^{(sa)}$ onto a CP map, but if $M_{N:1}^{(sa)}$ lies outside the domain where $\mathbb{T}_{N+1:1}^R$ is

well-defined, then the resulting map will not necessarily correspond to what would be observed in the actual experiment.

Nonetheless, restricted process tensors are – unlike the non-CP briefly maps discussed in the previous chapter – operationally well-defined and the possible ‘break-down’ of complete positivity is not fundamental but merely due to the description of the process we chose. This situation is similar to incomplete tomography of a quantum state; the probabilities obtained from probing a state with a POVM that is not informationally complete might be faithfully reproduced by a non-positive ‘density matrix’. This matrix would yield correct probabilities for each of the POVM elements used to probe the state in the first place, but would not contain any information about probabilities for other POVM elements. Quite obviously, the non-positivity of such a ‘density matrix’ is neither fundamental, nor does it imply negative probabilities, but is simply a remnant of the chosen representation.

An important example for a limited set of operations that always leads to a completely positive restricted process tensor is when the performable operations are only projective measurements in the computational basis. This is exactly the case for classical processes. Here, the CP maps $\{|k\rangle\langle k| \otimes |k\rangle\langle k|\}_{k=1}^d$ are self-dual, and a process tensor reconstructed with these duals is, by construction, positive. Nonetheless, such a restricted process tensor obtained from projective control operations *cannot* be meaningfully applied to any control operation outside the span of projective operations, even though it will *always* yield positive probabilities.

Finally, restricted process tensors do not necessarily satisfy causality constraints, *i.e.*, they are not trace preserving. While the structure of a restricted process tensor has to be such that no performable causally ordered control comb can have any influence on past measurement statistics, this requirement does not have to hold for deterministic control combs that the restricted process tensor can not be meaningfully applied to. As was the case for complete positivity, though, any restricted process tensor $\mathbb{T}_{N+1:1}^R$ can always be extended to a causally ordered object (namely $\mathbb{T}_{N+1:1}$).

Before we continue, it is instructive to briefly review what we have learnt about restricted process tensors so far: Restricted process tensors are distinctly operational objects in the sense that they contain the maximum amount of information that can be inferred about a process, given certain experimental limitations. For their construction, we give up axiomatic considerations about the properties that open dynamics should satisfy, and switch to a full input/output picture of dynamics; once the output state corresponding to any implementable sequence of operations can be predicted, the process at hand is fully described for practical purposes. This is exactly the descriptor that is provided by the restricted process tensor.

While some of the nice mathematical properties of full process tensors fail to hold in the restricted case, importantly, we still recover partial containment properties. Perhaps most notably, restricted process tensors have a clear-cut operational meaning, unlike NCP maps. *A priori*, though, it is unclear, what additional conclusions can be drawn from the information contained in a restricted process tensor. This question will be the subject of the subsequent sections.

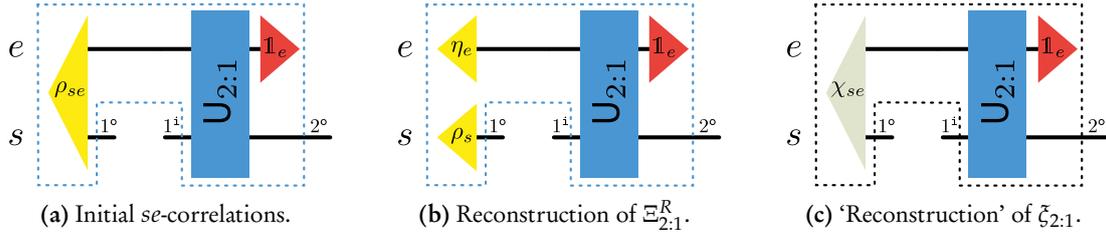


Figure 5.3: Different scenarios in the reconstruction of $\mathbb{T}_{2:1}$ and/or $\mathbb{T}_{2:1}^R$ (represented by the respective dotted lines). If there are initial se -correlations, then $\chi_{se} = \rho_{se} - \rho_s \otimes \eta_e \neq 0$. The corresponding ‘reconstruction’ of $\zeta_{2:1}$ depicted in (c) is fictitious, and done by post-processing of $\mathbb{T}_{2:1}$.

5.4 WITNESSES FOR SYSTEM-ENVIRONMENT CORRELATIONS FROM LIMITED CONTROL

One of the key features of the process tensor formalism is that it can describe the dynamics of open systems that are initially correlated with their environment. Additionally, besides providing a descriptor of the dynamics, it allows for the unambiguous *detection* of correlations and, more generally, memory effects. We shall now investigate, the extent to which similar statements can be made based on limited process tensors.

First, we discuss the detection of initial correlations between the system of interest and its environment. Such correlations are a generic feature of most experiments and represent a record of past system-environment interactions. Therefore, detecting initial correlations, in turn, implies detecting non-Markovian dynamics [2].

To see if it is still possible to witness se correlations by means of a restricted set R of control operations, consider an initial system-environment state (before preparation) of the form

$$\rho_{se} = \rho_s \otimes \eta_e + \chi_{se}, \quad (5.12)$$

where $\rho_s = \text{tr}_e \rho_{se}$, $\rho_e = \text{tr}_s \rho_{se}$, and $\chi_{se} = \rho_{se} - \rho_s \otimes \eta_e$ contains all initial correlations between the system of interest and its environment. If we had full experimental control, we could probe the correlations by reconstructing a one-slot process tensor $\mathbb{T}_{2:1}$ defined on times $\{t_1, t_2\}$, and check if it is of product form, *i.e.*, if $\mathbb{T}_{2:1} = \mathbb{L}_{2:1} \otimes \rho_s$, where $\mathbb{L}_{2:1} \in \mathcal{B}(\mathcal{H}_2^o \otimes \mathcal{H}_1^i)$ is a CPTP map from t_1 to t_2 and $\rho_s \in \mathcal{B}(\mathcal{H}_1^o)$ is the initial state of the system. Any deviation

$$\tilde{\zeta}_{2:1} := \mathbb{T}_{2:1} - \frac{1}{d} \text{tr}_{1^o}(\mathbb{T}_{2:1}) \otimes \text{tr}_{1^i 2^o}(\mathbb{T}_{2:1}) \quad (5.13)$$

from this product form implies the existence of initial correlations. It is straightforward to see that $\tilde{\zeta}_{2:1}$ can be expressed directly in terms of the se correlation term χ_{se} as

$$\tilde{\zeta}_{2:1} = \chi_{se} \star \mathbb{U}_{2:1} \star \mathbb{1}_e, \quad (5.14)$$

where $\mathbb{U}_{2:1}$ is the se unitary map between t_1 and t_2 [31] (see Fig. 5.3 for a graphical representation). $\tilde{\zeta}_{2:1}$ encapsulates the time evolution of the initial correlations χ_{se} . Importantly, $\tilde{\zeta}_{2:1} \neq 0$ implies that the initial state ρ_{se} was correlated, while $\tilde{\zeta}_{2:1} = 0$ merely implies that if initial correlations were present, the unitary $\mathbb{U}_{2:1}$ does not allow for their detection via local operations. In this sense, $\tilde{\zeta}_{2:1}$ is the maximal information about correlations that can be inferred via local operations alone.

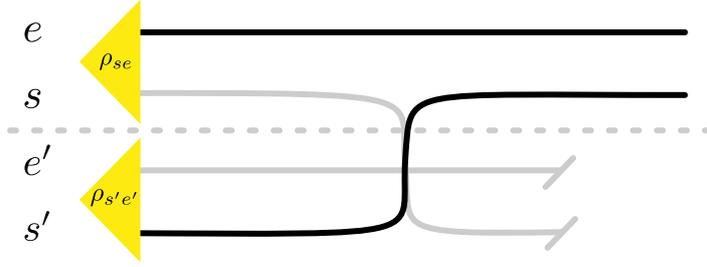


Figure 5.4: Preparation of a product state. Two copies of the initial state, ρ_{se} and $\rho_{s'e'}$, are created and their system degrees of freedom are swapped. Subsequently, the degrees of freedom of the second copy are disregarded. The resulting se state is of product form $\rho_s \otimes \eta_e$. For better delineation, the wires that correspond to degrees of freedom that are discarded are depicted in gray.

The direct derivation (5.13) of the correlation memory matrix $\xi_{2:1}$ from $T_{2:1}$ relies on the fact that the experimenter has full control over the system of interest and its derivation is not directly applicable to $T_{2:1}^R$. Indeed, simply reiterating the derivation (5.13) would, in general, yield false positives for witnessing correlations [2]. Even if the underlying process is Markovian, there is no reason, why the corresponding restricted process tensor would have to be of product form, and as such, a potential $\xi_{2:1}^R$ computed for a restricted process tensor $T_{2:1}^R$ via Eq. (5.13) might not vanish, even though the process is Markovian. In particular, this holds true, if the available control operations are not causal breaks, *i.e.*, not of product form.

However, there are several ways around this problem. Firstly, imagine an experimental situation that allows for the preparation of two copies of the initial state, ρ_{se} and $\rho_{s'e'}$, and admits a swap operation between the two states. In detail, let $\mathbf{S}_{ss'}$ be a swap operation between the system Hilbert spaces \mathcal{H}_s and $\mathcal{H}_{s'}$ of the two prepared initial states. With this, we can prepare an initial product state by swapping the system degrees of freedom of both states, and subsequently discarding the degrees of freedom of the second copy of the se state [2] (see Fig. 5.4):

$$\text{tr}_{s'e'} \left\{ \mathbf{S}_{ss'} [\rho_{se} \otimes \rho_{s'e'}] \mathbf{S}_{ss'}^\dagger \right\} = \rho_s \otimes \rho_{e'} \quad (5.15)$$

where $\text{tr}_{s'e'}$ denotes the trace with respect to the degrees of freedom of the second copy of the system-environment state. Now, using this product state as the initial state, we can experimentally reconstruct the restricted process tensor $\Xi_{2:1}^R$ for this situation (see Fig. 5.3b) and compare it to the original restricted process tensor $T_{2:1}^R$; if they do not coincide, then there are initial se correlations.

The corresponding correlation-memory matrix $\tilde{\zeta}_{2:1}^R = T_{2:1}^R - \Xi_{2:1}^R$ contains *all* the information about correlations that can be detected by local preparations that are an element of $\text{Span}(R)$. As already mentioned, this construction is not equivalent to simply reiterating Eq. (5.13) for restricted process tensor, *i.e.*, is not obtained by post-processing of $T_{2:1}^R$, but by experimental reconstruction of *both* $T_{2:1}^R$ and $\Xi_{2:1}^R$. Importantly, $\tilde{\zeta}_{2:1}^R$ by construction cannot yield false positives for the presence of initial system-environment correlations; if $\tilde{\zeta}_{2:1}^R \neq 0$, then there exist initial system-environment correlations. If, on the other hand $\tilde{\zeta}_{2:1}^R = 0$, no correlations that can be detected by local preparations in $\text{Span}(R)$ are present. Depending on the dimension d_w of $\text{Span}(R)$ and the total system-environment dynamics $U_{2:1}$, $\tilde{\zeta}_{2:1}^R$ and

$\xi_{2:1}$ can contain the same information about the existence of initial correlations; in general, though, they will differ (see Sec. 5.5.1 for explicit examples).

In principle, the swap operation enables the construction of witnesses for correlations for any set R of available local operations, without requiring any knowledge of the initial system state ρ_s (unlike, e.g., the similar correlation witnesses proposed in [168, 169, 199] and experimentally reconstructed in [200]). This construction of witnesses makes full use of the resources available to the experimenter, and $\xi_{2:1}^R$ is evidently the maximum of information about correlations that can be inferred based on operations in R . However, its reconstruction requires both the preparation of two copies of the same initial state as well as the availability of a swap operation between the respective system degrees of freedom.

If the experimental setup does not allow one to deduce $\xi_{2:1}^R$, the feasibility of witnesses for correlations depends on the set R of available operations. If $\text{Span}(R)$ contains operations $M_{1^0 1^1}^{(\alpha)}$ that decouple the system from its environment and map the system to a known state, then witnesses can be straightforwardly reconstructed. For example, this is the case when the set of available operations contains projective measurements (see Sec. 5.2.2 and [144]). Such operations in general have the form of causal breaks, *i.e.*,

$$M_{1^0 1^1}^{(\alpha)} = \rho_{1^1}^{(\alpha)} \otimes E_{1^0}^{(\alpha)}, \quad (5.16)$$

where $E_{1^0}^{(\alpha)}$ is a POVM element and $\rho_{1^1}^{(\alpha)}$ is a quantum state. Evidently, if the dynamics is Markovian, then

$$T_{2:1}^R \star E_{1^0}^{(\alpha)} = p_\alpha O_{2:1}^R, \quad (5.17)$$

where $p \in [0, 1]$ is the probability to measure the outcome corresponding to $E_{1^0}^{(\alpha)}$ and $O_{2:1}^R \in \mathcal{B}(\mathcal{H}_2^0 \otimes \mathcal{H}_1^1)$ is a (not necessarily positive) matrix that is independent of α . Basically, what Eq. (5.17) says is that if the dynamics is Markovian and the experimenter has access to causal breaks, they could perform partial tomography of the channel $L_{2:1}$ from t_1 to t_2 , yielding $O_{2:1}^R$. If they discover, that this channel (or rather the resulting matrix) depends on the measurement outcome α , *i.e.*, if Eq. (5.17) does not hold, then they know that there are initial correlations that influence the dynamics of the system.

In what follows, for sake of clarity, we will focus on the case where $\xi_{2:1}^R$ can actually be constructed, and compare its ability to detect initial correlations with that of $\xi_{2:1}^R$.

5.4.1 Witnesses for Non-Markovianity

If an experiment consists of more than one time-step, we can make assertions about the propagation of memory effects throughout the dynamics. As we have seen in the previous chapter, memory effects are present if for the same input state at t_α , but different histories (or trajectories [7]) of previous operations, the future statistics differ. Equivalently, expressed in terms of link products, a process is Markovian iff at every time $t_\alpha \leq N$, we have

$$T_{N+1:1} \star M_{\alpha:1} \propto T_{N+1:\alpha} \quad (5.18)$$

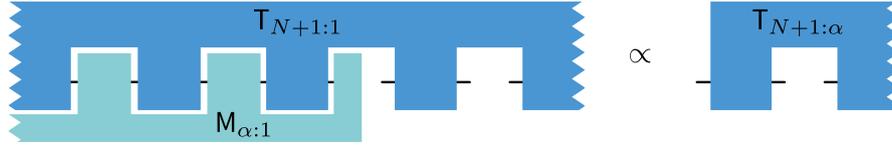


Figure 5.5: Markovian process tensor. For every Tester element $M_{\alpha:1}$ that ends on an output of the process tensor, the resulting object $T_{N+1:1} \star M_{\alpha:1}$ has to be – up to normalization – a proper process tensor $T_{N+1:\alpha}$ that is independent of $M_{\alpha:1}$.

where $T_{N+1:\alpha} \in \mathcal{B}(\mathcal{H}_{N+1}^0 \otimes \cdots \mathcal{H}_\alpha^1)$ is a proper process tensor and independent of $M_{\alpha:1} \in \mathcal{B}(\mathcal{H}_\alpha^0 \otimes \cdots \otimes \mathcal{H}_1^0)$. Importantly, the above equation only holds if $M_{\alpha:1}$ ends on an output space of the process tensor (see Fig. 5.5).

Demanding Eq. (5.18) to hold for all t_α then implies the product structure (4.40) of Markovian processes:

$$T_{N+1:1}^{\text{Markov}} = L_N \otimes L_{N-1} \otimes \cdots \otimes L_1 \otimes \rho_s, \quad (5.19)$$

where each L_α is a channel from t_α to $t_{\alpha+1}$.

For the investigation of the Markovianity of a process by means of restricted process tensors, there are – just like for the case of initial *se*-correlations – two cases that have to be distinguished.

Firstly, whenever $\text{Span}(R)$ contains at least one causal break, Eq. (5.18) can be evaluated directly; if we find different performable histories $M_{\alpha:1}, M'_{\alpha:1} \in R^{\otimes \alpha}$ (ending on an output of the process) such that at some time t_α

$$T_{N+1:1}^R \star M_{\alpha:1} \not\propto T_{N+1:1}^R \star M'_{\alpha:1} \quad (5.20)$$

then the process is non-Markovian. As for the case of full control, this criterion can be tested for in a finite number of experiments, but requires that the experimenter can perform operations that end on an output, *i.e.*, allow for the preparation of fresh states that are independent of the history. Naturally, the converse of Eq. (5.20) does not hold; if the restricted set of operations fails to detect a history dependence/memory, it does not mean that there is none, rather, that it could not be detected within the experimental limitations. An example of an experimental setup for which $\text{Span}(R)$ contains causal breaks, but does not allow to unambiguously decide for the existence of memory effects is one in which the experimenter can only perform projective measurements.

The second case that has to be investigated is the one where $\text{Span}(R)$ does not contain causal breaks and a direct evaluation of Eq. (5.20) is *not* possible. A prominent example of this is the case where R is the set of unitary operations (see Sec. 5.2.1). As the detection of non-Markovianity hinges on the availability of causal breaks, such a set of local operations seems to be inadequate for the investigation of the Markovianity of a process. However, as discussed above, such a set of operations can, in principle, be employed for the detection of *se* correlations, if an additional swap operation can be performed.

It is intuitively clear that, if correlations can be detected at any time via local operations, the process must be non-Markovian [2]. Correlations constitute a memory of past interactions, and their detection implies that memory plays a non-negligible role in the process. Consequently, non-Markovianity can be tested for by checking for system-environment correlations at each

time t_α , for varying earlier sequences of operations $M_{\alpha:1}$. Again, the converse statement does not hold. On the one hand, a restricted set of local operations does not necessarily allow for the detection of all locally detectable correlations. On the other hand, a process can also be non-Markovian without any system-environment correlations being present at any point in time [105].

Additionally, it is questionable if an experimental setup that does not allow for causal breaks offers the possibility of a swap operation with an additional auxiliary system. Consequently, the resort to the detection of non-Markovianity via the detection of *se* correlations by means of a swap operation is rather mentioned for completeness than potential applicability. For a feasible *direct* detection of non-Markovianity, causal breaks are inescapable.

A very interesting alternative approach for the detection and quantification of non-Markovian effects based on restricted experimental control has recently been proposed in [148]. There, the authors trained a machine learning algorithm to decide, whether or not a given dynamics was Markovian. Specifically, under the assumption that the experimenter can only perform one of the Pauli matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ at each time t_α (and a measurement at the final time t_{N+1}), *i.e.*, only reconstruct a restricted process tensor, they trained the algorithm on a set of known circuits to decide both the presence of memory, as well – when it existed – the size of the corresponding environment, for new, unknown dynamics. Naturally, this approach is limited by the size of the environment that the algorithm can be trained on. Here, we will not follow this route, and stick with the restricted correlation matrices $\tilde{\zeta}_{2:1}^R$, that can, in principle be reconstructed, and possess a clear-cut operational interpretation.

5.5 EXAMPLES AND APPLICATIONS OF RESTRICTED PROCESS TENSORS

Having introduced restricted process tensors in the previous sections, it is now time to illustrate the concepts we encountered for some instructive examples, chosen such that the effects we aim to study are present and can be straightforwardly investigated. Consequently, we are not going to present examples with fully realistic dynamics, as the corresponding computational complexity would overshadow the conceptual points we want to make. This fact notwithstanding, the reconstruction of restricted process tensors and of the witnesses for correlations and memory effects that we have presented in the previous sections, is, in practice, independent of the complexity of the underlying dynamics, and scales with the size of the system of interest and the number of time steps ‘only’.³ As such, even though our examples are toy models, the corresponding *experimental* reconstruction of $T_{N+1:1}^R$ would, in principle, be no more difficult than for any other kind of dynamics.

5.5.1 Reconstruction of Qubit Dynamics

It is instructive to start the illustration of the reconstruction of restricted process tensors and correlation witnesses with a low dimensional and computationally accessible example. For ease of notation, for the most part, we restrict the investigation to the one-slot process tensors $T_{2:1}^U$

³ As for the case of full process tensors, it scales exponentially with the number of times it is defined on.

and $\mathbb{T}_{2:1}^P$, reconstructed from local unitary and projective maps, respectively. The generalization to restricted N -slot process tensors is always possible in a straightforward way.

The simplest conceivable open quantum system is the case of a qubit coupled to another qubit that serves as the environment. While in general the initial system-environment state before preparation is considered constant and part of the dynamics, here, we shall investigate a family of initial states in order to explicitly analyze the set of detectable correlations. We choose the seven-dimensional family of X -states [201, 202] as a blueprint for initial system-environment states ρ_{se} . X -states are states of the form

$$\rho_{se}^X = \begin{pmatrix} a_{11} & 0 & 0 & a_{14} \\ 0 & a_{22} & a_{23} & 0 \\ 0 & a_{23}^* & a_{33} & 0 \\ a_{14}^* & 0 & 0 & a_{44} \end{pmatrix} \quad (5.21)$$

in a given basis (in our case the eigenbasis of $\sigma_z^s \otimes \sigma_z^e$). Eq. (5.21) describes a valid quantum state iff ρ_{se}^X has unit trace, $a_{22}a_{33} \geq |a_{23}|^2$ and $a_{11}a_{44} \geq |a_{14}|^2$. The family of X -states includes both entangled states and separable, as well as product states. Hence, it is well-suited for the analysis of the detectability of correlations by means of local unitary operations or projective measurements. For our system-environment Hamiltonian we choose

$$H_{se} = \omega \left(\sigma_x^s \otimes \sigma_x^e + \sigma_y^s \otimes \sigma_y^e + \sigma_z^s \otimes \sigma_z^e \right), \quad (5.22)$$

where $\omega \in \mathbb{R} \setminus \{0\}$. These choices of H_{se} and ρ_{se}^X allow for an analytical construction of $\mathbb{T}_{2:1}$, $\mathbb{T}_{2:1}^U$ and $\mathbb{T}_{2:1}^P$, but are also non-trivial enough to exhibit the features we would like to investigate, in particular the effects of initial system-environment correlations.

We can now explicitly construct the full one-step process tensor, as well as the restricted ones. To this end, we compute the output states for a full basis of the space $\mathcal{B}(\mathcal{H}_1^{\dagger} \otimes \mathcal{H}_1^0)$ of local operations at time t_1 . Such a possible basis is given by the set $\{Q_k \otimes Q_\ell^T\}_{k,\ell=1}^2$, where the pure states

$$Q_1 = \frac{1}{2}(\mathbb{1} + \sigma_z), \quad Q_2 = \frac{1}{2}(\mathbb{1} + \sigma_x), \quad Q_3 = \frac{1}{2}(\mathbb{1} - \sigma_x), \quad Q_4 = \frac{1}{2}(\mathbb{1} + \sigma_y) \quad (5.23)$$

constitute a basis of $\mathcal{B}(\mathcal{H}_1^{\dagger})$ and $\mathcal{B}(\mathcal{H}_1^0)$. Measuring the state of the system after the time $\Delta t := t_2 - t_1$ has lapsed would then yield the corresponding output states

$$\rho^{(k\ell)'} = \mathbb{1}_e \star \mathbb{U}_{2:1} \star Q_{k\ell} \star \rho_{se}^X, \quad (5.24)$$

where $Q_{k\ell} = Q_k \otimes Q_\ell^T$, and $\mathbb{U}_{2:1}$ is the Choi matrix of the unitary map $\mathcal{U}_{2:1}[\rho_{se}] = e^{-iH_{se}\Delta t} \rho_{se} e^{iH_{se}\Delta t}$. Each of the maps $Q^{(k\ell)}$ corresponds to a projection of the system state on Q_ℓ , followed by a reparation of Q_k . The full one-step process tensor $\mathbb{T}_{2:1}$ can be calculated, using Eq. (4.25), as

$$\mathbb{T}_{2:1} = \sum_{k,\ell=1}^4 \rho^{(k\ell)'} \otimes q_k^* \otimes q_\ell^*, \quad (5.25)$$

with $\text{tr}(q_m^\dagger Q_n) = \delta_{mn}$.

Accordingly, in order to construct $\mathbb{T}_{2:1}^U$ based on the set of unitary preparations, we fix a basis of the space spanned by the unitary maps. Here, we choose the basis $\{\mathcal{V}_\beta\} = \{\mathcal{V}_0, \mathcal{V}_{(k,\pm)}, \mathcal{V}_{(k+\ell+1,+)}\}$ (with $k, \ell \in \{1, 2, 3\}$ and $k < \ell$), derived in App. C.1. However, in principle, any set of $(d^2 - 1)^2 + 1 = 10$ linearly independent unitary operations would yield the same results. The output states corresponding to each of these unitary preparations are given by

$$\zeta^{(\beta)'} = \mathbb{1}_e \star \mathbb{U}_{2:1} \star \mathbb{V}_\beta \star \rho_{se}^X, \quad (5.26)$$

and the corresponding restricted process tensor $\mathbb{T}_{2:1}^U$ can be constructed via

$$\mathbb{T}_{2:1}^U = \sum_{\beta=1}^{10} \zeta^{(\beta)'} \otimes \mathbf{v}_\beta^*, \quad \text{where} \quad \text{tr}(\mathbf{v}_\beta^\dagger \mathbf{v}_\gamma) = \delta_{\beta\gamma}. \quad (5.27)$$

In the same vein, we construct $\mathbb{T}_{2:1}^P$ by using the basis of the set of projective maps introduced in Eq. (C.10). Importantly, each of these prescriptions is an operational one, *i.e.*, the reconstruction that we are carrying out numerically, could be conducted in exactly the same way in a laboratory.

Given $\mathbb{T}_{2:1}$, the correlation-memory matrix $\zeta_{2:1}$ can be readily derived by using Eq. (5.13); the restricted correlation-memory matrix $\tilde{\zeta}_{2:1}^U$ ($\tilde{\zeta}_{2:1}^P$) is obtained by constructing the one-step process tensor $\Xi_{2:1}^U$ ($\Xi_{2:1}^P$) for the uncorrelated initial state $\tilde{\rho}_{se}^X = \rho_s^X \otimes \eta_e^X$ and subtracting it from $\mathbb{T}_{2:1}^U$ ($\mathbb{T}_{2:1}^P$). As the explicit matrix representations of $\mathbb{T}_{2:1}$, $\mathbb{T}_{2:1}^U$, and $\mathbb{T}_{2:1}^P$ are not particularly insightful in their own right, we refrain from providing them here,⁴ and rather report the corresponding results with respect to the detection of initial correlations.

It turns out that, except for trivial total dynamics ($\omega t = n\frac{\pi}{2}$), the restricted correlation-memory matrices $\tilde{\zeta}_{2:1}^U$ and $\tilde{\zeta}_{2:1}^P$ are equal to zero iff [2]

$$a_{23} = a_{14} = 0 \quad \text{and} \quad a_{22}a_{33} = a_{11}a_{44}. \quad (5.28)$$

However, we have $\tilde{\zeta}_{2:1}^U \neq \tilde{\zeta}_{2:1}^P$ and the total correlation-memory matrix $\zeta_{2:1}$ differs from both $\tilde{\zeta}_{2:1}^U$ and $\tilde{\zeta}_{2:1}^P$; nonetheless, $\zeta_{2:1}$ can detect exactly the same kinds of correlations, as it is also equal to zero iff Eq. (5.28) is fulfilled.

X -states that satisfy Eq. (5.28) are product states, which means that in this particular case *any* correlations (classical or quantum) between the system and the environment could be detected with both $\tilde{\zeta}_{2:1}^U$ and $\tilde{\zeta}_{2:1}^P$ (and hence also with the unrestricted correlation-memory matrix $\zeta_{2:1}$).

On the other hand, the ‘do nothing’ correlation-memory matrix $\tilde{\zeta}_{2:1}^I$, *i.e.*, the correlation-memory matrix that can be constructed without performing any local operation (except for the swap $\mathbf{S}_{ss'}$ necessary to create an uncorrelated system-environment state), would be zero iff $\Im(a_{23}) = 0$ (*i.e.*, $a_{23} \in \mathbb{R}$). Unsurprisingly, performing local operations substantially increases the set of detectable correlations.

In this example, unitary preparation and projections can reveal exactly the same kinds of initial correlations as the full set of local operations. This is not generally true. Consider, for example, a two-qubit state with correlation matrix $\chi_{se} = \kappa \sigma_z^s \otimes \sigma_x^e$, where $\kappa \in \mathbb{R} \setminus \{0\}$, and a

⁴ The corresponding calculations are – as already mentioned – not computationally demanding and can be carried out at ease.

total unitary evolution between t_1 and t_2 that is given by the swap operator \mathbf{S} between system and environment. The action of $\tilde{\zeta}_{2:1}^U$ on an arbitrary unitary map V reads

$$\tilde{\zeta}_{2:1}^U \star V = \mathbb{1}_e \star \mathbf{S} \star V \star \chi_{se} = \kappa \operatorname{tr}_e(\sigma_x^s \otimes \mathcal{V}[\sigma_z^e]) = 0, \quad (5.29)$$

where \mathbf{S} is the Choi state of the swap operator, and hence $\tilde{\zeta}_{2:1}^U = 0$. On the other hand, $\tilde{\zeta}_{2:1} \neq 0$; for example, if we consider the causal break operation $|0\rangle\langle 0| \otimes |1\rangle\langle 1|$, we obtain

$$\begin{aligned} \tilde{\zeta}_{2:1} \star (|0\rangle\langle 0| \otimes |1\rangle\langle 1|) &= \kappa \langle 1| \sigma_z |1\rangle \operatorname{tr}_e(\sigma_x \otimes |0\rangle\langle 0|) \\ &= -\kappa \operatorname{tr}_e(\sigma_x \otimes |0\rangle\langle 0|) = -\kappa \sigma_x. \end{aligned} \quad (5.30)$$

In this case, the fact that the correlations cannot be detected by unitary transformations alone, stems from a particular interplay between the total unitary evolution (the swap operation \mathbf{S}) and the correlation matrix χ_{se} .

5.5.2 Restricted Process Tensors and Quantum Control

An important field where a restricted set of performable local operations and the presence of non-Markovian effects come into play together is that of quantum control of open systems (see for example Ref. [149] for an introduction). Here, generally speaking, the goal is to steer the system of interest to a desired final state by means of local, time-dependent Hamiltonians, which can be controlled by the experimenter. For microscopic models that assume knowledge of the total system-environment Hamiltonian, the impact of these local Hamiltonians on the dynamics of the system can often be readily deduced. However, if only local information like, for example, a master equation description, is at hand, it is in general unclear how to include the influence of a local operation into the description [127]. The process tensor approach is tailored to solve this problem in an operational manner.

The presence of memory effects is of particular importance for dynamical decoupling experiments [152], where the local Hamiltonians are employed in such a way that they average out the influence of the environment and, effectively, keep the system decoupled from its environment. This is only possible if memory effects are present [127]. Under the assumption that the time span over which the local Hamiltonians act is small compared to typical time scales of the dynamics of the system (*i.e.*, the Hamiltonians basically act at fixed times as ‘kicks’ of infinite strength [203]), such an experimental setup can be described as a mapping from an initial state $\rho_s \in \mathcal{B}(\mathcal{H}_0^{\mathfrak{h}})$ of the system at time t_0 and a sequence of unitary maps $\{\mathcal{V}_\alpha\}_{\alpha=1}^N$, that act on the system at times $\{t_\alpha\}_{\alpha=1}^N$, to a final state

$$\rho'_s = \mathcal{T}_{N+1:0}[\mathcal{V}_N, \dots, \mathcal{V}_1, \rho_s] = \mathbb{T}_{N+1:0} \star \mathcal{V}_N \star \dots \star \mathcal{V}_1 \star \rho_s. \quad (5.31)$$

To demonstrate the applicability of the restricted process tensor framework to quantum control and dynamical decoupling, here, we will show how it can be used to find an ideal decoupling sequence in the scenario where decoupling is required at a single fixed time t_{N+1} (as opposed to decoupling for all times).

To this end, we shall consider a sequence of unitary maps $\mathcal{V}_1, \dots, \mathcal{V}_N$ to be decoupling, if at the time t_{N+1} the state ρ'_s of the system is related to the input state ρ_s at t_0 by a known unitary

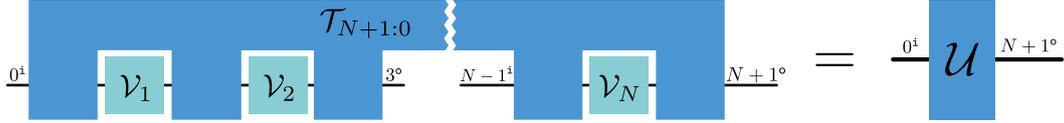


Figure 5.6: Goal of dynamical decoupling. By applying the correct sequence of unitary operations $\mathcal{V}_1, \dots, \mathcal{V}_N$, the resulting overall dynamics should become a known unitary – and thus correctable – dynamics \mathcal{U} .

map \mathcal{U} , *i.e.*, $\rho'_s = \mathcal{U}[\rho_s]$ for all initial system states ρ_s (see Fig. 5.6). In this case, as \mathcal{U} is known and reversible, the effect of the coupling between system and environment on the state of the system can be fully corrected for.

In order for dynamical decoupling to be operationally meaningful, must to assume that the initial state at t_0 is controllable by the experimenter, and thus, in what follows, we will consider process tensors with an additional open wire at t_0 , that, as already assumed in Eq. (5.31), constitutes a mapping from the controllable inputs $\{\rho_s, \mathcal{V}_1, \dots, \mathcal{V}_N\}$ to the final output state at time t_{N+1} . Its experimental reconstruction is achieved by measuring the output states for a basis $\{\rho_s^{(\mu)}\}_{\mu=1}^{d^2}$ of input states and a basis of sequences of unitaries, *i.e.*,

$$\mathbb{T}_{N+1:0}^U = \sum_{\vec{j}, \mu} \rho^{(\mu, \vec{j})'} \otimes \mathbf{v}_{\vec{j}}^* \otimes \mathbf{r}_{\mu}^* \quad (5.32)$$

where $\{\mathbf{r}_{\mu}\}_{\mu=1}^{d^2}$ is the dual set to $\{\rho^{(\mu)}\}_{\mu=1}^{d^2}$, $\rho^{(\mu, \vec{j})'}$ denotes the output state corresponding to the sequence $\{\rho_s^{(\mu)}, \mathcal{V}_{j_1}, \dots, \mathcal{V}_{j_N}\}$ of inputs, and $\{\mathbf{v}_{\vec{j}} = \mathbf{v}_{j_1} \otimes \dots \otimes \mathbf{v}_{j_N}\}$ are the duals to the sequences of unitary operations. Now, if a sequence of unitaries $\mathcal{Z}_1, \dots, \mathcal{Z}_N$ is a successful decoupling sequence, then

$$\mathbb{T}_{N+1:0}^U \star \mathcal{Z}_1 \star \dots \star \mathcal{Z}_N = \mathbb{U}, \quad (5.33)$$

where $\mathbb{U} \in \mathcal{B}(\mathcal{H}_{N+1}^o \otimes \mathcal{H}_0^i)$ is the Choi state of a unitary map (see Fig. 5.6). Hence, given the restricted process tensor of a process, it is then merely a numerical sampling problem to find such a successful decoupling sequence (if it exists for the chosen time steps). To explicitly illustrate this description of dynamical decoupling in terms of a restricted process tensor, we reiterate the shallow pocket model discussed in Ref. [127].

For this example, let $H_{se} = \frac{\xi}{2} \sigma_z^s \otimes \hat{x}$ be the total, time-independent Hamiltonian for a qubit (the system) coupled to a particle on a line (the environment), where \hat{x} is the position operator. The system is initially prepared in state $\rho_s(0)$ and uncorrelated with the environment, which we choose to be in state $\eta_e = |\Psi\rangle\langle\Psi|$, with $\langle x|\Psi\rangle = \sqrt{\frac{\gamma}{\pi}} \frac{1}{x+i\gamma}$ and $\gamma > 0$. The free evolution of the system state, *i.e.*, the evolution without intermediate local operations, is given by

$$\rho_s(t) = \begin{pmatrix} \rho_{00}(0) & \rho_{01}(0)e^{-g\gamma t} \\ \rho_{01}^*(0)e^{-g\gamma t} & \rho_{11}(0) \end{pmatrix} \quad \text{with} \quad \rho_{kl}(0) = \langle k|\rho_s(0)|l\rangle, \quad (5.34)$$

which constitutes a purely dephasing dynamics [127]. A possible experimental decoupling procedure could consist of a preparation of the initial system state $\rho_s(0)$, a free evolution of the system-environment state generated by H_{se} for a time Δt , a local unitary operation \mathcal{Z} , and,

finally, a tomography of the system state after another free evolution for a time Δt . Let $\mathbb{T}_{2:0}^U$ be the restricted process tensor for this experiment. We have

$$\mathcal{T}_{2:0}^U[\rho_s, \mathcal{Z}] = \text{tr}_e \{ \mathcal{U}_{\Delta t} [\mathcal{Z} [\mathcal{U}_{\Delta t} [\rho_s(0) \otimes |\Psi\rangle\langle\Psi|]]] \} = \mathbb{T}_{2:0}^U \star \rho_s \star Z \quad (5.35)$$

where $\mathcal{U}_{\Delta t}[\rho_{se}] = e^{-iH_{se}\Delta t} \rho_{se} e^{iH_{se}\Delta t}$ and we have omitted an identity map on the environment. A unitary map \mathcal{Z} that leads to successful decoupling would satisfy

$$\mathbb{T}_{2:0}^U \star Z = U, \quad (5.36)$$

where U is a unitary map. Having reconstructed the process tensor for this experimental scenario, one could numerically search for the best decoupling operation, for example by sampling unitaries \mathcal{Z} and checking if Eq. (5.36) is satisfied. Here, due to the computationally simple structure of the problem, this question can be answered analytically [127]. Choosing a unitary matrix $Z_{(a,b)} = a\sigma_x + b\sigma_y$, with $|a|^2 + |b|^2 = 1$, $a, b \in \mathbb{R}$, and corresponding unitary map $\mathcal{Z}_{(a,b)}[\rho_s] = Z_{(a,b)}\rho_s Z_{(a,b)}^\dagger$, we obtain

$$\mathbb{T}_{2:0}^U \star Z_{(a,b)} = \bar{Z}_{(a,b)}, \quad (5.37)$$

where $\bar{Z}_{(a,b)}$ is the Choi state of the map $\bar{Z}_{(a,b)}[\rho_s] = Z_{(a,b)}^\dagger \rho_s Z_{(a,b)}$. Consequently, any local operation of the form $\mathcal{Z}_{(a,b)}$ decouples the system of interest from its environment for the given process. This can also be shown directly from the total Hamiltonian H_{se} [127]. However, if the se Hamiltonian is unknown, or computationally intractable, this is no longer an option. On the other hand, the restricted process tensor can be reconstructed without any knowledge of the underlying Hamiltonian, and allows for a numerical search of a sequence $\mathcal{Z}_{N:1}$ of unitary operations, such that Eq. (5.33) holds. Even if perfect decoupling is not possible, this procedure still allows one to find the best implementable decoupling sequence, where ‘best’ in this case means that the resulting overall map is closest to the set of unitary operations.

In contrast, a description of the dynamics in terms of a master equation would fail to reproduce these results (as shown in Ref. [127]). The master equation of the free evolution of the system is given by

$$\dot{\rho}_s(t) = \mathbb{L}[\rho_s(t)] = -g \frac{\gamma}{4} [\sigma_z, [\sigma_z, \rho_s(t)]] , \quad (5.38)$$

where \mathbb{L} is the Lindbladian of the time evolution. The resulting time evolution of ρ_s is then given by $\rho_s(t) = e^{\mathbb{L}t}[\rho_s(0)]$. Naïvely, when implementing an intermediate local operation $\mathcal{Z}_{(a,b)}$ after free evolution for a time Δt , one might assume that the evolution of the system state from $t = 0$ to $t = 2\Delta t$, including the unitary operation in the middle would be described by

$$\rho_s(2\Delta t) = \left(e^{\mathbb{L}\Delta t} \circ \mathcal{Z}_{(a,b)} \circ e^{\mathbb{L}\Delta t} \right) [\rho_s(0)] . \quad (5.39)$$

However, given that $e^{\mathbb{L}\Delta t}$ is a purely dephasing channel, we conclude that $e^{\mathbb{L}\Delta t} \circ \mathcal{Z}_{(a,b)} \circ e^{\mathbb{L}\Delta t}$ cannot be equal to a unitary transformation, and Eq. (5.39) is thus not the correct description of the dynamics with intermediate intervention. As soon as the dynamics are non-Markovian, standard master equations fail to capture the influence that local operations at intermediate time steps have on the dynamics of the system of interest [80].

This example, in turn, also provides further proof for our statement that there are CP-divisible dynamics that are not Markovian. Master equations of Lindblad type lead to CP dynamics that satisfy a semi-group property. Thus, as in this case all the involved maps are invertible, the dynamics is iCP-divisible. It can also be checked straightforwardly, that it is oCP-divisible. However, the system can be decoupled by means of intermediate operations, and the dynamics is thus non-Markovian [80, 127].

5.6 CLASSICAL RESOURCES

Throughout this chapter, we have frequently alluded to the fact that classical stochastic processes are a particular set of restricted process tensors, namely those pertaining to situations where the experimenter only has projective measurements in a fixed (classical) basis at their disposal to probe the process. Evidently, limitation to projective measurements does not magically make a process classical. Rather, the set of classical processes with respect to a given measurement basis form a subset of all possible processes. Here, building on the intuition we gained in the previous sections for process tensors and limited control, we will define explicitly, what we mean by a classical process, and provide their full characterization. In this sense, this section is an extension of the discussion of classical processes and the breakdown of Kolmogorov conditions that we carried out in Ch. 3, but with the advantage, that we are now in a position to make full use of the process tensor formalism, and can thus derive all structural properties of classical processes.

When probing classical processes (without interventions), experimental control at every time t_α is limited to projective measurements in a fixed orthogonal basis, denoted by $\{|x\rangle_\alpha\}_{x_\alpha=1}^d$. Now, performing measurements at times $t_\alpha \in \Lambda_K$ in this fixed basis, the experimenter collects joint probability distributions $\mathbb{P}_{\Lambda_K}(\mathbf{x}_{\Lambda_K})$, where \mathbf{x}_{Λ_K} is a sequence of outcomes at times Λ_K . If the underlying process is classical, then these probability distributions satisfy the Kolmogorov properties we discussed in Ch. 3. As satisfaction of these conditions is the *only*⁵ defining property of a classical process, following Ref. [80] we will define classical processes accordingly:

DEFINITION 5.1: *A process is classical on times Λ_K , if its joint probability distributions obtained from sequences of projective measurements in the fixed basis $\{|x_\alpha\rangle_{x_\alpha=1}^d$ satisfy Kolmogorov conditions, i.e., if*

$$\mathbb{P}_{\Lambda_k}(\mathbf{x}_{\Lambda_k}) = \sum_{\Lambda_K \setminus \Lambda_k} \mathbb{P}_{\Lambda_K}(\mathbf{x}_{\Lambda_K}), \quad \forall \Lambda_k \subseteq \Lambda_K, \quad (5.40)$$

where $\sum_{\Lambda_K \setminus \Lambda_k}$ denotes the sum over all times in $\Lambda_K \setminus \Lambda_k$ and \mathbf{x}_{Λ_k} is the restriction of \mathbf{x}_{Λ_K} to times Λ_k .

If Kolmogorov conditions hold for the joint probability distributions that an experimenter deduces, then there is a possible – albeit potentially complicated – classical explanation for the experimental observations. If they do not hold, then the experiment can only be described by

⁵ Up to some technical regularity properties of the joint probability distributions.

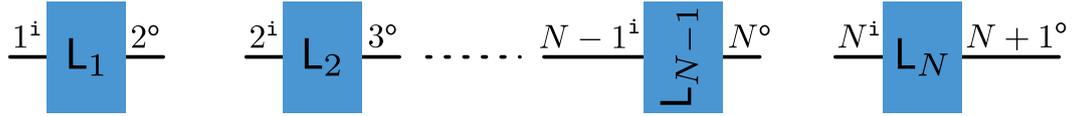


Figure 5.7: Markovian process that leads to classical statistics. If the maps $\{L_\alpha\}$ satisfy the NCGD condition, then a classical experimenter, *i.e.*, an experimenter that can only prepare states that are diagonal in a fixed orthogonal basis, and only performs projective measurements in said basis, cannot distinguish this process from a classical one.

employing quantum mechanics.⁶ Evidently, Def. (5.1) is purely operational in nature. Basically, it asserts that ‘if it looks like a classical process, swims like a classical process, and quacks like a classical process then it *is* a classical process’. Consequently, there can be processes that are implemented by means of quantum resources, but will be indistinguishable to an experimenter that can only perform projective measurements in the fixed basis $\{|x_\alpha\rangle\}$.

While one might misunderstand this fact as a weakness of the above definition, we should rather consider it a strength. It provides us with an experimentally probable criterion for classicality, that does not rely on inaccessible entities – like, for example, the unitaries governing the underlying process – and thus offers a clear-cut delineation between classical and genuinely quantum processes. Additionally, it does not *a priori* define classicality based on coherence, a mathematical property, that has been claimed to be a genuinely quantum trait. In this sense, this definition is similar to the definition of Markovianity in quantum mechanics, that we gave above, which relied on logical conditional independence of observable joint statistics, instead of the satisfaction of some mathematical properties, like. *e.g.*, CP-divisibility.

5.6.1 Classical Markovian Dynamics

Having our definition of classical processes at hand, we can now answer the question regarding what the set of classical process tensors looks like. To this end, in a first step, it is instructive to reiterate the results of Ref. [80], where this question has been answered for the case of Markovian dynamics. As discussed, a Markovian process tensor is of the form

$$\mathbb{T}_{N+1:1}^{\text{Markov}} = L_N \otimes L_{N-1} \otimes \cdots \otimes L_1, \quad (5.41)$$

where each of the maps $L_\alpha \in \mathcal{B}(\mathcal{H}_{\alpha+1}^o \otimes \mathcal{H}_\alpha^i)$ is a CPTP map, and for simplicity of the discussion, we assume that the process tensor starts with an open wire (see Fig. 5.7). Now, an experimenter that only has access to classical resources to probe the process, can prepare an initial state at t_1 , that is diagonal in the fixed basis $\{|x_1\rangle\}_{x_1=1}^d$ and perform projective measurements in this basis at later times. For ease of notation, and as we will encounter it frequently in what follows, we will denote the set of all quantum states that are diagonal in the classical basis $\{|x_\alpha\rangle\}$, *i.e.*, the set of all *incoherent* states by Θ :

$$\Theta := \left\{ \rho \mid \rho = \sum_{x_\alpha=1}^d p_{x_\alpha} |x_\alpha\rangle \langle x_\alpha|, \quad \text{with } p_{x_\alpha} \geq 0 \text{ and } \sum_{x_\alpha=1}^d p_{x_\alpha} \leq 1 \right\}, \quad (5.42)$$

⁶ Importantly, this only holds true, if the experimenter can be sure that they only perform projective measurements, and do not perform active interventions. Otherwise, as we have seen in Ch.3, Kolmogorov conditions can also be violated in classical physics.

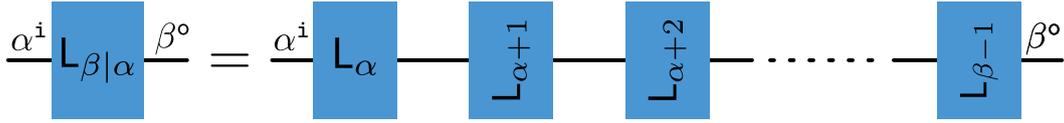


Figure 5.8: Definition of the maps $L_{\beta|\alpha}$. The maps $L_{\gamma|\beta}$ are defined in an analogous way.

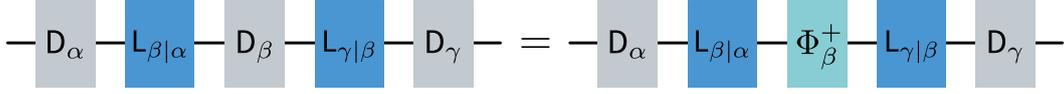


Figure 5.9: NCGD dynamics. Dephasing maps at t_1 and t_{N+1} imply that only incoherent states can be fed into the process, and only projective measurements in the classical basis can be performed. For these available experimental resources, the identity map at t_α is indistinguishable from the completely dephasing map.

where, for generality, we also allow for subnormalized states. If a state is not diagonal in the fixed basis, we shall consider the off-diagonal terms to be its *coherences*. Evidently, if none of the maps $\{L_\alpha\}$ in Eq. (5.41) can create coherences, *i.e.*, if $\rho \in \Theta$ implies $\mathcal{L}_\alpha[\rho] \in \Theta$, then the resulting dynamics can be understood as a time evolution of a probability vector (the diagonal of the state of the system), and it will look classical to an experimenter that can only probe it by classical means.

Quantum maps that cannot create coherences out of incoherent states are called *maximally incoherent operations* (MIO) in the literature and play an important role in the resource theory of coherence [204, 205]. While maximal incoherence of the maps $\{L_\alpha\}$ is sufficient for a Markovian process to be classical, it is not necessary. The process can still appear classical if the maps $\{L_\alpha\}$ can create coherences, as long as these coherences cannot be picked up at a later point by means of projective measurements in the classical basis. Put differently, if the process is classical, then none of the maps L_α creates coherences that can be transformed into populations, *i.e.*, enact changes of the diagonal entries of the state of the system, at a later time.

Specifically, let $\mathcal{L}_{\beta|\alpha} = \mathcal{L}_{\beta-1} \circ \cdots \circ \mathcal{L}_\alpha$ and $\mathcal{L}_{\gamma|\beta} = \mathcal{L}_{\gamma-1} \circ \cdots \circ \mathcal{L}_\beta$ be the dynamics from t_α to t_β and from t_β to t_γ , respectively (see Fig. 5.8). Then, we have the following definition, due to Smirne *et al.* [80]:

DEFINITION 5.2: A Markovian dynamics is considered *non-coherence-generating-and-detecting* (NCGD), if for all $t_\alpha < t_\beta < t_\gamma$ we have

$$\mathcal{D} \circ \mathcal{L}_{\gamma|\beta} \circ \mathcal{D} \circ \mathcal{L}_{\beta|\alpha} \circ \mathcal{D} = \mathcal{D} \circ \mathcal{L}_{\gamma|\beta} \circ \mathcal{L}_{\beta|\alpha} \circ \mathcal{D}, \quad (5.43)$$

where $\mathcal{D}[\rho] = \sum_x \langle x | \rho | x \rangle |x\rangle \langle x|$ is the completely dephasing map in the classical basis at the respective times (t_α , t_β , and t_γ).

Intuitively, if the dynamics is NCGD, the presence of coherences – if they exist – cannot be detected at any point by means of measurements in the classical basis. Equivalently, Eq. (5.43) can be read as stating, that if the dynamics is NCGD, then the completely dephasing map \mathcal{D}_β in between $\mathcal{L}_{\gamma|\beta}$ and $\mathcal{L}_{\beta|\alpha}$ is indistinguishable from an identity map for a classical observer (see Fig. 5.9 for a graphical representation). This understanding of Eq. (5.43) coincides exactly with the Kolmogorov conditions, that require that summing over measurement outcomes – which is the same as applying the completely dephasing map – is indistinguishable from not

performing a measurement. Consequently, Markovian dynamics is classical in the sense of Def. 5.1 iff it is NCGD [80].

In anticipation of the generalization to non-Markovian dynamics, it is insightful to rephrase the definition of NCGD in terms of Choi states. A Markovian dynamics is NCGD if for all $t_\alpha < t_\beta < t_\gamma$, we have

$$D_\gamma \star L_{\gamma|\beta} \star D_\beta \star L_{\beta|\alpha} \star D_\alpha = D_\gamma \star L_{\gamma|\beta} \star \Phi_x^+ \star L_{\beta|\alpha} \star D_\alpha, \quad (5.44)$$

where $L_{y|x}$ is the Choi state of $\mathcal{L}_{y|x}$, and D_x and Φ_x^+ are the Choi states of the completely dephasing map and the identity map at t_x , respectively. Importantly, Eq. (5.44) completely determines the structural requirements that the matrices $L_{y|x}$ have to satisfy to be NCGD (in detail, these requirements correspond to the ones mentioned in Lemma 1 of [80] for the maps $\mathcal{L}_{y|x}$). One advantage of the quantum comb formalism lies in the fact that similar structural requirements can be readily derived in the case of non-Markovian dynamics. Before we do so, it is insightful to briefly discuss the above result about the connection between classicality and NCGD.

First and foremost, the results of [80] establish a clear link between coherence and classicality, that stems from logical reasoning, rather than from the assumption that coherence is *a priori* an inherently quantum trait. Classical dynamics does *not* imply that there is no coherence created at any time in the process [80]. Rather, if any coherence is created, its presence cannot be detected by the classical means available to the experimenter. In particular, maximally incoherent operations form a strict subset of maps that lead to NCG dynamics, and thus – in the Markovian case – result in classical statistics.

Secondly, the reasoning that led to the connection between coherence and classicality cannot be straightforwardly employed in the non-Markovian case; for the case of Markovian dynamics, the classicality of a process can be decided on the level of quantum channels, as all higher order joint probability distributions can be constructed from two point correlations, which, in turn, are encapsulated by the maps $L_{y|x}$. Consequently, in the memoryless case, the relation between coherences in the system of interest, and the classicality of the underlying process is direct, since there are no non-classical correlations between the system and its environment that permeate memory effects. As we will see below, neither of these properties holds in the non-Markovian case. In particular, even if the state of the system is diagonal in the classical basis at each time, the process does not necessarily satisfy Kolmogorov conditions if the underlying dynamics is non-Markovian.

5.6.2 Maximally Incoherent Combs (MIC)

Before comprehensively characterizing the set of classical non-Markovian combs, it is insightful to analyze the interplay of coherence and classicality for non-Markovian processes. While in the Markovian case, their relation is relatively direct, it turns out to be more subtle when memory effects play a role.

As we discussed, in the memoryless setting, if the maps $\{\mathcal{L}_\alpha\}$ are MIO, then the underlying dynamics is classical. The concept of maximally incoherent operations can straightforwardly be

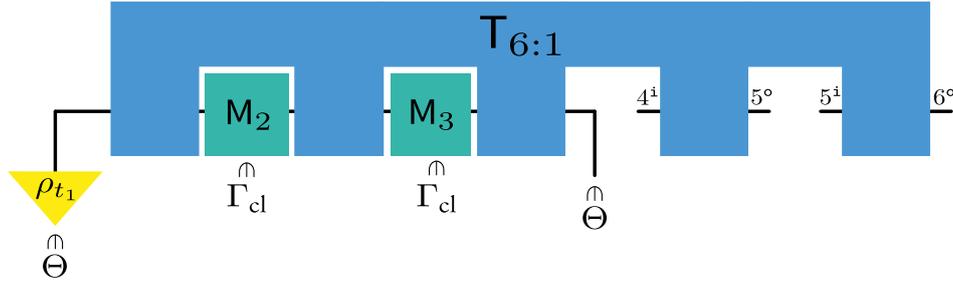


Figure 5.10: Maximally incoherent comb. If all the inputs to the comb (here, $T_{6:1}$) are classical, and the system state at every time is classical, then no coherence can be ‘extracted’ from the comb, and it is maximally incoherent.

generalized to the non-Markovian case. We will consider a comb maximally incoherent, if it cannot create coherent system states for any classical control operations (including interventions), that are ‘plugged’ into its slot.

Specifically, as before, let Θ be the set of (potentially subnormalized) incoherent states in the fixed computational basis $\{|x_\alpha\rangle\}$ and let the experimenter have access to purely classical operations only. For example, they could perform projective measurements in the computational basis, or perform classical interventions. Consequently, in terms of Choi states, the set $\tilde{\Gamma}_{\text{cl}}$ of fully classical operations is given by

$$\tilde{\Gamma}_{\text{cl}} = \left\{ \sum_{x_\alpha, y_\alpha} p_{y_\alpha|x_\alpha} |y_\alpha\rangle\langle y_\alpha| \otimes |x_\alpha\rangle\langle x_\alpha| \right\}, \quad (5.45)$$

where $p_{y_\alpha|x_\alpha} \geq 0$, $\sum_{y_\alpha} p_{y_\alpha|x_\alpha} \leq 1$, and $|y_\alpha\rangle\langle y_\alpha| \otimes |x_\alpha\rangle\langle x_\alpha|$ is the Choi state of an operation that prepares state $|y_\alpha\rangle\langle y_\alpha|$ upon measuring the system in state $|x_\alpha\rangle\langle x_\alpha|$. As it should classically also be allowed to *not* perform an operation, we add the identity operation Φ^+ to the set of classical operations, *i.e.*, $\Gamma_{\text{cl}} = \tilde{\Gamma}_{\text{cl}} \cup \{\Phi^+\}$.⁷ We then have the following natural extension of the definition of MIO to the non-Markovian case:

DEFINITION 5.3 (MAXIMALLY INCOHERENT COMB (MIC)): *The comb/process tensor of a process defined on times $\{t_{N+1}, \dots, t_1\}$ is called a maximally incoherent comb (MIC) if for any incoherent initial state $\rho_1 \in \Theta$ at t_1 , any time $t_\beta \in \{t_{N+1}, \dots, t_1\}$, and any sequence of classical operations $\{M_\alpha \in \Gamma_{\text{cl}}\}$ that were performed at times $\{t_{\beta-1}, \dots, t_2\}$, the resulting state ρ_β at t_β is incoherent, *i.e.*, $\rho_\beta \in \Theta$.*

A graphical representation of this definition can be found in Fig. 5.10. Intuitively, a comb is an MIC if no coherence can be extracted from it by means of classical resources/operations. Quite obviously, for one time step, when the comb $T_{N+1:1}$ is simply a quantum channel L_1 , the definition of MIC coincides with the definition of maximally incoherent operations, and, consequently, in the Markovian case, maximally incoherent combs satisfy Kolmogorov conditions. However, somewhat counterintuitively, in the non-Markovian case, an MIC does not necessarily lead to joint probability distributions that satisfy Kolmogorov conditions. To see this, consider the following example:

⁷ Defined in this way, the set Γ_{cl} is not convex, which is somewhat unsatisfactory from a resource theoretic point of view. This ‘problem’ could be alleviated by replacing Γ_{cl} by its convex hull. Such a replacement would not affect any of the subsequent arguments, though.

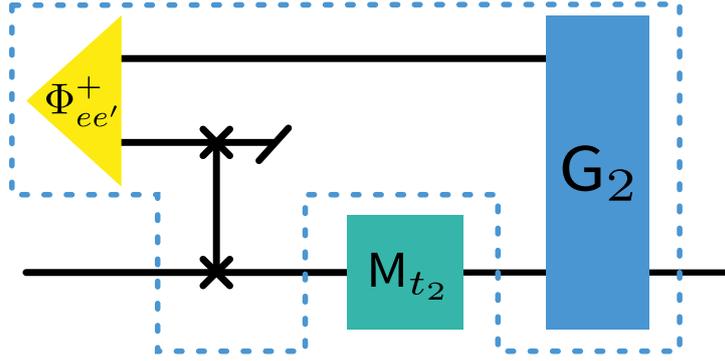


Figure 5.11: Non-classical MIC. The state of the system is classical at every step of the process (maximally mixed at t_2 , and a classical mixture of projectors onto $|0\rangle$ and $|1\rangle$ at t_3), so the corresponding process tensor (blue outline) is MIC. However, the inferred statistics do not satisfy the Kolmogorov condition.

EXAMPLE 5.1 (MIC $\not\Rightarrow$ KOLMOGOROV CONDITIONS): *If the dynamics is described by an MIC, the state of the system does not display any coherences at any point in time if the inputs to the process are classical. Consider the following circuit (see also Fig. 5.11): Let the initial state of the environment be given by a (normalized) maximally entangled two qubit state $\Phi_{ee'}^+ \in \mathcal{B}(\mathcal{H}_e \otimes \mathcal{H}_{e'})$ and let the system be a qubit.⁸ Between t_1 and t_2 the state of the system is swapped with one half (corresponding to the label e) of the environment state, i.e., the state of the system at t_2 is $\rho_2 = \mathbb{1}/2$, independent of the system state that was initialized at t_1 . Between t_2 and t_3 , the system and the second half of the environment (the part labeled by e') undergo a CPTP map G_2 (which could – in principle – be dilated to a unitary, but for conciseness, we restrict ourselves to the relevant part of it), that yields output $|0\rangle\langle 0|$ on the system, when system and environment are in the state Φ^+ , and $|1\rangle\langle 1|$ otherwise. Written in its Choi form, we have*

$$G_2 = |0\rangle\langle 0| \otimes \Phi^+ + |1\rangle\langle 1| \otimes (\mathbb{1} - \Phi^+) . \quad (5.46)$$

It is easy to check that G_2 is indeed CPTP, and the state of the system at t_3 is a convex mixture of $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$, i.e., the resulting comb is MIC.⁹ However, it does not satisfy the Kolmogorov condition. To see this, consider the probabilities for a measurement in the computational basis at t_3 , with no operation performed at t_2 and a random default state fed into the process at t_1 . In this case, the system-environment state before G_2 is equal to Φ^+ , which means that we have $\rho_3 = |0\rangle\langle 0|$. Consequently, the measurement in the computational basis at t_3 yields the probabilities

$$\mathbb{P}(0, t_3) = 1 \quad \text{and} \quad \mathbb{P}(1, t_3) = 0 . \quad (5.47)$$

On the other hand, performing a measurement at t_2 and discarding the outcomes, amounts to performing the completely dephasing map D_2 . After this map, i.e., right before G_2 , the system-environment state is of the form

$$\rho_2^{se'} = \frac{1}{2} \sum_{i=1}^2 |i\rangle\langle i| \otimes |i\rangle\langle i| = \frac{1}{2} (\Phi^+ + \Phi^-) , \quad (5.48)$$

⁸ Importantly, even though we denote them by the same symbol, the environment state should not be confused with the Choi state of an identity map.

⁹ For this example, we could even allow for non-classical states and operations at t_1 and t_2 , respectively, and still not see any coherence in the state of the system at t_3 .

where $\Phi^- = (\sigma_z \otimes \mathbb{1})\Phi^+(\sigma_z \otimes \mathbb{1})$ is a Bell state. Consequently, in this case the final system state ρ_3 at time t_3 is of the form $\rho_3 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$. Finally, the obtained probabilities, given that the completely dephasing maps \mathcal{D}_2 was performed at t_2 are of the form

$$\begin{aligned} \mathbb{P}^{\mathcal{D}_2}(0, t_3) &= \sum_{x_2} \mathbb{P}(0, t_3; x_2, t_2) = \frac{1}{2} \\ \text{and } \mathbb{P}^{\mathcal{D}_2}(1, t_3) &= \sum_{x_2} \mathbb{P}(1, t_3; x_2, t_2) = \frac{1}{2}, \end{aligned} \quad (5.49)$$

which does not coincide with (5.47). Even though the state of the system is incoherent at every time, i.e., appears to be classical, the multi-time statistics do not satisfy the Kolmogorov condition. Another case of non-Markovian dynamics that does not allow for the extraction of coherences, but is non-classical nonetheless, can be found in [80]; here, the examples for non-Markovian dynamics that are NCGD but not classical also constitute an *ante litteram* example of non-classical MICs [206].

Unlike in the Markovian case, where the set of operations that do not create coherences is contained in the set of NCGD dynamics, and as such leads to statistics that satisfy Kolmogorov conditions, this is not the case when memory effects play a non-negligible role. Put differently, when memory effects are present, the connection between coherences and the classicality of the underlying dynamics is not straightforward anymore.

In general, it is *both* the coherences of the system state, *as well as* the non-classical correlations between the system and its environment that can lead to non-classical behavior. Intuitively, while the completely dephasing map leaves the system unchanged if the underlying dynamics is MIC, it does not necessarily leave the overall *system-environment* state invariant. Specifically, if the comb is an MIC, then at every time t_α we have $\mathcal{D}_\alpha[\rho_\alpha^s] = \mathcal{I}[\rho_\alpha^s]$, but not necessarily $(\mathcal{D}_\alpha \otimes \mathcal{I})[\rho_\alpha^{se}] = \mathcal{I}[\rho_\alpha^{se}]$ for all times t_α ; Even if the state of the system is classical at all times, the completely dephasing map can still be *invasive*, and as such, its influence can potentially be differentiated from the ‘do-nothing’ operation at some later point in time. While it is not necessary for the satisfaction of Kolmogorov conditions, that the action of the completely dephasing map and the identity map coincide, it is sufficient, and as such directly relates the classicality of dynamics to *quantum discord* [207, 208]:

LEMMA 5.1: *If the system-environment state has the form*

$$\rho_\alpha^{se} = \sum_{m=1}^d p_\alpha^m \Pi_\alpha^m \otimes \eta_\alpha^m, \quad (5.50)$$

at every time t_α , where $\{p_\alpha^m\}$ are probabilities that add to unity, $\{\Pi_\alpha^m\}$ are orthogonal projectors in the classical basis, and $\{\eta_\alpha^m\}$ are quantum states on the environment, then the underlying dynamics is classical, i.e., satisfies the Kolmogorov conditions.

Before we prove this statement, it is insightful to discuss the form of the *se*-states defined in Eq. (5.50). States of this form have vanishing quantum discord [207–210], i.e., they do not display *quantum* correlations between the system and the environment. For a general zero-discord state, the set $\{\Pi_{t_j}^m\}$ could be any set of mutually orthogonal projectors, and the state is considered classical, since there exists a measurement with d perfectly distinguishable outcomes that overall leaves the total state undisturbed [207, 210] (see also the proof below). As we only

allow for measurements in a fixed basis, it is not sufficient for us to have a system-environment state that has vanishing discord, but it has to vanish in the correct basis, *i.e.*, the one in which the experimenter’s measurements act. Consequently, we will call states of the form (5.50) discord zero states *with respect to the classical basis*, keeping in mind that, in general, quantum discord is not basis dependent, but obtained by a minimization procedure over all possible measurement scenarios. Throughout the remainder of this chapter, whenever we consider a state to be of discord zero, we will always implicitly mean that it can be represented as (5.50) with the projectors being diagonal in the classical basis. Importantly, this basis dependence mirrors the basis dependence of coherence, which is also always defined with respect to a fixed classical basis.

Proof. Let the *se*-state at every time t_α be of the form

$$\rho_\alpha^{se} = \sum_{m=1}^d p_\alpha^m \Pi_\alpha^m \otimes \eta_\alpha^m, \quad (5.51)$$

For this state, the completely dephasing map on s has the same effect as the ‘do-nothing’ channel, *i.e.*,

$$\mathcal{D}_\alpha \otimes \mathcal{I}_e \left[\sum_m p_\alpha^m \Pi_\alpha^m \otimes \eta_\alpha^m \right] = \mathcal{I}_s \otimes \mathcal{I}_e \left[\sum_m p_\alpha^m \Pi_\alpha^m \otimes \eta_\alpha^m \right]. \quad (5.52)$$

This equivalence of the action of the identity map and the completely dephasing map directly implies that a measurement in the classical basis overall does not disturb the system-environment state. Consequently, if the system-environment state has vanishing discord with respect to the classical basis at all times, the resulting statistics satisfy the Kolmogorov conditions. \square

While the above lemma is unsurprising, it sheds light on the properties that a general non-Markovian dynamics has to satisfy to appear classical. We will see that classical non-Markovian processes mirror Markovian ones, but with coherence replaced by discord. We have already seen that maps L_α that do not produce coherences lead to classical Markovian dynamics. Analogously, system-environment unitary dynamics that cannot create discord (with respect to the classical basis), will evidently induce classical non-Markovian dynamics. However, in the Markovian case, it is not necessary, that all maps L_α are MIOs for the dynamics to be classical, but rather, the dynamics has to be such that if coherences are created at some point, they cannot be picked up by classical measurements at some later point. In the same vein, it is not necessary for classical non-Markovian dynamics that all system-environment unitaries are only non-discord creating, but it suffices that discord that is created cannot be detected at any later time. In clear analogy to the Markovian case, we shall dub such dynamics non-discord-generating-and-detecting (**NDGD**). We will define the set of NDGD maps in the following section.

Importantly, with our understanding of discord in mind, incoherent system states can be considered to be discord zero states of the form Eq. (5.50), with a trivial environment, and NCGD dynamics are a special case of NDGD dynamics, for the case where no memory is propagated. As we will see now, NDGD dynamics naturally extend the concept of classicality to the non-Markovian case, and dynamics display classical statistics if they are NDGD.



Figure 5.12: Concise dilation of an open process. By absorbing the initial state and the final trace into the respective CPTP maps, every open process can be represented as a sequence of CPTP maps \mathcal{G}_α (under the assumption of an initially uncorrelated system-environment state). Once these maps are known, all joint probability distributions can be computed.

5.6.3 Non-Discord-Generating-and-Detecting Dynamics and Classical Processes

In the Markovian case, classicality of the process can be decided on the level of the CPTP maps as in the absence of memory all higher order probability distributions can be obtained from the initial state ρ_1 at time, and the transformation maps $L_{\beta|\gamma}$ and $L_{\gamma|\beta}$. It suggests itself to employ this intuition in the non-Markovian case. As we have seen in the previous chapter, every non-Markovian dynamics can be dilated to a concatenation of an initial (potentially correlated) system-environment state, and unitary se -dynamics, interspersed by the operations on the system alone, performed at times $\{t_\alpha\}$. Equivalently, to make notation more concise, we can consider this open dynamics as a concatenation of *CPTP maps*, interspersed by the operations on the system alone, performed at times $\{t_\alpha\}$, and the respective CPTP maps are connected by wires that can transport memory (see Fig. 5.12 for reference). In what follows, we will denote these CPTP maps by \mathcal{G}_α to clearly distinguish them from the Markovian case. On this dilated level, the dynamics is Markovian – there are no additional external wires can carry memory forward – and all higher order joint probability distributions could be built up when the individual CPTP maps \mathcal{G}_α and the initial state of the environment ρ_1^{se} are known. For simplicity, and easier connection to the Markovian case, in what follows we shall assume that the initial system environment state at time t_1 is uncorrelated. A generalization to correlated initial states is straightforward, and rather a notational than a conceptual obstacle. With this, we can define NDGD dynamics:

DEFINITION 5.4 (NDGD): *An open dynamics on times $\{t_\alpha\}_{\alpha=1}^{N+1}$ with CPTP maps $\{\mathcal{G}_\alpha\}$ is called non-Discord-generating-and-detecting (NDGD) if it satisfies*

$$D_\alpha \star G_{\beta|\alpha} \star D_\beta \star G_{\gamma|\beta} \star D_\gamma = D_\alpha \star G_{\beta|\alpha} \star \Phi_\beta^+ \star G_{\gamma|\beta} \star D_\gamma \quad (5.53)$$

for all $t_1 < t_\alpha < t_\beta < t_\gamma \leq t_{N+1}$, where $G_{y|x}$ is the Choi state of the CPTP map $\mathcal{G}_{y|x} = \mathcal{G}_{y-1} \circ \dots \circ \mathcal{G}_{x+1} \circ \mathcal{G}_x$ and D_x are completely dephasing maps acting on the system.

A graphical representation of this definition can be found in Fig. 5.13. Formally, it is equivalent to the definition of NCGD dynamics, with the difference, that the involved intermediary maps between times are now the system-environment maps, instead of the CPTP maps $L_{y|x}$ that acted on the system alone.

Analogously to the case of NCGD, an NDGD dynamics, *i.e.*, a dynamics that satisfies Eq. (5.53), cannot create discord (with respect to the classical basis) that can be detected at

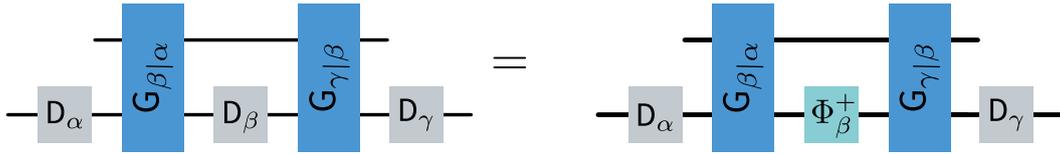


Figure 5.13: NDGD. From the perspective of a classical observer, the identity map at any time t_j cannot be distinguished from the completely dephasing map. Any discord (with respect to the classical basis) that is present in the initial se -state, or created by the system-environment unitaries, cannot be detected by a classical observer.

a later time by means of classical measurements. Or, equivalently, an experimenter that can only perform measurements in the classical basis, cannot distinguish between a completely dephasing map and an identity map. As such, it is the natural extension of NCGD to the non-Markovian case. Indeed, for the case of memoryless dynamics, Eq. (5.53) is the same as Eq. (5.44), and the definitions of NCGD and NDGD coincide. Somewhat unsurprisingly, we then have the following theorem:

THEOREM 5.1 (NDGD AND CLASSICALITY): *A non-Markovian dynamics is classical, if it is NDGD.*

The proof of this theorem can be found in App. C.4. In order to further elucidate the relation of discord and classicality in stochastic processes, it is insightful to discuss the proximity of Thm. 5.1 to the corresponding results in [80] for the Markovian case. Thm. 5.1 establishes the importance of the role of quantum discord for the classicality of non-Markovian dynamics. While in the memoryless case, it is coherence – or the impossibility of detection thereof – that makes a process classical; here, this role is played by discord, with the only difference that instead of describing the dynamics in terms of maps that are solely defined on the system of interest, we are forced to dilate the dynamics to a system-environment space, where it is Markovian. Consequently, the classicality of a process cannot be decided based on a master equation that describes the evolution of the system alone [80]. However, given a Hamiltonian that generates the corresponding system-environment dynamics, it can be decided if a dynamics is classical by checking the validity of Eq. (5.53).

It would be desirable, if NDGD were a sufficient *and* necessary criterion for the classicality of non-Markovian dynamics. However, this is not the case. NDGD as defined in Eq. (5.53) is a statement about the entire system-environment dynamics, and holds for any possible input state on the environment. However, by means of projective measurements on the system alone, one only has access to the system part, and the dynamics cannot be fully probed. Consequently, the criterion (5.53) will, in general, be too strong for a given experimental scenario. We provide an example of dynamics that is not NDGD, but nonetheless leads to classical dynamics in App. C.4.

This fact notwithstanding, Thm. 5.1 completes the results of [80] and provides as much of a clear connection between discord and classical processes as can be obtained. We now finish the chapter with a comprehensive analysis of temporal correlations that can be present in a classical process, and explicitly connect the reconstruction of classical processes to the framework of restricted process tensors.

5.6.4 Temporal Correlations in Classical Processes

To end this chapter, it remains to show, what ‘non-classical’ temporal correlations can still persist, despite the dynamics being considered classical. Here, ‘non-classical’ has to be put in quotation marks, as, up until this point, we have defined classicality in purely operational terms; if the statistics a classical observer obtains from an experiment look classical, we consider the underlying dynamics to be classical. We shall continue to take this distinctly operational perspective on the classicality of a process. Nonetheless, it is an interesting question to ask, what structural properties classical combs $\{\mathbb{T}_{N:1}\}$ satisfy. It turns out that most classical combs are not diagonal in the classical basis, and as such, for the remainder of this section, we shall consider the corresponding off-diagonal terms to correspond ‘non-classical’ temporal correlations. Put differently, in accordance with the overall theme of this chapter, in this section we analyze the set of correlations that an experimenter with only restricted, classical resources – projective measurements in a fixed orthogonal basis – is blind to. For lack of a better term, we shall continue to dub these correlations ‘non-classical’, but we will keep the quotation marks to clearly distinguish this algebraic notion of ‘non-classicality’ from the operationally well-defined one that we advocated for in the previous sections.

To give meaning to the idea of ‘non-classical’ temporal correlations, we return to the proof of the classical Kolmogorov extension theorem, that we provided in App. A.3. There, we used the fact that for any classical joint probability distribution $\mathbb{P}_{N:1}(x_{j_N}, \dots, x_{j_1})$ on times $\{t_1, \dots, t_N\}$, we can construct a corresponding process tensor

$$\mathbb{T}_{N:1} = \sum_{j_N, \dots, j_1} \mathbb{P}_{N:1}(x_{j_N}, \dots, x_{j_1}) \mathbb{1}_N^{\dagger} \otimes |x_{j_N}\rangle\langle x_{j_N}| \otimes \dots \otimes \mathbb{1}_1^{\dagger} \otimes |x_{j_1}\rangle\langle x_{j_1}|, \quad (5.54)$$

where, for simplicity, we assume that $\mathbb{T}_{N:1}$ has no open input and output wires. Evidently, the process tensor defined in Eq. (5.54) is diagonal in the classical product basis, and yields the correct probabilities when applied to a sequence of classical projective measurements:

$$\text{tr} \left[\left(\mathbb{P}_{x_{j_N}} \otimes \dots \otimes \mathbb{P}_{x_{j_1}} \right) \mathbb{T}_{N:1} \right] = \mathbb{P}_{N:1}(x_{j_N}, \dots, x_{j_1}), \quad (5.55)$$

where $\mathbb{P}_{x_{j_\alpha}} = |x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|$ is the Choi state of a projective measurement in the fixed basis at time t_α with outcome x_{j_α} . Additionally, we can see that the action of the ‘do-nothing’ channel at any time t_α coincides with the action of a completely dephasing map, and as such, this comb satisfies Kolmogorov conditions.

Now, there are a lot of terms $\chi_{N:1}$ that could be added to $\mathbb{T}_{N:1}$ defined in Eq. (5.54) without changing the resulting statistics when the process is probed by projective measurements in the classical basis. In particular, any *proper* process tensor

$$\tilde{\mathbb{T}}_{N:1} = \mathbb{T}_{N:1} + \chi_{N:1}, \quad (5.56)$$

with

$$\text{tr} \left[\left(\mathbb{P}_{x_{j_N}} \otimes \dots \otimes \mathbb{P}_{x_{j_1}} \right) \chi_{N:1} \right] = 0, \quad \forall \mathbb{P}_{x_{j_N}} \otimes \dots \otimes \mathbb{P}_{x_{j_1}} \quad (5.57)$$

will still satisfy Eq. (5.55), *i.e.*, lead to the same classical statistics. Unsurprisingly, this situation is reminiscent of the case of restricted process tensors we discussed in the previous sections.

Restricted process tensors are well-defined on the span of available operations, but cannot make assertions about operations that lay in the orthogonal complement. Analogously, here, projective measurements in the classical basis are blind to correlations that lie outside their span. We will denote the orthogonal complement of the span of sequences in the classical basis, *i.e.*, the set of $\chi_{N:1} \in \mathcal{B}(\mathcal{H}_N^i \otimes \cdots \otimes \mathcal{H}_1^o)$ that satisfy Eq. (5.57) by $\mathcal{Q}^{(N)}$.

While the addition of a term $\chi_{N:1} \in \mathcal{Q}^{(N)}$ to a given process tensor does not change the observed statistics when the process is probed at times $\{t_1, \dots, t_N\}$, it might still lead to a violation of Kolmogorov conditions. To see this clearly, let us consider for a moment the case of three times $\{t_1, t_2, t_3\}$. The corresponding process tensor is of the form

$$T_{3:1} = \sum_{j_3, j_2, t_1} \mathbb{P}_3(x_{j_3}, x_{j_2}, x_{j_1}) \mathbb{1}_{t_3}^i \otimes |x_{j_3}\rangle\langle x_{j_3}| \otimes \mathbb{1}_{t_2}^i \otimes |x_{j_2}\rangle\langle x_{j_2}| \otimes \mathbb{1}_{t_1}^i \otimes |x_{j_1}\rangle\langle x_{j_1}| \quad (5.58)$$

$$+ \chi_{3:1}, \quad (5.59)$$

where $\mathbb{P}_3(x_{j_3}, x_{j_2}, x_{j_1})$ is the experimentally observable joint probability distribution and $\chi_{3:1} \in \mathcal{Q}^{(3)}$. Now, if the process is classical, then, for instance, a ‘do-nothing’ operation at time t_2 has to be indistinguishable from a completely dephasing map for a classical observer, *i.e.*,

$$T_{3:1} \star P_{x_{j_1}} \star \Phi_2^+ \star P_{x_{j_3}} - T_{3:1} \star P_{x_{j_1}} \star D_2 \star P_{x_{j_3}} = 0 \quad \forall P_{x_{j_1}}, P_{x_{j_3}}, \quad (5.60)$$

which, in turn, implies

$$\chi_{3:1} \star P_{x_{j_1}} \star \Phi_2^+ \star P_{x_{j_3}} - \chi_{3:1} \star P_{x_{j_1}} \star D_2 \star P_{x_{j_3}} := \chi_{3:1} \star P_{x_{j_1}} \star A_2 \star P_{x_{j_3}} = 0, \quad (5.61)$$

for all projective measurements $P_{x_{j_1}}, P_{x_{j_3}}$, where we have defined the matrix $A_2 = \Phi_2^+ - D_2 = \sum_{j_2 \neq k_2} |j_2\rangle\langle k_2| \otimes |j_2\rangle\langle k_2|$. Eq. (5.61) constitutes an additional requirement for the terms $\chi_{3:1}$, in order for the process tensor to be classical.

In general, this requirement has to hold for *any* subset $\Delta \subseteq \Delta_N := \{t_1, \dots, t_N\}$ of times. Setting $\bar{\Delta} = \Delta_N \setminus \Delta$, we can define the corresponding set \mathcal{S} of correlation terms as

$$\mathcal{S} = \{ \chi_{N:1} \in \mathcal{B}(\mathcal{H}_N^i \otimes \cdots \otimes \mathcal{H}_1^o) \mid \chi_{N:1} \star_{\alpha \in \Delta} A_\alpha \star_{\beta \in \bar{\Delta}} P_{x_{j_\beta}} = 0, \forall \Delta, P_{x_{j_\beta}} \}. \quad (5.62)$$

While this definition is somewhat clunky, it exactly expresses the intuition we built above algebraically; in particular, correlation terms $\chi_{N:1} \in \mathcal{S}$ cannot lead to violations of the Kolmogorov conditions. Finally, with this we can fully characterize the set of classical process tensors on the set Δ_N :

THEOREM 5.2 (CLASSICAL PROCESS TENSORS): *A process tensor $T_{N:1}$ defined on times Δ_N leads to classical statistics iff it is of the form*

$$T_{N:1} = \sum_{j_N, \dots, j_1} \mathbb{P}_N(x_{j_N}, \dots, x_{j_1}) \mathbb{1}_{t_N}^i \otimes |x_{j_N}\rangle\langle x_{j_N}| \otimes \cdots \otimes \mathbb{1}_{t_1}^i \otimes |x_{j_1}\rangle\langle x_{j_1}| + \chi_{N:1}, \quad (5.63)$$

where $\mathbb{P}_N(x_{j_N}, \dots, x_{j_1})$ is a joint probability distribution and $\chi_{N:1} \in \mathcal{Q} \cap \mathcal{S}$.

Admittedly, this characterization is not the poster child of simplistic beauty, but nonetheless provides a full characterization of classical process tensors on N times. Additionally, the trouble

in providing a simple and intuitively insightful characterization of the set of classical processes is in line with similar troubles in the characterization of classical Markovian processes [80], and as such not all that surprising. Despite the technical obstacles, the possibility of unambiguously defining the set of classical processes, demonstrates the versatility of the theory of restricted process tensors, that we have relied upon heavily in the above derivation

5.7 LIMITED RESOURCES – SUMMARY

In this chapter, we have first brought the framework of process tensors closer to experimental reality by analyzing its restriction to limited control scenarios, and then used the resulting insights to discuss the definition and theory of classical processes in light of this analysis. While general non-Markovian quantum dynamics can be unambiguously described and characterized experimentally if the experimenter has unlimited control, we have seen that the situation is more layered when this is not the case.

On the positive side, in the case of limited experimental control, it is still possible – independent of the existence of memory effects – to reconstruct a restricted process tensor, as long as the dimension of the space that is spanned by the available operations is known. The thusly obtained restricted process tensor contains the maximal amount of information about the process that can be inferred locally, based on the set of available operations. Additionally, restricted process tensors still satisfy partial containment properties. They can be applied to any operation that lies in the span of the available operations, and provide an operationally meaningful complete dynamical description of the underlying dynamics. Surprisingly, the set of operations a restricted process tensor can be applied to can exceed the experimentally available ones. For example, if the set of available manipulations coincides with the set of unitary maps, the reconstructed process-tensor can, *e.g.*, be applied to any sequence of unital operations.

On the negative side, restricted process tensors are in general not completely positive – or rather, complete positivity is not even a meaningful property for these objects – and do not allow one to make assertions on the nature of the underlying process based on their structural properties. Nonetheless, we saw that if one further local operation, a swap with an identically prepared system, is performable, or if the span of the set of performable operations contains operations that decouple the system from its environment, it is possible to construct operationally well-defined witnesses for initial system-environment correlations and the non-Markovianity of a process. We have seen how to make maximal use of the available operations in the construction of these witnesses and illustrated their applicability for two extremal cases: the set of unitary operations, where no information about the system can be inferred from the operation; and the set of projective measurements, where information about the system is obtained, but it collapses to a pure state in the process. In both cases, the reconstructed witnesses detect initial correlations, as well as the non-Markovianity of the underlying process.

The quality of these witnesses, *i.e.*, their ability to detect correlations, depends crucially on the dimension of the space that is spanned by the available local operations and the interplay between correlations and the total unitary dynamics. However, the conjecture suggests itself

that for any reasonable scenario, *i.e.*, a general total unitary dynamics and a set of performable local operations that is not ‘too small’, it is always possible to detect system-environment correlations by means of the experimentally realizable local operations. Total unitary dynamics that prevent correlations from local detection should be mere pathological examples.

Finally, we applied the theoretical apparatus we developed to the case of experimental control, where the experimenter can only perform projective operations in a fixed orthogonal basis. Here, we considered restricted process tensors through the lens of classical processes and provided an operationally clear-cut definition of said processes, and, banking on ideas from Ch. 3, developed a full characterization of classical process tensors. Furthermore, for comprehensiveness, we connected classicality of processes to the concept of quantum discord, thereby extending to their full completion corresponding results for Markovian dynamics and coherence.

Restricted process tensors bridge the gap between the theoretical description of non-Markovian quantum dynamics and experimental reality and offer a versatile toolbox to discuss many different experimental situations in the same language. Consequently, there are manifold of possible extensions and applications of the restricted process tensor framework.

On the applied side, it naturally suggests itself for the description of quantum control and dynamical decoupling. In order to simplify the calculations, we have demonstrated this for a time-independent total Hamiltonian. In practice, however, the restricted process tensor can be reconstructed experimentally for any conceivable total dynamics. Given the restricted process tensor, it is then simply a numerical sampling problem to find the optimal sequence of operations that steers the system as close as possible to a desired final state. Our framework is also flexible enough to describe decoupling experiments and it allows one to search for the sequence of local operations that comes closest to achieving decoupling at some fixed final time. While this is not the original aim of dynamical decoupling, it nonetheless provides a new perspective: if decoupling at selective points in time is sufficient, decoupling schemes based on restricted process tensors might prove more efficient and less error-prone than traditional schemes that rely on the implementation of decoupling cycles much faster than typical correlation times [152].

Even if perfect decoupling will in general not be possible for randomly chosen time steps t_n , the process tensor approach is nonetheless fruitful: it opens up an avenue to benchmarking the deviation from perfect decoupling for a given choice of time steps and translates the question of whether perfect decoupling is possible to an inversion problem of the process tensor for the underlying process.

On the more theoretical side, the results and concepts we developed for classical process tensors pave the way to a fully-fledged resource theory of non-classicality of non-Markovian processes.¹⁰ We have mentioned that the concept of MIO can be extended to the non-Markovian case. Analogously, since any process tensor possesses a Kraus decomposition [17, 32], it would analogously be possible to extend the definition of incoherent operations (IO) [212], and strictly incoherent operations (SIO) [213, 214] to the non-Markovian case. The former would be MICs that admit at least one set of corresponding Kraus operators, such that none of the individual Kraus operators allows one to extract coherence from the process. The latter would

¹⁰ A first step in this direction, albeit with a very different aim in mind has recently been proposed in [211].

be MICs that allow for a set of corresponding Kraus operators that are all incoherent, where incoherent in this context means that they are diagonal in a classical product basis. Our results would then provide the ideal basis for the operational interpretation of the results of such a resource theory. However, work in these direction will have to be relegated to the future.

Now, having discussed open quantum processes in detail, it is time to drop one of the major assumption this discussion was based on – global temporal order – and see where it leads us.

CAUSALLY INDEFINITE PROCESSES

Throughout this thesis we have frequently hinted at the fact, that many of the derived results do not rely on the existence of a global temporal order. For example, for the generalized extension theorem to hold, it is not a prerequisite that the sets Λ_k that the corresponding joint probability distributions are defined on are times that signify a chronology of events. Analogously, the framework of higher order quantum maps is not erected on the principle of causal ordering as its foundation; rather, temporal order imposes structural requirements on the corresponding quantum combs, and we have seen that causally disordered combs can describe situations in which not all the elements of the underlying quantum network are deterministic.

On the other hand, for the entire discussion of open quantum processes in the previous two chapters, we always had an underlying circuit dilation and clear experimental scenario of successive interrogations of the system of interest in mind. In light of this, allusion to temporally disordered processes seems more like a mathematical afterthought, than a deeper statement in its own right. After all, what would it mean for a process to lack a clear causal order, and how would we describe such processes mathematically? Or, more fundamentally, is it even possible to construct a physical theory without global order that does not run into paradoxes?

Putting temporal order into question seems somewhat unnatural, as it is one of the basic pillars that both our everyday understanding of the world, as well as our physical theories, are built on. Events, no matter how complicated the underlying dynamical theory, seem to happen in a causal succession, and there is a clear arrow of time that defines in which direction they can influence each other. Although according to our intuition, causality seems built into the fabric of reality its position as an axiom can and should be scrutinized. Future experiments might challenge the idea that causal order is fundamental, and may reduce it to a property that exists locally but is violated globally.

For instance, an experiment could consist of two parties (Alice and Bob) conducting measurements in their separated laboratories. The temporal order of events would be heralded by the joint probability distributions of their measurement outcomes. While Alice and Bob experience a well-defined temporal order in their respective laboratories, it is fathomable that a third party (Charlie) that receives measurement data from both Alice and Bob is unable to assign a relative causal order to them based on the received data.

Specifically, if we consider for a moment, that both Alice's and Bob's experiment consists of receiving a quantum system, manipulating it (*i.e.*, implementing an instrument), and sending the resulting system forward (see Fig. 6.1), then Alice and Bob will always be sure about the

causal order of events in their respective laboratories, but the joint probabilities they measure may not allow one to deduce a relative ordering between both laboratories. Importantly, phrasing this scenario in terms of instruments and probability distribution already suggests that all of these scenarios can be described by higher order quantum maps. Then, ‘all’ we have to do to capture potential exotic temporal correlations is to relax the causality constraints we imposed on them, to reflect only local, rather than global temporal ordering.

An experimental example of a causally unordered process is the quantum switch, theoretically introduced in [53] and experimentally realized in [215–217]. Besides the quantum switch, no other exotic causal structure has been implemented experimentally so far. Nonetheless, the mathematical description of such structures is well developed [53, 54, 63, 64] and is subject to active research, both from a foundational perspective (see, e.g., Refs. [42, 158, 218–223]), as well as in terms of possible applications (see, e.g., Refs. [224–227]). As already alluded to, the main mathematical object to represent such general processes are higher order quantum maps – dubbed *process matrices* in this context – introduced in [54] for two parties and later extended to multiple parties in Ref. [42]. Consequently, we will be able to discuss causally disordered process with only minimal alterations to the formalism we have been using throughout this thesis.

Importantly, in Ref. [54], the authors showed that there exist so-called *causally non-separable* process matrices – i.e., valid process matrices that cannot be written as a probabilistic mixture of causally ordered ones. These causally non-separable process matrices respect local causality, i.e., in the respective laboratories of Alice and Bob in the example above, but go beyond what can be described by quantum mechanics that abides by global causal order. Additionally, causally non-separable process matrices also encapsulate processes that can violate *causal inequalities*, i.e., they do not allow for an underlying causal model. As an aside, this situation for temporal correlations is reminiscent of the analogous case for spatial correlations; quantum states that cannot be written as a probabilistic mixture of product states are called non-separable and can violate Bell inequalities.¹

By definition, no process that is compatible with a global causal order exhibits correlations that are obtained from a causally non-separable process matrix.² On the other hand, in our discussion of higher order quantum maps, we have already seen that processes that do not satisfy causality constraints can be simulated non-deterministically, i.e., by conditioning the collection of data on an additional measurement outcome. For example, Charlie might measure an additional system that he possesses, which has previously interacted with both Alice and Bob. He could choose to only record the data he receives from Alice and Bob when the measurement of his system yields a particular outcome. Even if the global causal ordering of Alice’s and Bob’s laboratory is fixed, the data that Charlie records could lead him to believe that there is no temporal ordering between Alice and Bob. In this way, *any* process matrix – causally

¹ In the same way in which there are entangled states that do not violate a Bell inequality [228], causal non-separability is a necessary, but not sufficient, prerequisite for the violation of a causal inequality.

² This seemingly innocuous statement has recently been put into question by the results of [229].

ordered or not – can be implemented experimentally by a quantum circuit (*i.e.*, a causally ordered process) with an additional measurement [15, 17, 52, 111].³

In this chapter, building on the theory of open quantum system dynamics we developed earlier, we shall consider the task of experimentally implementing processes with indefinite causal order by means of conditioning. We will answer two natural questions that suggest themselves: Given a process matrix, what is its implementation in terms of an open process (with measurement)? What *resources* are necessary to simulate a causally non-separable process? In this sense, our motivation is less foundational – we shall not discuss the *meaning* of the potential absence of global temporal order – but rather operational: Assuming the set of permissible process matrices as given, what resources would one need to recreate any one of them experimentally. Nonetheless, we will see that this investigation provides a clear connection between the quantumness of the required implementation resources and the causal non-separability of the process matrix that is to be simulated.

Specifically, we will obtain a general implementation scheme for arbitrary causally disordered processes which requires a genuinely tripartite entangled initial state. Moreover, we will derive necessary and sufficient conditions for a general circuit with measurement to yield a valid process, and give an explicit example of causally non-separable process matrices that can be simulated with a probability that exceeds 50%, thus outperforming previous simulation protocols. Finally, we shall see that – independent of the implementation scheme – the simulation of causally non-separable process matrices requires both genuine tripartite entanglement in the initial state, as well as nonlocal unitary dynamics, *i.e.*, it requires the underlying causal process to be non-Markovian and genuinely quantum. Importantly, the results we encounter in this chapter, which are based on Ref. [3], provide a constructive way to experimentally simulate arbitrary process matrices and establish a clear connection between entanglement, non-Markovianity and causal non-separability.

Consequently, this chapter can be read in two different ways. On the one hand, as an exercise in applying the framework used to describe open quantum processes to other fields, thus demonstrating its versatility. On the other hand – more interestingly – it can be seen as connecting the areas of causally ordered and causally disordered processes via the resources that are needed to simulate the latter by means of the former. Before we proceed with a brief recapitulation of the structural properties of causally ordered processes, a short disclaimer is in order. Firstly, our discussion of causally disordered processes will be distinctly device dependent: that is, we will assume that we know the dimensions of the involved systems, and trust the devices that Alice and Bob use to make their measurements. This assumption allows us to use the process matrix framework as a descriptive tool (for device-independent discussions of causal disorder, see, for example, Refs. [42, 234]). Secondly, we will mostly consider the two-party case, as the definition of causally ordered processes becomes cumbersome in the multipartite case [42, 235], and definitive results cannot be easily obtained.⁴ We shall mention explicitly, whenever results apply to the multipartite case as well. With these disclaimers out

³ In a slightly different context, schemes involving conditioning of data are also actively investigated both theoretically as well as experimentally with respect to the simulation of closed timelike curves (see, *e.g.*, Refs. [230–233]).

⁴ This situation is reminiscent of the complexity of multipartite entanglement as compared to the bipartite case.

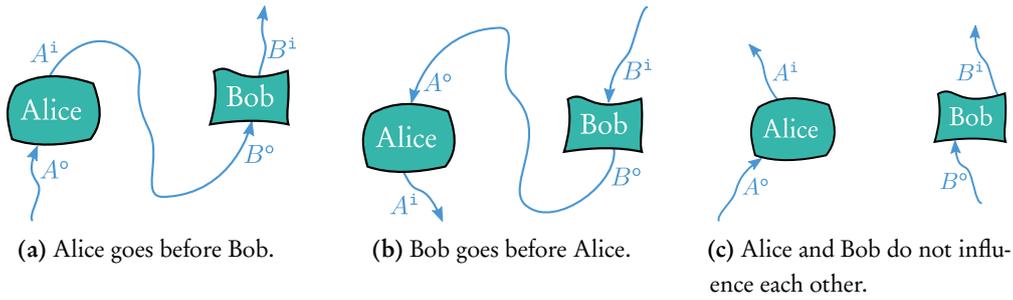


Figure 6.1: Possible global temporal orders for Alice’s and Bob’s respective experiments. The blue arrows depict a possible physical system entering and leaving the respective laboratories. (a) If Alice goes before Bob, her experiment can influence the statistics of Bob’s measurements. (b) If Bob goes before Alice, his experiment can influence the statistics of Alice’s measurements. (c) If Alice’s and Bob’s experiments are independent of each other, neither experimenter can influence the outcome statistics of the other. For explanation of the notation used in the text, inputs and outputs to the laboratories are labeled by the respective Hilbert spaces they are defined on.

of the way, we can now discuss causally disordered processes from the perspective of open quantum system dynamics.

6.1 CAUSALLY ORDERED PROCESSES – RECAP

In order to render the structure of causally disordered processes more transparent, connect them to the theory of open quantum processes, and provide a meaningful definition of causal indefiniteness, we shall briefly reiterate the structural properties of causally ordered processes that we derived in Ch. 2. In this way, the explicit requirements that are dropped, and the ensuing relaxed structural properties of process matrices become perspicuous. To keep notational overhead minimal, we will phrase what follows for two parties only. The generalization to more parties is always straightforwardly possible.

As already alluded to in the introduction to this chapter, we shall consider two laboratories A (Alice’s laboratory) and B (Bob’s laboratory), in which both parties can implement instruments $\mathcal{J}_A = \{M_x^{A^i A^o} \in \mathcal{B}(\mathcal{H}_A^i \otimes \mathcal{H}_A^o)\}$ and $\mathcal{J}_B = \{M_y^{B^i B^o} \in \mathcal{B}(\mathcal{H}_B^i \otimes \mathcal{H}_B^o)\}$.

In correspondence with the notation we employed throughout this thesis, X^o (X^i) labels the Hilbert space of the system that enters (leaves) laboratory X (see Fig. 6.1). Additionally, for better orientation, we shall frequently garnish the maps we consider with additional superscripts to make clear what Hilbert spaces they live on, and $X^{Y^i Y^o}$ implies that $X \in \mathcal{B}(\mathcal{H}_Y^i \otimes \mathcal{H}_Y^o)$. For succinct notation, whenever there is no risk of confusion, we will set $\mathbb{A} = A^i A^o$ and $\mathbb{B} = B^i B^o$ when applicable, *i.e.*, $M_x^{\mathbb{A}} \in \mathcal{B}(\mathcal{H}_A^i \otimes \mathcal{H}_A^o)$ and $M_y^{\mathbb{B}} \in \mathcal{B}(\mathcal{H}_B^i \otimes \mathcal{H}_B^o)$, and we shall drop the superscripts whenever they are not needed for further clarification. Notably, considering laboratories as the backdrop of our considerations has the advantage of introducing labels for events without the allusion to time or temporal order.

Now, as we have already discussed for the case of causally ordered processes, Alice and Bob can perform generalized measurements in their respective laboratories and collect the joint probability distributions $\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B)$ to measure outcomes x (y) in Alice’s (Bob’s)

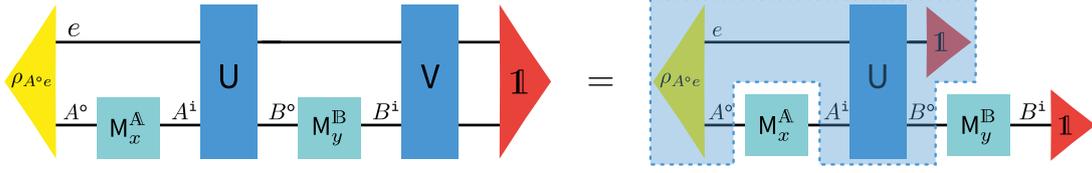


Figure 6.2: Causally ordered process $A \prec B$ without an open output wire. Every causal process can be represented as an open quantum circuit with a trace over the final output line. Due to this final trace, the corresponding process tensor decomposes into $W_{A \prec B}^{B^i B^0 A^i A^0} = \mathbb{1}_{B^i} \otimes W_{A \prec B}^{B^0 A^i A^0}$, where the superchannel $W_{A \prec B}^{B^0 A^i A^0}$ (blue box with dotted borders) is causally ordered.

laboratory, given that the instrument \mathcal{J}_A (\mathcal{J}_B) was used to interrogate the respective system of interest. This joint probability can be computed via

$$\begin{aligned} \mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) &= \text{tr} \left[\left(M_x^{A^i A^0 \top} \otimes M_y^{B^i B^0 \top} \right) W^{B^i B^0 A^i A^0} \right] \\ &= M_x^{A^i A^0} \star M_y^{B^i B^0} \star W^{B^i B^0 A^i A^0}, \end{aligned} \quad (6.1)$$

where we have introduced the *process matrix* $W^{B^i B^0 A^i A^0} \in \mathcal{B}(\mathcal{H}_{B^i} \otimes \mathcal{H}_{B^0} \otimes \mathcal{H}_{A^i} \otimes \mathcal{H}_{A^0})$ [54].

Formally, Eq. (6.1) looks equivalent to the corresponding equations for temporally ordered processes, and we have, for example, already encountered it in Eq. (3.16) when discussing the GET. As before, Eq. (6.1) corresponds to a generalized Born rule [87]. However, here, we will not assume a fixed temporal ordering between events happening in Alice’s and Bob’s laboratories, respectively.⁵ To emphasize this potential lack of global causal ordering, in this chapter, we will denote processes by W instead of C or T .

As we have seen, global causal ordering would impose structural properties on $W^{B^i B^0 A^i A^0}$. For example, a general process $W_{A \prec B}^{B^i B^0 A^i A^0}$ where Alice’s experiment comes before Bob’s – *i.e.*, Alice’s output system becomes Bob’s input system (see Fig. 6.2 for reference) – would be described by a process matrix of the form

$$W_{A \prec B}^{B^i B^0 A^i A^0} = \mathbb{1}_{B^i} \otimes W_{A \prec B}^{B^0 A^i A^0}, \quad (6.2)$$

where $W_{A \prec B}^{B^0 A^i A^0}$ is a causally ordered superchannel, *i.e.*, satisfies

$$\text{tr}_{B^0} (W_{A \prec B}^{B^0 A^i A^0}) = \mathbb{1}_{A^i} \otimes \rho_{A^0}, \quad (6.3)$$

with ρ_{A^0} the initial system state in Alice’s laboratory [17, 31]. This requirement guarantees that, overall, Bob has no influence on Alice’s statistics, *i.e.*, Alice’s statistics are independent of the CPTP map that Bob implements in his laboratory (see Ch. 2). Analogously, if $B \prec A$, we obtain the same requirements, but with the labels of A and B interchanged.

Finally, if the experiments of Alice and Bob are completely independent of each other (denoted by $A || B$), then

$$W_{A || B}^{B^i B^0 A^i A^0} = \mathbb{1}_{A^i B^i} \otimes \rho_{A^0 B^0}, \quad (6.4)$$

where $\rho_{A^0 B^0}$ is the initial (potentially correlated) state of Alice and Bob (see Fig. 6.3). Neither of

⁵ In fact, in Ch. 3, when we discussed the generalized extension theorem, we did not make this assumption either. In the current chapter, though, we will finally make the structural consequences of a lack of global temporal order manifest.

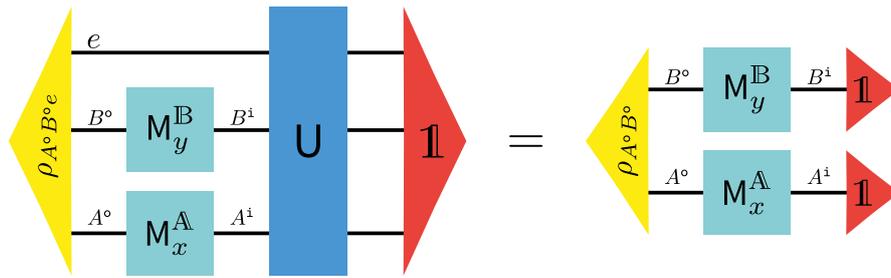


Figure 6.3: Process where Alice and Bob are independent. Events in neither laboratory can influence the outcome statistics in the other laboratory, respectively. However, Alice and Bob can share a correlated initial state $\rho_{A^o B^o}$.

the causally ordered processes $W_{A \prec B}$, $W_{B \prec A}$, and $W_{A \parallel B}$ with the structural properties defined above display any causal anomalies. Additionally, one can imagine an experimental situation, where the causal structure is not known with certainty, or depends on an exterior statistical parameter (like, for example, the outcome of a coin flip). Such a scenario would be described by a convex combination of causally ordered process matrices, which implies the following definition due to Oreshkov *et al.* [54]:

DEFINITION 6.1 (CAUSALLY SEPARABLE PROCESSES): A process $W^{B^i B^o A^i A^o}$ that can be written as

$$W^{B^i B^o A^i A^o} = q W_{A \prec B}^{B^i B^o A^i A^o} + (1 - q) W_{B \prec A}^{B^i B^o A^i A^o} \quad (6.5)$$

$$= q \left(\mathbb{1}_{B^i} \otimes W^{B^o A^i A^o} \right) + (1 - q) \left(\mathbb{1}_{A^i} \otimes W^{A^o B^i B^o} \right), \quad q \in [0, 1] \quad (6.6)$$

is called *causally separable*.

Evidently, the nomenclature stems from the similar definition of separable quantum states. There is still some level of ambiguity in Def. 6.1, as a process matrix $W_{A \parallel B}$ can be either considered to belong to the term $W_{A \prec B}$, or the term $W_{B \prec A}$ in Eq. (6.5). Importantly, though, the property of causal separability of a process is independent of this choice.

The set of causally separable process matrices comprises all processes that one would deem permissible if global causal order was a fundamental trait of nature. However, it has been shown – and we will reiterate the corresponding arguments shortly – that there are process matrices outside the set of causally separable process matrices, that nonetheless do not lead to logical paradoxes [53, 54]. These kinds of process matrices and their simulation shall be the topic of the remainder of this chapter, and we will – interchangeably – call them causally non-separable, causally indefinite, causally disordered, or simply acausal processes.

6.2 LOCAL CAUSAL ORDER

There is no fundamental principle that demands the global causal ordering of events (say, for the sake of concreteness, between the laboratories A and B). However, it seems reasonable not to drop the idea of causality altogether. Evidently, an experimenter that performs a measurement has a clear notion of their local causal ordering of events. To see this, consider a laboratory where one can open a hatch h_1 to let a system enter, then performs a measurement on said

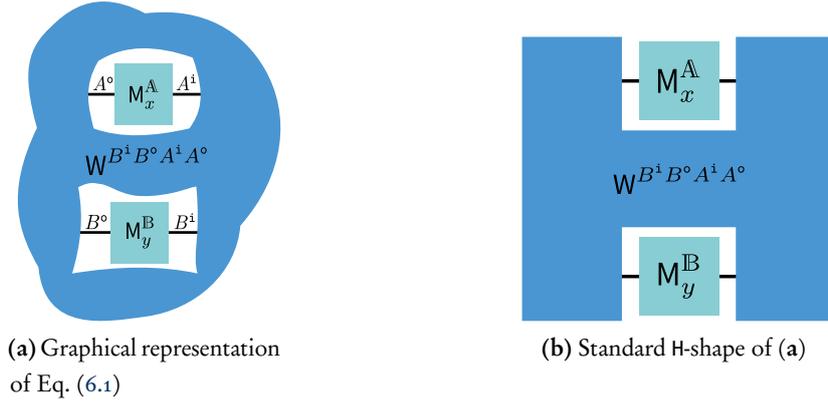


Figure 6.4: Sketch of causal disorder. Alice’s and Bob’s laboratory are embedded in some unknown spacetime structure. Locally, they experience causal order, but globally, there does not have to be a clear ordering between events. To emphasize the potential absence of global causal order, the corresponding process matrices are drawn in a H-shape that does not suggest any *a priori* ordering between events.

system, and afterwards lets it leave the laboratory through another hatch h_2 . Independent of the overall spacetime that the laboratory is embedded in, which could dictate a clear causal ordering between events in different laboratories or not, our experimenter can distinguish between inputs and outputs in their respective laboratory. Put differently, they have certainty about the fact that the system *first* entered through h_1 and *afterwards* left through h_2 , and not the other way round. Additionally, this statement has to hold independently of what other experimenters in different laboratories do. Consequently, any meaningful theory should preserve this local notion of causality, and not, for example, allow for influence from the output at h_2 on the input at h_1 . On the other hand, there is no *a priori* reason for a fixed temporal ordering between events in different laboratories.⁶

We shall formalize the above statements for the two party case. The ‘spacetime’ that Alice and Bob are embedded in is described by a process matrix $W^{B^i B^o A^i A^o}$ that contains all temporal correlations between events in both laboratories (see Fig. 6.4). Now, demanding that local causality holds in Alice’s laboratory independently of what Bob does, is equivalent to demanding that $W^{B^i B^o A^i A^o} \star M^B$ yields a causally ordered process on Alice’s side for all CPTP maps M^B on Bob’s side, *i.e.*:

$$W^{B^i B^o A^i A^o} \star M^B = W_B^{A^i A^o} = \mathbb{1}_{A^i} \otimes \rho_{A^o|B}, \quad \forall \text{ CPTP maps } M^B, \quad (6.7)$$

where the subscript $|B$ implies that the state ρ_{A^o} can depend on the CPTP map that Bob implemented. For a graphical representation, see Fig. 6.5. Analogously, Bob has to experience local causality, independent of what Alice does in her laboratory, which implies

$$W^{B^i B^o A^i A^o} \star M^A = W_B^{A^i A^o} = \mathbb{1}_{A^i} \otimes \rho_{A^o|B} \quad \forall \text{ CPTP maps } M^A. \quad (6.8)$$

Combining these two conditions into one, we thus obtain the definition of valid process matrices for the two party case due to Oreshkov *et al.* [54] (generalization to the many party case is straightforward [42]):

⁶ We are somewhat lax in our usage of laboratory. In general, it is sufficient to think of it as a ‘small’ region of spacetime. For a more thorough discussion, see, for example, Ref. [229].

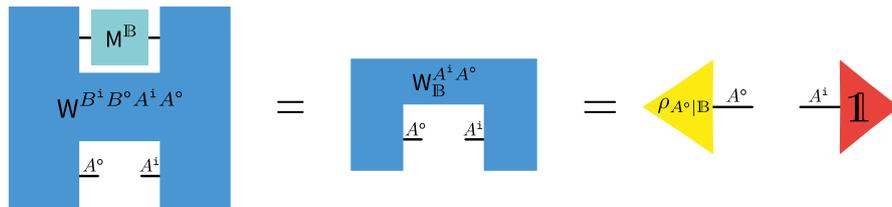


Figure 6.5: Local causality of process matrices. Independent of the CPTP map M^B that Bob implements, the process on Alice’s side is causally ordered. The subscript $|B$ signifies that, the resulting causally ordered process can depend on the CPTP map M^B . We shall depict process matrices in the form of a ‘H’ to emphasize that they do not have to possess a global causal ordering.

DEFINITION 6.2 (VALID PROCESS MATRICES): A matrix $W^{B^i B^o A^i A^o} \in \mathcal{B}(\mathcal{H}_B^i \otimes \mathcal{H}_B^o \otimes \mathcal{H}_A^i \otimes \mathcal{H}_A^o)$ is a valid process matrix if it satisfies

$$W^{B^i B^o A^i A^o} \geq 0, \quad (6.9)$$

and

$$W^{B^i B^o A^i A^o} \star M^A \star M^B = \text{tr} \left[\left(M^{A^T} \otimes M^{B^T} \right) W^{B^i B^o A^i A^o} \right] = 1, \quad (6.10)$$

for all CPTP maps $\{M^A, M^B\}$.

Positivity of $W^{B^i B^o A^i A^o}$ ensures that probabilities are positive⁷, while condition (6.10) enforces local causality in the respective laboratories. Importantly, though, it does not fix a *global* temporal order between Alice and Bob. Inserting the two CPTP maps $M^A = \frac{1}{d_{A^i}} \mathbb{1}_{A^i} \otimes \mathbb{1}_{A^o}$ and $M^B = \frac{1}{d_{B^i}} \mathbb{1}_{B^i} \otimes \mathbb{1}_{B^o}$ into Eq. (6.10), we see that any valid process matrix satisfies

$$\text{tr}(W^{B^i B^o A^i A^o}) = d_{A^i} d_{B^i}. \quad (6.11)$$

Naturally, all causally separable processes satisfy Reqs. (6.9) and (6.10), but there are processes that are valid processes but not causally separable.

It is important to note the structural similarities and differences between the conditions (6.9) and (6.10) for process matrices, and the causality conditions imposed on deterministic quantum combs. Process matrices yield unit probability on the *affine span* of the set of product CPTP maps, *i.e.*, all valid combs that can be written as $\sum_\gamma \kappa_\gamma M_\gamma^B \otimes M_\gamma^A$, where all M_γ^B, M_γ^A are CPTP and $\sum_\gamma \kappa_\gamma = 1$; the set of all deterministic combs that can be decomposed in this way coincides with the set of no-signaling operations [53, 182, 237] between Alice and Bob. Consequently, valid process matrices form the *dual set* to the set of non-signaling processes [183]. On the other hand, every causally ordered (say, $A \prec B$) deterministic comb yields unit probability when applied to another deterministic comb of the same ordering (and such that no open wires remain). In this particular sense, the set of causally ordered processes is self-dual.⁸

⁷ More precisely, positivity of $W^{B^i B^o A^i A^o}$ is *sufficient* for positive probabilities, but not necessary [236]. Demanding $W^{B^i B^o A^i A^o} \geq 0$ can be justified under the additional assumption that Alice and Bob can initially share a maximally entangled state in addition to the spatio-temporal correlations that are given by $W^{B^i B^o A^i A^o}$ [54, 225].

⁸ Strictly speaking, this is not entirely correct, as the original comb and the comb it acts on have to have slightly different structures, so that no open wires remain after contraction. However, in the logic of this comparison, we can neglect this technical difference.

Additionally, since the set of no-signaling maps is strictly smaller than the set of causally ordered combs, the set of admissible process matrices W^{BA} is strictly larger than the set of temporally ordered processes. As we have argued in Ch. 2, every CPTP map M^{A} (and analogously M^{B}) map can be written as a linear combination of generalized Pauli matrices Γ , *i.e.*,

$$M^{\text{A}} = \frac{1}{d_{A^i}} \mathbb{1}_{A^i} \otimes \mathbb{1}_{A^o} + \sum_{\alpha=1}^{d_{A^i}^2-1} a_{\alpha} \Gamma_{\alpha}^{A^i} \otimes \mathbb{1}_{A^o} + \sum_{\beta=1}^{d_{A^i}^2-1} \sum_{\gamma=1}^{d_{A^o}^2-1} c_{\beta\gamma} \Gamma_{\beta}^{A^i} \otimes \Gamma_{\gamma}^{A^o}, \quad (6.12)$$

where, importantly, terms of the form $\mathbb{1}_{A^i} \otimes \Gamma^{A^o}$ are absent. Consequently, Eq. (6.10) defines what terms can and cannot appear in a similar decomposition of a valid process matrix W^{BA} . We list the explicit restrictions that local causality imposes on process matrices in App. D.1. Notably, the set of permissible Pauli terms for causally ordered process matrices is strictly smaller than the corresponding set for general ones, which only guarantee local, but not global, temporal order.

REMARK 6.1: *As already mentioned, we will not be concerned with the ontological meaning of the potential absence of global causal ordering. However, before we embark on the more applied analysis of their simulation, a brief remark on the concept of causally indefinite processes is in order. A priori, there is no fundamental physical concept that guarantees or forces global causal ordering.⁹ Due to satisfaction of Eq. (6.10), causally indefinite process matrices do not lead to paradoxical situations and can therefore not directly be refuted on logical grounds. This makes them a potentially fruitful area of research in their own right.*

More broadly, throughout the history of physics and science in general, the investigation of concepts that are not fundamentally forbidden, but seemingly against common sense, has produced highly celebrated results. Two prime examples that spring to mind are entanglement, and the theory of general relativity. While the former concept follows from the vector space structure of quantum mechanics, but implies highly counterintuitive consequences, the latter starts from the minimal assumption that, locally, one can always transform to a free-falling coordinate system, but this is not necessarily the case globally. Above all, in light of this second example, the reasoning that led to the minimal requirements for valid process matrices seems to point into the right direction.

These arguments notwithstanding, up to this point, no such causal anomalies have been observed in nature – outside of experimental simulations – and it remains to see if the concept of causal indefiniteness can withstand further theoretical scrutiny [158]. Nonetheless, even if causally indefinite processes were to turn out to be incompatible with more fundamental principles, like, for example, purification postulates, their investigation would still advance our understanding of causality and temporal order in general.

With this theoretical prelude out of the way, we can now turn to more concrete questions surrounding process matrices, and connect them to the theory of open quantum system dynamics. This analysis can be carried out independently of the actual existence of causally indefinite processes, and everything we shall discuss is operationally well-defined, and can be realized within the framework of causally ordered processes we encountered in the previous chapters.

⁹ For a somewhat opposing view, and a thorough discussion of role of causality in physics, see, *e.g.*, Ref. [238].

6.3 CONDITIONAL SIMULATION OF CAUSALLY INDEFINITE PROCESSES

Having rigorously introduced the framework for causally disordered processes, we shall now focus on their simulation by means of causally ordered processes, *i.e.*, by means of open quantum system dynamics. By definition, if the process to be simulated does not adhere to the same causal ordering as the circuit invoked for its simulation, then, such a simulation cannot be achieved deterministically. However, as we have already seen, every process matrix can be simulated probabilistically by a quantum circuit with additional postselection [15, 17, 52, 111].

Explicitly, this means that the deducible joint probability distributions $\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B)$ corresponding to a given process matrix W^{BA} can be simulated by conditioning the outcome probabilities of Alice and Bob in a *causally ordered process* on an additional measurement outcome of a measurement performed (say, by Charlie) on an environmental degree of freedom. This means, that for every process matrix W^{AB} , there is a causally ordered process and a conditioning on the environment, such that $\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = p(x, y, \mu_{\text{succ}} | \mathcal{J}_A, \mathcal{J}_B, \mathcal{J}_C) / p(\mu_{\text{succ}} | \mathcal{J}_C)$, where μ_{succ} is the outcome of Charlie's measurement that we condition on.¹⁰ Here, and in what follows, we will distinguish between probabilities \mathbb{P} , obtained from process matrices, and probabilities p , obtained from the causally ordered process we use for the simulation

In this section, we will give a direct, and constructive proof (in the spirit of the one given in [15]) that every process matrix can be simulated by a circuit with postselection, and, importantly, provide a *concrete* implementation scheme. With respect to existing results of this type, our construction is advantageous, as it yields a higher probability of success. Subsequently, we shall analyze how, and under what circumstances, a *valid* process matrix emerges in general from a conditioned circuit. This analysis will then enable an investigation of the necessary resources to simulate a causally inseparable process.

6.3.1 Conditional Simulation of Arbitrary Causally non-Separable Processes

Causally ordered circuits, like the ones depicted in Figs. 6.2 and 6.3, yield a joint probability distribution $p(x, y | \mathcal{J}_A, \mathcal{J}_B)$ to obtain the outcomes x and y in Alice's and Bob's laboratory, respectively. Now, to simulate causally disordered processes, we allow for an additional final measurement with a *fixed* instrument \mathcal{J}_C on the environment, carried out by Charlie (see Fig. 6.6). With this additional measurement on the environment, one obtains a joint probability distribution $p(x, y, \mu | \mathcal{J}_A, \mathcal{J}_B, \mathcal{J}_C)$. In what follows, for compact notation, we will omit the explicit dependence on \mathcal{J}_C to emphasize that we consider Charlie's instrument to be fixed and part of the underlying process. By *conditioning* $p(x, y, \mu | \mathcal{J}_A, \mathcal{J}_B)$ on a measurement outcome on the environment, *i.e.*, by only recording data when the measurement on the environment yields a *specific* outcome, it is possible to simulate any process – causally non-separable or not. We have the following theorem [3]:

¹⁰ Importantly, conditioning on the other measurement outcomes $\mu \neq \mu_{\text{succ}}$ can also lead to valid process matrices (see, *e.g.* Ex. 6.1).

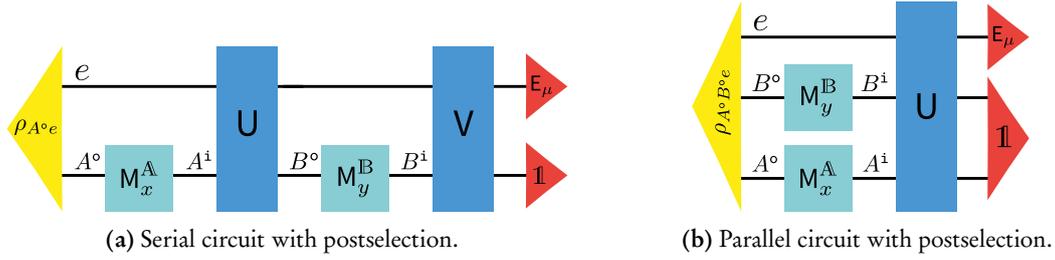


Figure 6.6: Causally ordered circuit with an additional measurement on the environment. After Alice’s and Bob’s experiments have concluded, and the respective output systems and the environment have gone through the corresponding system-environment unitaries, Charlie measures the environment. The respective outcome μ for the fixed instrument \mathcal{J}_C corresponds to the POVM element E_μ . For simulation of causally disordered processes, Alice’s and Bob’s results are only recorded when Charlie obtains the ‘correct’ measurement outcome.

THEOREM 6.1: *Any process matrix W^{BA} can be simulated by a circuit of the form of Fig. 6.7, with an initial state*

$$\rho_{A^o B^o e} := \rho_{s^o e} = \frac{1}{d_{A^o B^o}} \Phi_{A^o}^+ \otimes \Phi_{B^o}^+ \otimes \Pi_r^{(0)}, \quad (6.13)$$

where $\Pi_r^{(0)} = |0\rangle\langle 0|$ is a pure state of an environment r , $\Phi_{X^o}^+ \in \mathcal{B}(\mathcal{H}_{X^o} \otimes \mathcal{H}_{X^o})$ is an unnormalized maximally entangled state, and $d_{A^o B^o} = d_{A^o} d_{B^o}$. After the instruments \mathcal{J}_A and \mathcal{J}_B are applied, the systems and the environment evolve through the unitary U that satisfies

$$\text{tr}_e[\Pi_r^{(0)} U] = \sqrt{\lambda_{\max}}^{-1} \sqrt{W^{\text{BA}^\top}}, \quad (6.14)$$

where λ_{\max} is the largest eigenvalue of W^{BA} . The desired process is simulated by measuring the environment in the computational basis and conditioning on the outcome 0. The probability of success is

$$p_{\text{succ}} = \frac{1}{d_{A^o B^o} \lambda_{\max}}. \quad (6.15)$$

Before we prove the theorem, it is important to emphasize that it is constructive; for any given process matrix W^{BA} , it allows one to find an explicit circuit plus conditioning procedure that will yield the same statistics as the process matrix. This circuit is depicted in Fig. 6.7, where, in the spirit of open quantum system dynamics, we delineate between different spaces: the initial system space s^o includes one half of $\Phi_{A^o}^+$ and $\Phi_{B^o}^+$, i.e., $s^o = A^o B^o$, while the other half, along with r , makes up the environment e .

Additionally, this theorem is very close in spirit to Neumark’s theorem [159–161], that states that every POVM element can be implemented by means of a dilated process. Unsurprisingly, this theorem is a crucial ingredient to the proof of our statement.

Proof. For a given process matrix W^{BA} , probabilities are calculated via

$$\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = \text{tr} \left[\left(M_x^A \otimes M_y^B \right) W^\top \right], \quad (6.16)$$

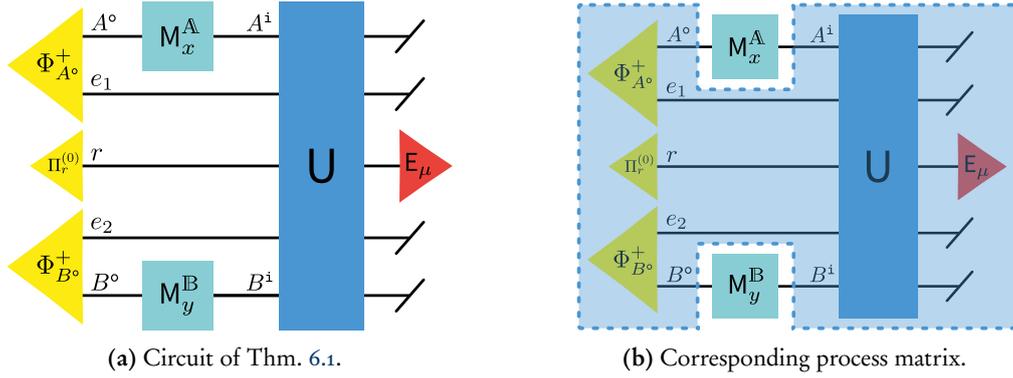


Figure 6.7: Circuit that can simulate any process matrix. For better delineation between what is considered system, *i.e.*, what can be manipulated by Alice and Bob, respectively, and the environment, the inaccessible halves of $\Phi_{A^0}^+$ and $\Phi_{B^0}^+$ are labeled by e_1 and e_2 , such that $e = re_1e_2$, with the understanding that $\mathcal{H}_{e_1} \cong \mathcal{H}_{A^0}$ and $\mathcal{H}_{e_2} \cong \mathcal{H}_{B^0}$. For better orientation, in the right panel the resulting process matrix is shaded in blue, yielding the H form that we use to depict them.

where we have set $W := W^{\text{BA}}$. Now, rolling back the Choi isomorphism, this can be rewritten as

$$\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = \text{tr} \left\{ \left[(\mathcal{M}_x^A \otimes \mathcal{I})[\Phi_{A^0}^+] \otimes (\mathcal{M}_y^B \otimes \mathcal{I})[\Phi_{B^0}^+] \right] W^T \right\} \quad (6.17)$$

which has the form of a quantum circuit with a final measurement. In detail, since W^T is positive we can think of it as a POVM element (up to normalization). By Neumark's theorem, there is a unitary U and a projector $\Pi_r^{(0)} := |0\rangle\langle 0|$ such that $\sqrt{\alpha W^T} = \text{tr}_r[\Pi_r^{(0)} U]$, where $\alpha > 0$ depends on W and is chosen such that $\mathbb{1} - \alpha W^T \geq 0$, *i.e.*, αW^T is a proper POVM element. Putting it all together, we obtain

$$\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = \frac{1}{\alpha} \text{tr} \left\{ \Pi_r^{(0)} U \left[(\mathcal{M}_x^A \otimes \mathcal{M}_y^B) [\Phi_{A^0}^+ \otimes \Phi_{B^0}^+] \otimes \Pi_r^{(0)} \right] U^\dagger \right\} \quad (6.18)$$

The right-hand side of (6.18), up to a normalization factor $d_{A^0 B^0}$, describes a circuit with a measurement on r in the computational basis that yields 0. The probability to measure 0 on the environment is

$$p(0) = \sum_{x,y} p(x, y, 0) = \frac{\alpha}{d_{A^0 B^0}} \sum_{x,y} \mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = \frac{\alpha}{d_{A^0 B^0}}, \quad (6.19)$$

where $p(x, y, 0)$ is the probability to measure x, y and 0 in Alice's, Bob's, and Charlie's laboratory, respectively. We shall see below, that for proper process matrices, as Eq. (6.19) already implies, this probability does not depend on the instruments that Alice and Bob employ. The *maximum* success probability p_{succ} is hence obtained for $\alpha = \lambda_{\text{max}}^{-1}$, where λ_{max} is the maximal eigenvalue of W . \square

Importantly for potential experimental implementations, the result of Thm. 6.1 is robust; due to the linearity of the scheme, a deviation of order ε of the experimentally prepared initial state, or the employed unitary U , from the target one leads to a deviation from the desired

process matrix of the same order. Making the conditioning process more manifest, we can succinctly express Eq. (6.18) as

$$\mathbb{P}(x, y | \mathcal{J}_A, \mathcal{J}_B) = \frac{1}{p(0)} p(x, y, 0 | \mathcal{J}_A, \mathcal{J}_B), \quad (6.20)$$

which explicitly corresponds to a conditional probability. Quite naturally, we can also express these statements in terms of a link product; for every process matrix W , there exists an initial state $\rho_{s^0 e}$ and a unitary U , such that

$$W = \frac{1}{p(0)} \rho_{s^0 e} \star U \star \Pi_r^{(0)} \star \mathbb{1}_{\bar{r}}, \quad (6.21)$$

where $\Pi_r^{(0)} = |0\rangle\langle 0|$ and $\mathbb{1}_{\bar{r}}$ denotes the final trace over all the degrees of freedom except for the ones corresponding to the environment r (see Fig. 6.7a). However, the notation we chose makes the connection to open quantum system dynamics more explicit. We shall rely heavily on the link product in the derivations below. It is easy to see that, unlike most of the following results, Thm. 6.1 does not only hold for the bipartite case, but can readily be extended to the multipartite one, thus providing a very general simulation scheme.

The simplest implementation of Theorem 6.1 is one where the ancilla is a qubit. Defining $X := W^T / \lambda_{\max}$ and $X_{\#} := \mathbb{1} - X$, a possible unitary U which implements the desired process matrix W , as one of two possible process matrices $\{W, W_{\#}\}$, can be written as

$$U = \sqrt{X} \otimes |0\rangle\langle 0| - \sqrt{X_{\#}} \otimes |0\rangle\langle 1| + \sqrt{X_{\#}} \otimes |1\rangle\langle 0| + \sqrt{X} \otimes |1\rangle\langle 1|. \quad (6.22)$$

This choice of U is indeed well-defined and unitary, as $[\sqrt{X}, \sqrt{X_{\#}}] = 0$, and $X_{\#}$ is a positive operator. Conditioning on the outcome 0 yields W , whereas conditioning on 1 yields the process matrix $W_{\#}$. As such, the case where the conditioning happens on a qubit already includes the case of higher dimensional environments; if r does not correspond to a qubit, then one outcome in the computational basis can be labeled as 0, while all the other other ones can be grouped together as outcome 1, thus, effectively, providing a qubit space from the perspective of postselection.

We conclude this section by illustrating Thm. 6.1 for an explicit example.

EXAMPLE 6.1: *In [54], Oreshkov, Costa and Brukner, introduced the following process matrix W_{OCB} as an example for a causally indefinite process:*

$$W_{\text{OCB}}^{B^1 B^0 A^1 A^0} = \frac{1}{4} \left(\mathbb{1}_{B^1 B^0 A^1 A^0} + \frac{1}{\sqrt{2}} \left(\sigma_z^{B^0} \otimes \sigma_z^{A^1} + \sigma_z^{B^1} \otimes \sigma_x^{B^0} \otimes \sigma_z^{A^0} \right) \right), \quad (6.23)$$

where σ_a^X are Pauli matrices on the Hilbert space $\mathcal{H}_X = \mathbb{C}^2$, and we have omitted the respective identity matrices. The experimental situation that W_{OCB} describes is one where the input and output spaces of Alice and Bob are qubits, respectively. It is not obvious at first glance that W_{OCB} is indeed causally indefinite, i.e., cannot be written as a convex mixture of causally ordered processes. This fact was shown in [54] directly, by constructing an information theoretic game that can be won by exploiting W_{OCB} with higher probability than would be possible with any causally ordered process or convex combinations thereof.

For this process matrix, we can choose α in (6.18) to be equal to 2. Consequently:

$$\begin{aligned}\sqrt{X} &= \frac{1}{2} \left(\mathbb{1}_{B^1 B^0 A^1 A^0} + \frac{1}{\sqrt{2}} \left(\sigma_z^{B^1} \otimes \sigma_x^{B^0} \otimes \sigma_z^{A^0} + \sigma_z^{B^0} \otimes \sigma_z^{A^1} \right) \right), \\ \sqrt{X_\#} &= \frac{1}{2} \left(\mathbb{1}_{B^1 B^0 A^1 A^0} - \frac{1}{\sqrt{2}} \left(\sigma_z^{B^1} \otimes \sigma_x^{B^0} \otimes \sigma_z^{A^0} + \sigma_z^{B^0} \otimes \sigma_z^{A^1} \right) \right).\end{aligned}\quad (6.24)$$

The corresponding unitary U for simulation can be constructed using (6.22), and the maximal probability $p(0)$ of success for the outcome 0 on the environment for this choice of unitary is equal to $1/2$. The process matrix $W_\#$ that one would obtain by conditioning on the outcome 1 is given by $W_\# = p(1)X_\# = \frac{1}{2}\mathbb{1} - W$. Both W and $W_\#$ are causally non-separable; conditioning data on either of the outcomes 0 or 1 on the environment, Charlie could not ascribe a causal order to Alice and Bob's actions based on the joint probability distributions he obtains. However, as expected, the average of W and $W_\#$ is causally ordered.

In anticipation of the resource analysis of the simulation procedure, it is instructive to investigate the properties of the employed building blocks for the simulation. By construction, the initial state ρ_{s^0e} exhibits genuine tripartite entanglement, as it is pure and entangled across all possible bipartitions $\{A^0 : B^0e, B^0 : A^0e, e : A^0B^0\}$. On top of that, it is easy to check, that the constructed U is nonlocal, i.e., it cannot be written as $U^{A^1} \otimes Z^{B^1e}, U^{B^1} \otimes Z^{A^1e}$ or $U^{B^1A^0} \otimes Z^e$. For this example, it is even tripartite entangling.¹¹ We will see in Sec. 6.4 that both of these properties – initial tripartite entanglement and a nonlocal unitary – are necessary requirements for the simulation of causally non-separable process matrices.

6.3.2 Conditional Circuits and Valid Processes

In the previous section, we have shown that every process – causally separable or not – can be obtained via a particular circuit with additional conditioning. On the other hand, not every circuit with measurement yields a *valid* process. The following theorem fixes the set of circuits that lead to a valid process matrix when conditioned on an outcome μ on the environment [3]:

THEOREM 6.2: *A circuit with measurement on the environment yielding outcome μ leads to a valid process matrix via conditioning iff the success probability $p(\mu)$ does not depend on the choices of instruments \mathcal{I}_A and \mathcal{I}_B .*

We have already encountered this independence in the derivation of Thm. 6.1. In a slightly different context, this fact has also been discussed in [111], where the authors pointed out that valid process matrices can be simulated by two-time states that have the property that the probability rule becomes linear, i.e., probabilities do not depend on the choice of instruments. Below, we provide a direct proof of this statement. Thm. 6.2 implies that the conditioned simulation of a process matrix is well-defined; since the probability of success does not depend on Alice's and Bob's instruments, the reconstructed process matrix is independent of how Alice and Bob choose to run their respective experiments. Consequently, conditioning is well-defined and clear-cut experimental prescription. However, this also means that the respective circuits are highly fine-tuned; an arbitrary circuit with measurement would almost always lead to

¹¹ In general, nonlocality of a unitary operation is necessary for it to be entangling, but not sufficient.

success probabilities that depend on the choices of instruments, or, put another way, would lead to a reconstructed process that violates local causality.

Proof. We prove this statement for the serial case depicted in Fig. 6.6a, with two system-environment unitary maps U and V . The parallel case, as well as the generalization to the multi-party scenario, follows in a similar vein. The probability to measure x, y and μ in a run of a general two-step circuit with measurement on the environment (with a fixed instrument \mathcal{J}_C) is given by:

$$p(x, y, \mu | \mathcal{J}_A, \mathcal{J}_B) = \text{tr}(\rho_{A^0e} \star U \star M_x^A \star V \star M_y^B \star E_\mu). \quad (6.25)$$

Now, if the probability of success, $p(\mu)$, is independent of the instruments that Alice and Bob use, we can compute it via $p(\mu) = \sum_{x,y} p(x, y, \mu | \mathcal{J}_A, \mathcal{J}_B)$. Notably, due to the lack of Kolmogorov conditions in quantum mechanics, this would generally not be possible. Setting

$$W^{\text{BA}} := \frac{1}{p(\mu)} \rho_{A^0e} \star U \star V \star E_\mu \star \mathbb{1}_{\bar{e}}, \quad (6.26)$$

where $\mathbb{1}_{\bar{e}}$ corresponds to the final trace over all but the environmental degrees of freedom (see Fig. 6.6a), we see that

$$W^{\text{BA}} \star M^A \star M^B = 1 \quad \forall \text{CPTP maps } M^A, M^B. \quad (6.27)$$

As the thusly defined W^{BA} is positive by construction, it is hence a valid process matrix.

To prove the converse statement, let the circuit with measurement be such that, for a choice of (informationally complete) instruments $\tilde{\mathcal{J}}_A, \tilde{\mathcal{J}}_B$ on Alice's and Bob's side respectively, one obtains a valid process matrix W_μ by conditioning on Charlie's outcome. *Per se*, the respective success probability $\tilde{p}(\mu)$ can depend on the employed instruments $\tilde{\mathcal{J}}_A$ and $\tilde{\mathcal{J}}_B$. Using the arguments from above, we see that the reconstructed process matrix is given by

$$W_\mu = \frac{1}{\tilde{p}(\mu)} \rho_{A^0e} \star U \star V \star E_\mu \star \mathbb{1}_{\bar{e}}. \quad (6.28)$$

As, by assumption, W_μ is a valid process matrix, we know that it yields unit probability for any CPTP maps on Alice's and Bob's side, respectively. Thus, for *any* instruments $\mathcal{J}_A, \mathcal{J}_B$, with corresponding CP maps $\{M_x^A\}, \{M_y^B\}$, we have

$$\sum_{x,y} \text{tr}(\rho_{A^0e} \star U \star M_x^A \star V \star M_y^B \star E_\mu) = \tilde{p}(\mu) \sum_{x,y} \text{tr}[(M_x^A \otimes M_y^B) W_\mu^T] = \tilde{p}(\mu), \quad (6.29)$$

which implies that the success probability is independent of the instruments $\mathcal{J}_A, \mathcal{J}_B$. \square

While the process matrix obtained by conditioning on a particular measurement outcome can causally non-separable, the average process matrix is compatible with a definite causal order, as it must be. For example, if, as in Thm. 6.1, the conditioning is carried out for a parallel circuit (as the one depicted in Fig. 6.6b), then the overall process without postselection is of the form $\mathbb{1}_{B^i A^i} \otimes \rho_{B^0 A^0}$, and we have

$$\sum_{\mu} p(\mu) W_\mu = \mathbb{1}_{B^i A^i} \otimes \rho_{B^0 A^0} = W_{A||B}, \quad (6.30)$$

where W_μ is the process matrix corresponding to outcome μ , respectively. This is in agreement with the results from Ex. 6.1, where we had $p(0)W + p(1)W_\# = \frac{1}{4}\mathbb{1}_{BA}$ and $\rho_{B^0A^0} = \frac{1}{4}\mathbb{1}_{A^0B^0}$.

On the other hand, if, like in Fig. 6.6a, the underlying causally ordered circuit is of serial form (say, with order $A \prec B$), then we have

$$\sum_{\mu} p(\mu)W_{\mu} = \mathbb{1}_{B^i} \otimes W^{B^0A^iA^0} = W_{A \prec B}, \quad (6.31)$$

where $W^{B^0A^iA^0}$ is a causally ordered process tensor for Alice's laboratory. The requirements (6.30) and (6.31) impose causality constraints on the success probability of simulation [233], which we shall exploit in the following section to find maximal success probabilities for the simulation of particular process matrices.

6.3.3 Probability of Success

The probability of success for simulating a process matrix depends – amongst others – on its causal structure and the protocol that is employed for its implementation [233]. While we consider *maximal* success probabilities below, here, using the scheme provided in Sec. 6.3.1, we can show the following notable property:

REMARK 6.2: *With the protocol of Theorem 6.1, the success probability for the implementation of a process matrix that violates a causal inequality can exceed 1/2.*

Before we give a proof by example, it is necessary to briefly comment on causal inequalities. As we alluded to in the introduction of this chapter, the discussion of causal structures in terms of process matrices is distinctly instrument dependent, as we assume that we have knowledge of the respective dimension of the involved Hilbert spaces and that quantum mechanics holds in each of the laboratories. On the other hand, we could take a point of view where we do not trust the respective experimenters and do not assume that quantum mechanics necessarily holds locally, which is the instrument independent viewpoint. In this case, it is possible to derive inequalities that must be satisfied if there is a clear causal ordering (or convex mixture thereof) between events on the level of joint probability distributions only [54, 220, 221].

It has been shown, that there are causally non-separable process matrices that do not violate any causal inequalities [225]. Correspondingly, violation of a causal inequality is – in a sense – a stronger signature of acausality than causal non-separability. This is similar to the situation for quantum correlations, where there are entangled states that do not violate Bell inequalities (see, for example, Ref. [239]). In this chapter, except for the above remark, we are not concerned with this distinction, and only differentiate between causally separable and causally non-separable processes, *i.e.*, we will assume a device/instrument dependent point of view. However, as the above remark also holds for the violation of causal inequalities, and not just for causal non-separability, it is worth making the more general statement. We now give an explicit example to illustrate Rem. 6.2.

Proof. It has been shown in Ref. [240] that the process matrix of Ex. 6.1 can be mixed with a certain amount of *white noise* and remain causally non-separable. In detail, the process matrix

$$W' = \frac{\gamma}{4(\gamma+1)}\mathbb{1} + \frac{1}{1+\gamma}W_{\text{OCB}} \quad \text{is causally non-separable for } \gamma \in [0, \sqrt{2}-1], \quad (6.32)$$

where W_{OCB} is the process matrix defined in (6.23). In order to be able to implement W' with success probability $p = \alpha(d_{A_1}d_{A_2})^{-1}$ according to the procedure provided in Sec. 6.3.1, the relation $\mathbb{1} - \alpha W' \geq 0$ has to hold. The minimal eigenvalue of $\mathbb{1} - \alpha W'$ is equal to $4 + 4\gamma - 2\alpha - \alpha\gamma$. For $\gamma \rightarrow \sqrt{2}-1$, the maximal allowed α tends to $4\sqrt{2}/(1+\sqrt{2})$. Consequently, there are causally non-separable process matrices that can be implemented with a probability arbitrarily close to $p = \sqrt{2}/(1+\sqrt{2}) \approx 0.59$. As W' also violates a causal inequality for $\gamma < \sqrt{2}-1$ [240], this means that there are process matrices that violate causal inequalities that can be implemented with a success probability of more than 50%. \square

It is important to contrast this result with the scheme for the simulation of process matrices proposed in [52]; there, independent of the process matrix that is to be simulated, the proposed scheme always yields a success probability of 1/16.

Finally, before discussing the resources necessary for the simulation of causally non-separable process matrices, we briefly have a look at maximal success probabilities for the simulation of a given process matrix W .

As pointed out above, the process matrices W_μ obtained by conditioning on an outcome μ have to add up to a causally ordered process. For example, for the parallel case, we have $\sum_\mu p(\mu)W_\mu = \mathbb{1}_{A^i B^i} \otimes \rho_{A^o B^o}$, which, as all $W_\mu \geq 0$, implies

$$pW \leq \mathbb{1}_{A^i B^i} \otimes \rho_{A^o B^o}, \quad (6.33)$$

where W is the process matrix that is to be simulated and p is the corresponding success probability. Consequently, when searching for the maximal probability of simulation for a given W by means of a parallel circuit, we aim to

$$\begin{aligned} &\textbf{maximize:} && p \\ &\textbf{subject to:} && pW \leq \mathbb{1}_{A^i B^i} \otimes \rho_{A^o B^o} \\ &&& \text{tr } \rho_{A^o B^o} = 1, \quad \rho_{A^o B^o} \geq 0. \end{aligned}$$

This constitutes a semidefinite program (SDP) [241, 242] (albeit not in its standard form), that can be readily solved numerically, by using standard tools (for example CVX, a package for specifying and solving convex programs [243, 244]). For W_{OCB} , this evaluation yields – up to numerical precision – a maximal simulation probability of $p_{\text{max}} = 1/2$, which implies, that, in this particular case, the simulation procedure of Thm. 6.1 is already optimal.

Analogously, for the case of an underlying simulation circuit with causal ordering $A \prec B$, we obtain the following maximization problem:

$$\begin{aligned} &\textbf{maximize:} && p \\ &\textbf{subject to:} && pW \leq \mathbb{1}_{B^i} \otimes W^{B^o A^i A^o} \\ &&& \text{tr}_{B^o} W^{B^o A^i A^o} = \mathbb{1}_{A^i} \otimes \rho_{A^o}, \quad \text{tr } \rho_{A^o} = 1, \quad \rho_{A^o} \geq 0. \end{aligned}$$

This, again, constitutes an SDP that can readily be solved and yields a maximum success probability of $p_{\max} = \sqrt{2}(\sqrt{2} - 1) \approx 0.59$. The same result holds true for simulation via a circuit with causal order $B \prec A$. The serial case slightly outperforms the parallel one in the simulation of W_{OCB} , but the solution of the SDP has the drawback that it does not provide us with an explicit circuit for the ideal implementation. On the other hand, it offers us with the means to gauge the optimality of the simulation circuit provided by Thm. 6.1 for any given W .

It is now time to discuss the extent to which entanglement, nonlocal operations, and non-Markovian features are needed to simulate process matrices that are causally non-separable.

6.4 RESOURCES FOR CAUSALLY NON-SEPARABLE PROCESS MATRICES

The constructive procedure presented in Sec. 6.3.1 to simulate any process matrix W via conditioning requires both genuine tripartite entanglement, as well as a nonlocal unitary U . In this section, we shall see that the simulation of a causally non-separable process matrix via conditioning *always* requires both a genuinely entangled initial state, as well as a nonlocal unitary – no matter the strategy. As a first step, we will prove this statement for the special case of the parallel circuit depicted in Fig. 6.6b. This circuit is, in turn, a special case of a serial circuit, including two unitary evolutions (depicted in Fig. 6.6a), and it is natural to ask if the requirement of initial entanglement can be lifted if two intermediary evolutions are available. We show in Sec. 6.4.2 that this is not the case, and both initial entanglement and nonlocal unitaries are indeed crucial for the simulation of a causally non-separable process matrix.

6.4.1 The Parallel Case

The parallel case is described by a triple $\{\rho_{A^0 B^0 e} = \rho_{s^0 e}, U, E_\mu\}$, *i.e.*, an initial total state, an intermediary unitary dynamics and a conditioning on the environment. By Neumark's theorem, without loss of generality, we can choose E_μ to correspond to an orthogonal projection. Possible resources for the simulation of a causally non-separable process matrix are the initial state $\rho_{s^0 e}$, as well as the unitary U . We have the following theorem [3]:

THEOREM 6.3: *For the conditional simulation of a causally non-separable process matrix W by means of a parallel circuit, it is necessary that both the initial state $\rho_{s^0 e}$ is genuinely tripartite entangled as well as the unitary matrix U is nonlocal, *i.e.*, it cannot be written as a product operation in any possible bipartition.*

Proof. For the proof, we will make use two main ingredients. Firstly, we will employ the formal resemblance of entanglement and causal non-separability, and show that separability of the initial state, or locality of the unitary U directly translates to causal separability. Secondly, we will make use of the fact that certain terms cannot appear in the generalized Pauli decomposition of a valid process matrices. For example, in the decomposition of W , a term of the form $\Gamma^{B^1} \otimes \mathbb{1}_{B^0} \otimes \Gamma^{A^1} \otimes \mathbb{1}_{A^0}$ is not allowed, as it violates local causality. We shall dub such a term an $B^1 A^1$ term. Analogously, a term of the form $\Gamma^{B^1} \otimes \Gamma^{B^0} \otimes \mathbb{1}_{A^1} \otimes \mathbb{1}_{A^0}$, would be denoted an

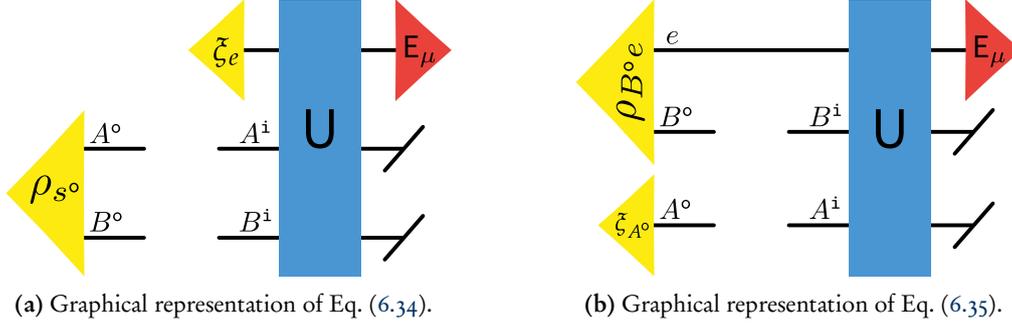


Figure 6.8: Structural consequences of an initial product state in a parallel circuit. Due to the product structure, the resulting process matrix decomposes into a tensor product, too, as can be seen by the presence of open wires that are disconnected from the rest of the circuit (the wires labeled by A^o and B^o in (a), and the wire labeled by A^o in (b).)

$B^i B^o$ term, etc. A full list of generalized Pauli terms that cannot exist in the decomposition of a valid process matrix is given in App. D.1.

Now, to prove the first part of the theorem, let U be an arbitrary unitary matrix, E_μ an arbitrary orthogonal projection on the environment, and let the initial system-environment state be of the product form $\rho_{s^o e} = \rho_{s^o} \otimes \zeta_e$. Following the arguments of the proof of Thm. 6.2, we see that for this scenario, we have

$$\begin{aligned} W_\mu &= \frac{1}{p(\mu)} (\rho_{s^o} \otimes \zeta_e) \star U \star E_\mu \star \mathbb{1}_{\bar{e}} \\ &= \frac{1}{p(\mu)} \rho_{s^o} \otimes (\zeta_e \star U \star E_\mu \star \mathbb{1}_{\bar{e}}) := \rho_{s^o} \otimes \Theta_{s^i}, \end{aligned} \quad (6.34)$$

where we have ordered the link product according to the elements that are contracted with one another (see Fig. 6.8a for a graphical representation) and we have set $s^i = A^i B^i$.

Local causality forbids terms of the form $A^i B^i$, B^i or A^i to appear in the process matrix W_μ (see App. D.1). If Θ_{s^i} is not proportional to $\mathbb{1}_{s^i}$, one of these terms is bound to appear in W_μ . Consequently, if W_μ is a proper process matrix, then it is of the form $\mathbb{1}_{s^i} \otimes \rho_{s^o}$, which is causally ordered (non-signaling).

A similar argument holds for the case $\rho_{s^o e} = \rho_{A^o} \otimes \zeta_{B^o e}$. Here, we have

$$\begin{aligned} W_\mu &= \frac{1}{p(\mu)} (\rho_{A^o} \otimes \zeta_{B^o e}) \star U \star E_\mu \star \mathbb{1}_{\bar{e}} \\ &= \frac{1}{p(\mu)} \rho_{A^o} \otimes (\zeta_{B^o e} \star U \star E_\mu \star \mathbb{1}_{\bar{e}}) := \rho_{A^o} \otimes \Theta_{B^i B^o A^i}. \end{aligned} \quad (6.35)$$

See Fig. 6.8b for a graphical representation. As before, local causality forbids certain Pauli terms; in particular the ones of the form B^i , $A^i B^i$, $B^i B^o$ and $A^i B^i B^o$. This forces $\Theta_{B^i B^o A^i}$ to be decomposable as $\Theta_{B^i B^o A^i} = \mathbb{1}_{B^i} \otimes \zeta_{B^o A^i}$ (where $\zeta_{B^o A^i} \in \mathcal{B}(\mathcal{H}_{B^o} \otimes \mathcal{H}_{A^i})$), which implies that the process matrix (6.35) is causally ordered (Alice goes before Bob). The same argument applies for an initial state of the form $\rho_{s^o e} = \rho_{B^o} \otimes \zeta_{A^o e}$. Consequently, any initial state $\rho_{s^o e}$ of the form

$$\rho_{s^o e} = q \rho_{A^o B^o} \otimes \zeta_e + t \rho_{A^o} \otimes \zeta_{B^o e} + (1 - q - t) \rho_{B^o} \otimes \zeta_{A^o e}, \quad (6.36)$$

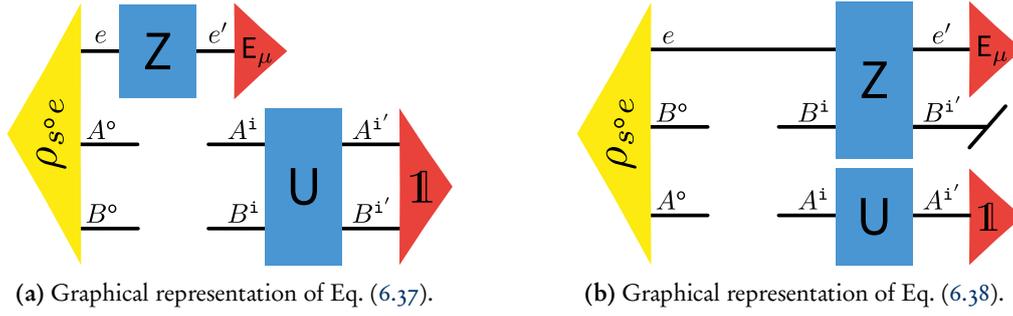


Figure 6.9: Structural consequences of a unitary of product form. Due to the product structure of the respective elements, the resulting process matrix decomposes into a tensor product, too, as can be seen by the presence of open wires that are disconnected from the rest of the circuit (the wires labeled by A^i and B^i in (a), and the wire labeled by A^i in (b).) For compactness, the superscripts of the unitary maps are suppressed in both figures.

does not lead to a causally non-separable process matrix for all coefficients $\{q, t \in [0, 1] \mid q + t \in [0, 1]\}$. As such, for the simulated process matrix to be causally non-separable, it is necessary that the initial state is genuinely tripartite entangled, which provides a direct link between genuine quantum resources and the causal indefiniteness of a process.

To prove the second part of the theorem, *i.e.*, the necessity of a nonlocal unitary, let $\rho_{s^0 e}$ be an arbitrary state and $U = U^{A^i B^i A^i B^i} \otimes Z^{e' e}$ a unitary of product form, where $U^{A^i B^i A^i B^i}$ is the Choi state of a unitary map that acts on Alice's and Bob's outputs, and $Z^{e' e}$ is the Choi state of a unitary map on the environment (see Fig. 6.9a). The resulting process matrix is given by

$$\begin{aligned}
 W_\mu &= \frac{1}{p(\mu)} \rho_{s^0 e} \star (U^{A^i B^i A^i B^i} \otimes Z^{e' e}) \star E_\mu \star \mathbb{1}_{A^i B^i} \\
 &= \frac{1}{p(\mu)} (\rho_{s^0 e} \star Z^{e' e} \star E_\mu) \otimes (U^{A^i B^i A^i B^i} \star \mathbb{1}_{A^i B^i}) \\
 &= \frac{1}{p(\mu)} \mathbb{1}_{A^i B^i} \otimes (\rho_{s^0 e} \star Z^{e' e} \star E_\mu), \tag{6.37}
 \end{aligned}$$

which is causally ordered (specifically, it is non-signaling).

Analogously, if the unitary map is of the form $U = U^{A^i A^i} \otimes Z^{B^i e' B^i e}$, we obtain the following process matrix:

$$\begin{aligned}
 W_\mu &= \frac{1}{p(\mu)} \rho_{s^0 e} \star (U^{A^i A^i} \otimes Z^{B^i e' B^i e}) \star E_\mu \star \mathbb{1}_{A^i B^i} \\
 &= \frac{1}{p(\mu)} \mathbb{1}_{A^i} \otimes (\rho_{s^0 e} \star Z^{B^i e' B^i e} \star E_\mu). \tag{6.38}
 \end{aligned}$$

See Fig. 6.9b for a graphical representation. Again, this process matrix is causally ordered (it allows signaling from Bob to Alice, but not the other way round). A similar argument holds for total unitaries of the form $U = U^{B^i B^i} \otimes Z^{A^i e' A^i e}$. Consequently only non-product – and as such nonlocal – unitaries lead to causally non-separable process matrices. \square

In agreement with the results in Sec. 6.3.1 genuine tripartite entanglement does not mean that Alice and Bob have to initially share entanglement amongst each other. However, the

total state of the environment, Alice, and Bob has to be entangled in any possible bipartition. Genuine tripartite entanglement in the initial state constitutes a quantum memory of the past, that can be used to implement a causally non-separable process. In other words, pre-shared quantum memory is a crucial resource for the simulation of causal non-separability.

Non-product unitaries are signaling [44, 182, 245], which makes the above theorem perspicuous; causal non-separability can only be simulated if a resource is available that enables communication between Alice, Bob and the environment. Such a unitary propagates the initial memory in a detectable way. Consequently, it is the *non-Markovianity* of the underlying circuit that enables the simulation of causally indefinite processes. Having these results for the parallel case at hand, we are now in a position to discuss, if the requirements of initial entanglement and nonlocal unitaries can be relaxed in the serial case, *i.e.*, when two intermediate unitary dynamics are available.

6.4.2 The Serial Case

In the previous sections, we have analyzed the implementation of causally non-separable processes by means of a parallel circuit with additional conditioning. Evidently, if we allow for *any* possible initial state, the parallel circuit is a special case of the *serial* one (depicted in Fig. 6.6a), *i.e.*, a circuit with two intermediary unitaries.

For such a circuit, the resources available for the simulation of a causally indefinite process matrix are the entanglement of the initial state between the system and the environment, and the nonlocality of the two unitary maps U and V . It is then natural to ask, if a serial circuit with measurement allows us to relax the resource requirements for the simulation of causally non-separable process matrices. This question is answered by the following theorem [3]:

THEOREM 6.4: *The conditional simulation of a causally non-separable process matrix by means of a serial circuit requires initial system-environment entanglement and nonlocal intermediate system-environment unitaries.*

Proof. Without loss of generality, we will show this statement for the case of an underlying circuit of causal order $A \prec B$. The prove proceeds in the same manner as the proof of Thm. 6.3: Let $\rho_{A^o e} = \rho_{A^o} \otimes \xi_e$ be the initial system-environment state and let U and V be arbitrary system-environment unitary maps. The resulting process matrix obtained by conditioning on outcome μ on the environment is (see Fig. 6.10a):

$$\begin{aligned} W_\mu &= \frac{1}{p(\mu)} (\rho_{A^o} \otimes \xi_e) \star U \star V \star E_\mu \star \mathbb{1}_{\bar{e}} \\ &= \frac{1}{p(\mu)} \rho_{A^o} \otimes (\xi_e \star U \star V \star E_\mu \star \mathbb{1}_{\bar{e}}), \end{aligned} \quad (6.39)$$

where, as before, $\mathbb{1}_{\bar{e}}$ denotes the trace over all spaces but the final environment (labeled by e'' in Fig. 6.10a). Eq. (6.39) means that for the serial case with initial product state, the resulting process matrix is of the same form as the one in Eq. (6.35) in the proof of Thm. 6.3, and it thus has to be causally ordered ($A \prec B$) for the same reasons.

Analogously, we can prove the necessity of nonlocal unitaries U and V . Firstly, let $V = V^{B^i B^i} \otimes Z^{e'' e'}$. It is straightforward to see, that in this case, the resulting process

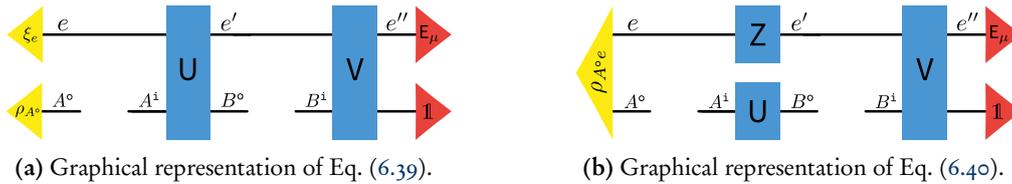


Figure 6.10: Structural consequences of product state and product unitary (serial case). Due to the product structure, the resulting process matrix decomposes into a tensor product, too, as can be seen by the presence of open wires that are disconnected from the rest of the circuit (the wires labeled by A^0 in (a) and the wires A^1 and B^0 in (b)) For compactness, the superscripts of the unitary maps are suppressed in both figures.

matrix W_μ is one-way signaling ($A \prec B$), and consequently causally ordered. Secondly, setting $U = U^{B^0 A^1} \otimes Z^{e' e}$, we obtain

$$W_\mu = \frac{1}{p(\mu)} \rho_{A^0 e} \star Z^{e' e} \star V \star E_\mu \star \mathbb{1}_{\bar{e}} \otimes U^{B^0 A^1} := \Theta^{B^1 A^0} \otimes U^{B^0 A^1}. \quad (6.40)$$

A graphical representation of this equation can be found in Fig. 6.10b.

As $U^{B^0 A^1}$ is the Choi matrix of a unitary map, up to normalization, it corresponds to a maximally entangled state, which implies that terms of the form $B^0 A^1$ appear in its decomposition. This fact forces $\Theta^{B^1 A^0}$ in Eq. (6.40) to satisfy certain structural requirements. As terms of the form $B^1 B^0 A^0$ and $B^1 B^0 A^1 A^0$ cannot appear in the decomposition of W_μ , we must have $\Theta^{B^1 A^0} = \mathbb{1}_{B^1} \otimes \Theta_{A^0}$, which means that the resulting process matrix in Eq. (6.40) is causal ($A \prec B$). \square

As for the parallel case, the nonlocality of the system-environment unitaries is clear. If the *first* unitary U was of product form, local causality in Alice’s laboratory would automatically dictate a global order between the two laboratories. The nonlocality of the *second* system-environment unitary V enables communication between Bob and the environment, which is necessary to ‘blur’ the causal order between Alice and Bob.

Put differently, for a process that does not allow to store information in the environment and access it at a later time, local causality fixes the global temporal order, and our theorems show the importance of genuine pre-existing quantum memory, and system-environment unitaries that transport memory in a detectable way. As such, we have shown what we set out to show in this chapter: independent of the strategy, genuine quantum non-Markovianity is necessary for the simulation of (bipartite) causally non-separable processes.

6.5 CAUSALLY INDEFINITE PROCESSES – SUMMARY

In principle, quantum mechanics is compatible with the existence of processes that lack a definitive predefined causal order. To date, however, no such processes has been found in nature or has been realized experimentally, besides the quantum switch [53, 215, 216]. In this chapter, we set out to analyze the less ambiguous question of providing a constructive procedure to probabilistically simulate every causally disordered process by means of a causally ordered circuit with additional conditioning. With respect to previous results of this type, the strategy

we found yields a higher probability of success, facilitating the experimental simulation of causal anomalies.

The simulation of causally indefinite processes can be obtained by a simple circuit with measurement on the environment. Importantly – in contrast to the results of [52, 111] – the conditioning procedure we encountered happens on the environment, and not on the outputs of Alice and Bob; consequently, Charlie can decide whether or not to record data, without having direct access to Alice’s or Bob’s degrees of freedom. Additionally, beyond the proof of existence, we provided a *constructive* way to obtain a triple of initial state, unitary evolution and measurement outcome on the environment that yields the desired given process matrix W .

At first glance, conditioning might seem like a cherry-picking of data to obtain statistics that display causal anomalies. However, what we have provided is not a mathematical post-processing procedure, performed offline, but an experimental procedure: data is collected whenever the measurement on the environment yields the correct outcome, and we have thus predominantly used the term ‘conditioning’ throughout this chapter, rather than ‘postselection’.

This understanding of the conditioning process makes causality become an emergent average property. For example, for the conditioning process presented in Ex. 6.1, both process matrices W and $W_{\#}$ obtained by conditioning on the two possible outcomes 0 and 1 are causally non-separable, but their average $p(0)W + p(1)W_{\#}$ is – as it should be – causally ordered. We can give further credence to this statement by considering the overall process tensor with an open line on the environment (*i.e.*, before it is contracted with a measurement) that one obtains for this example. It simply is of the form $p(0)W \otimes |0\rangle\langle 0| + p(1)W_{\#} \otimes |1\rangle\langle 1|$, where $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are pure orthogonal states on the environment, and can be considered flags that herald the two different causally disordered processes. Stretching the operational meaning of causality, we can then go a step further and interpret this decomposition in the following way: In every run of the experiment, one of the two disordered processes $\{W, W_{\#}\}$ ‘exists’, and Charlie’s measurement outcome simply reveals which one of the two it was. Looked at under this light, causality really is an average property, that cannot be meaningfully attributed to the individual runs of the experiment. Importantly, this argument relies on the fact that the above decomposition is a convex one, and it would not hold for coherent superpositions of causally disordered processes.

However, interpretations like the above one have to be taken with a grain of salt. Despite the idea of causality (or the lack thereof) in each individual run being the guiding principle of the definition of causally separable processes [42, 235], one should not attach too much meaning to this concept; after all, causality is fundamentally a concept that can only be probed by changing parameters and observing the effect of this change on the outcomes on other events, which cannot be decided in a *single* run of an experiment.

We have seen that the simulation of causally unordered processes is highly fine-tuned. A randomly chosen combination of initial state, unitary evolutions and conditioning almost always leads to a matrix W_{μ} that violates local causality. Put differently, there are spatial correlations that cannot be sensibly understood as temporal correlations [246]. A natural question that suggests itself is that of robustness of the conditioning procedure: Are there circuits that yield causally non-separable process matrices for every conceivable conditioning on the environment? Evidently, Ex. 6.1 is not such a case, as conditioning in the $\{|+\rangle, |-\rangle\}$

basis results in a maximally incoherent process matrix (for either outcome). While numerically, one can construct situation with a wide range of post-selection procedures that result in causally indefinite processes, it is unclear if there are circuits that are fully robust when it comes to the basis that the conditioning occurs in [247].

Finally, we analyzed in detail the resources necessary to probabilistically implement a causally non-separable process matrix. Our results show that the implementation of causally unordered processes requires both genuine tripartite entanglement in the initial state as well as nonlocal unitary dynamics. The requirement of initial entanglement cannot be lifted even if we allow for more nonlocal communication by introducing a second system-environment unitary. Initial entanglement represents a genuine quantum memory of the past, while a nonlocal unitary dynamics allows for a detectable propagation of this quantum memory. Hence, the obtained results – loosely speaking – establish that only genuinely quantum non-Markovian processes allow for the simulation of causally non-separable processes via conditioning. This result, however, only holds for the two-party case; if more parties are involved, causal inequalities can be violated with purely classical processes [248].

While the connection of entanglement, nonlocality and causal inseparability is insightful, it by no means provides a resource theory of causal disorder. The same holds true for the analysis of the impact of causally indefinite spatio-temporal correlations on computational and communication tasks [53, 226, 249]. A fully-fledged resource theory, with a well-defined measure of acausality¹², is still missing. Its future development will shed further light on the operational meaning and potential applications of causal indefiniteness.

¹² In [95], causal robustness is suggested as a measure that satisfies certain desirable properties. However, it arguably lacks a clear operational interpretation.

CONCLUSIONS

Within this thesis, we have looked at general stochastic processes from a multitude of angles, and established higher order quantum maps as their natural descriptor. It has turned out that, fundamentally, many frameworks that apply to apparently different situations are sides of the same coin, and can be accommodated in our generalized extension theorem.

The generalized extension theorem (GET) we have derived not only provides an ‘umbrella theorem’ that establishes an axiomatic underpinning for causal modeling and stochastic processes, but also advocates for the distinctly operational point of view that we have adopted throughout. A process (classical, quantum, or beyond) is fully described once the outputs to all implementable inputs can be predicted, and, as soon as one has a descriptor that satisfies this requirement, no further information can be learnt about it. Following this approach, we have been able to recover generalized Kolmogorov conditions, and provide the minimal requirements for the existence of a quantum stochastic process.

Experimentally, these descriptors are meaningful due to the linearity of quantum mechanics, which allows for their reconstruction in a finite number of experiments. We can use the general ideas of process tomography to make deterministic quantum combs amenable to experimental investigation when only limited experimental resources are available. Additionally, besides merely describing the stochastic process at hand, the structural properties of process tensors enable the characterization and quantification of persistent memory effects. In particular, we have given a full characterization of quantum Markov order, and provided an operational understanding of CP divisibility. Importantly, in all of our discussions, the guiding principle has not been the resulting mathematical properties, but the logical concepts, such as conditional independence in the case of Markov order. Structural properties were – as they should be – the end, rather than the starting point of our considerations.

Naturally, the first question for future research that comes to mind, is that of reconstruction complexity. The number of measurements necessary for the reconstruction of a process tensor scales exponentially with the number of time steps, which seems to put a pin into the feasibility of their actual experimental reconstruction. Consequently, in the future, compressed sensing ideas, as they have been developed for the tomographic reconstruction of sparse or low rank density matrices [188, 250] will have to be generalized to the temporal setting, to efficiently recover process tensors from experimental data.

On the other hand, quite often, it will be sufficient to gauge the non-Markovianity of a process without attempting its full reconstruction. We have encountered oCP divisibility as a witness of non-Markovianity that, in addition, allows one to decide what temporal correlations

can and cannot be present. Evidently, there is a large range of possible, operationally well-defined alternative witnesses to be discovered based on the process tensor framework, that are potentially better suited to the task, and easier to access experimentally. In general, our operational approach will lead to witnesses, that not only reveal the non-Markovianity of a process, but make quantitative assertions about the existing memory effects.

More fundamentally, the GET offers us a versatile toolbox to abstractly model quantum stochastic processes and their continuous limit, without the need to resort to a system-environment Hamiltonian. We have provided the example of a $1 \rightarrow 3$ adapter, that, at each iteration, triples the number of time steps the process tensor acts on. Such a generation of quantum stochastic processes makes the resulting underlying field theory accessible. For example, it could potentially be obtained in a similar way as continuum limits of quantum lattice systems [251].

Apart from this possibility of explicitly analyzing the underlying process, the combination of the GET and the concept of restricted process tensors has also provided us with an operationally clear-cut way to determine the boundary between classical and quantum processes. The corresponding analysis in Ch. 6 demonstrates the intimate relation between discord and classicality in temporal processes, and enables the comprehensive classification of classical stochastic processes. With said classification at hand, further analysis of the structural properties of classical process tensors suggests itself, as does the development of a resource theory of classicality for temporal processes. Such a theory would allow for a clear quantification of the value of non-classical correlations in time, and would thus be of tremendous technological importance.

Finally, in our discussion of causally disordered processes, we have discovered a direct connection between multipartite entanglement, non-locality, and the simulability of causal non-separability. These results open up many interesting novel research directions. As we have seen, the constructive approach we provided is not necessarily ideal, and optimization of the simulation circuit would simplify future experimental implementation. Additionally, while we analyzed the resources that go into the simulation of causal indefiniteness, we have not fully investigated how stable the simulation procedure is. Obtaining a clearer, more quantitatively tangible relation, between the quantumness of the employed resources and the causal non-separability of the simulated process matrix, would shed further light on the interconvertibility of quantum correlations and causal disorder.

Intimately intertwined with this question is that of development of a fully-fledged resource theory of causal disorder. Such a theory would allow one to value the causal non-separability of a given process matrix, and give an operational meaning to the degree of causal indefiniteness displayed by a process. This, in turn, would give more credence to the concept of causally disordered processes *per se*.

Switching to an operational description of temporal processes by means of higher order quantum maps makes them amenable to direct analysis, and answers many open questions. Luckily, though, for theoretical physicists working in the field, it allows one to phrase even more questions than it answers, and sheds light on a multitude of fruitful future research avenues.

Part I

APPENDIX



CAUSAL MODELING AND GENERAL STOCHASTIC PROCESSES

A.1 DEFINITION OF STOCHASTIC PROCESSES

For completeness, here we give the definition of a stochastic process as it can be found in literature (see, e.g., [70, 71]), using the concept of *random variables* X , which are defined as follows [71]:

DEFINITION A.1 (RANDOM VARIABLE): *Let (Ω, F) and (Ω', F') be measurable spaces. A map $X: \Omega \rightarrow \Omega'$ is called a random variable, if $\forall f' \in F'$, we have $X^{-1}(f') \in F$.*

Random variables are useful for several reasons. In general, the σ -algebra F could be some very abstract set, and it is often helpful to introduce a mapping $X: \Omega \rightarrow \Omega'$ to some new, more tangible, σ -algebra F' . For example, the set Ω' is often taken as the real line, i.e., $X: \Omega \rightarrow \mathbb{R}$, which is helpful, as real numbers allow for further computation [71]. In this sense, the mapping X could be something as benign as a mapping $X: \{\text{H}, \text{T}\} \rightarrow \{0, 1\}$ from the outcomes H (heads) and T (tails) to the integers 0 and 1 for a coin flip. On the other hand, X could also specify, what aspects of an experiment are of interest. For example, the original sample space could contain the outcomes $\{1, \dots, 6\}$, but one might only be interested in whether or not the outcome was odd or even. The corresponding random variable could then be of the form $X(1) = X(3) = X(5) = 0$ and $X(2) = X(4) = X(6) = 1$.

In order to be meaningfully defined, the mapping X should respect measurability. If the space Ω' is equipped with a σ -algebra F' , then, for every $f' \in F'$, we must have $X^{-1}(f') \in F$; every pre-image of a measurable set f' has to be an element of the σ -algebra F , and as such measurable.

As mentioned, in practice the image space is quite often the real line \mathbb{R} , and the space \mathbb{R} has a natural σ -algebra B , the σ -algebra corresponding to the set of open subsets of \mathbb{R} [72] (the so-called *Borel σ -algebra*).

If (μ, Ω, F) is a probability space, then X induces a probability measure \mathbb{P} on (Ω', F') via

$$\mathbb{P}(f') = \mu(X^{-1}[f']), \quad (\text{A.1})$$

and as such induces a new probability space $(\mathbb{P}, \Omega', F')$. Using the the notion of a random variable we then can define the concept of a classical stochastic process:

DEFINITION A.2 (CLASSICAL STOCHASTIC PROCESS): *Let (Ω, F) be a measurable space and let Λ be a set. A stochastic process is a mapping*

$$X: \Omega \times \Lambda \rightarrow \mathbb{R}, \quad (\text{A.2})$$

where $X(t) : \Omega \rightarrow \mathbb{R}$ is a random variable for every $t \in \Lambda$. For a fixed $\omega \in \Omega$ the curve $t \mapsto X(\omega, t)$ is called a trajectory of the stochastic process (see Fig. 3.1).

The functional dependency $X(t)$ contains all statistical correlations of the process at hand, like, e.g., its memory length, and the joint probability distribution \mathbb{P}_{Λ_k} that we used in our Def. 3.4 of classical stochastic processes relates to the above definition via

$$\mathbb{P}_{\Lambda_k}(x_k, \dots, x_1) = \mathbb{P}(X(t_k) \in x_k, \dots, X(t_1) \in x_1), \quad (\text{A.3})$$

where $\Lambda_k = \{t_k, \dots, t_1\}$, the events x_k, \dots, x_1 are elements of F' , and $\mathbb{P}(X(t_k) \in x_k, \dots, X(t_1) \in x_1)$ is the probability for the random variable to take values that lie in x_α at t_α . In this sense, directly defining stochastic processes in terms of probability spaces – as we did in the main text – has the advantage of incorporating the mapping $(\Omega, F) \rightarrow (\Omega', F')$ given by the random variable, at the expense of losing the flexibility that random variables provide.

Naturally, this equivalence carries over to the general case of $|\Lambda| = \infty$, and we have

$$\mathbb{P}_\Lambda(\mathbf{x}_\Lambda) = \mathbb{P}(X(t) \in x_t), \quad \text{with } t \in \Lambda. \quad (\text{A.4})$$

For all intents and purposes, both definitions of classical stochastic processes are equivalent. In practice, we are rarely ever concerned with random variables or their trajectories *per se*, but only with actually measurable quantities. For example, coming back to the example of Brownian motion, it does not make sense to ask for the probability to measure a certain trajectory of the particle. Consequently, for our purposes, it proved to be more insightful to define classical stochastic processes without resorting to random variables.

A.2 PROOF OF THE GENERALIZED EXTENSION THEOREM (GET)

Here, we prove the general extension theorem for processes with interventions. The structure of the proof follows the derivation of the KET presented in [72]; given a family of compatible combs, we use the consistency condition to define a unique comb $\mathcal{C}_\Lambda^\sharp$, that contains all finite ones as ‘marginals’, on a large enough ‘container space’. Due to its properties (linearity and boundedness) $\mathcal{C}_\Lambda^\sharp$ can then be uniquely extended to a comb $\overline{\mathcal{C}}_\Lambda^\sharp$ on the closure of said container space.

Let Λ be a (possibly uncountable) set, $\{\Lambda_k\}_{\Lambda_k \subseteq \Lambda}$ the set of all finite subsets of Λ , and let $\{\mathcal{C}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda}$ be the corresponding family of combs. For ease of notation, we assume all CP maps that the combs act on to have the same input and output space, *i.e.*, $\mathcal{M}_{x_{j_k}} : \mathcal{B}(\mathcal{H}_\alpha) \rightarrow \mathcal{B}(\mathcal{H}_\alpha)$; a generalization to maps with distinct input and output spaces is straightforward.¹ We denote the space of these maps by $\mathcal{F}(\mathcal{H}_\alpha)$. Consequently, we have $\mathcal{C}_{\Lambda_k} : \mathcal{F}_{\Lambda_k} \rightarrow \mathbb{R}$, where $\mathcal{F}_{\Lambda_k} = \bigotimes_{\alpha \in \Lambda_k} \mathcal{F}(\mathcal{H}_\alpha)$ and $\bigotimes_{\alpha \in \Lambda_k}$ denotes a tensor product over all times $t_\alpha \in \Lambda_k$. Let the family of combs satisfy the consistency condition $\mathcal{C}_{\Lambda_k}[\bullet] = \mathcal{C}_{\Lambda_K}[\mathcal{I}_{\Lambda_K \setminus \Lambda_k} \bullet]$ for all finite $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$, where $\mathcal{I}_{\Lambda_K \setminus \Lambda_k} = \bigotimes_{\alpha \in \Lambda_K \setminus \Lambda_k} \mathcal{I}_\alpha$.

¹ Here, the restriction to CP maps is not of importance, but merely assumed to establish a clear connection to the quantum case. In principle, the maps $\mathcal{M}_{x_{j_k}}$ could be *any* linear maps between spaces of bounded operators.

Now, we ‘lift’ this family of combs to a comb $\mathcal{C}_\Lambda^\sharp$ that contains all of them as ‘marginals’. To this end, we define the inverse projection $\pi_{\Lambda_k}^{-1} : \mathcal{F}_{\Lambda_k} \rightarrow \mathcal{F}_\Lambda$, with

$$\pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}] = \xi_{\Lambda_k} \bigotimes_{\alpha \in \Lambda \setminus \Lambda_k} \mathcal{I}_\alpha \quad (\text{A.5})$$

for all $\xi_{\Lambda_k} \in \mathcal{F}_{\Lambda_k}$, which trivially extends any ξ_{Λ_k} to a corresponding operator that lies in \mathcal{F}_Λ ; Eq. (A.5) should be read as defining the operator that acts non-trivially on $\mathcal{B}(\bigotimes_{\alpha \in \Lambda_k} \mathcal{H}_\alpha)$, and trivially on all other spaces. The inverse projection operator $\pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}]$ exists and is unique for any finite $\Lambda_k \in \Lambda$ and all $\xi_{\Lambda_k} \in \mathcal{F}_{\Lambda_k}$ [252].

In the same way, we can define a partial inverse projection $\pi_{\Lambda_k \leftarrow \Lambda_K}^{-1} : \mathcal{F}_{\Lambda_k} \rightarrow \mathcal{F}_{\Lambda_K}$ for any two finite sets $\Lambda_k \subseteq \Lambda_K \subseteq \Lambda$, i.e., $\pi_{\Lambda_k \leftarrow \Lambda_K}^{-1}[\xi_{\Lambda_k}] = \xi_{\Lambda_k} \bigotimes_{\alpha \in \Lambda_K \setminus \Lambda_k} \mathcal{I}_\alpha$. In terms of partial inverse projections, the consistency property reads

$$\mathcal{C}_{\Lambda_k}[\xi_{\Lambda_k}] = \mathcal{C}_{\Lambda_K} \left[\pi_{\Lambda_k \leftarrow \Lambda_K}^{-1}[\xi_{\Lambda_k}] \right]. \quad (\text{A.6})$$

Let $\mathcal{F}_\Lambda^\sharp$ denote the set of all ‘lifted’ operators, i.e.,

$$\mathcal{F}_\Lambda^\sharp = \{ \xi \in \mathcal{F}_\Lambda : \xi = \xi_{\Lambda_k} \bigotimes_{\alpha \in \Lambda \setminus \Lambda_k} \mathcal{I}_\alpha, \text{ for some finite } \Lambda_k \}. \quad (\text{A.7})$$

It is straightforward to see that this set forms a vector space; for any $\alpha, \beta \in \mathbb{R}$ and $\xi = \pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}]$, $\zeta = \pi_{\Lambda_K}^{-1}[\zeta_{\Lambda_K}]$, we have

$$\alpha \xi = \pi_{\Lambda_k}^{-1}[\alpha \xi_{\Lambda_k}] \in \mathcal{F}_\Lambda^\sharp \quad \text{and} \quad \alpha \xi + \beta \zeta = \pi_{\Lambda_k \cup \Lambda_K}^{-1}[\alpha \xi_{\Lambda_k} + \beta \zeta_{\Lambda_K}] \in \mathcal{F}_\Lambda^\sharp, \quad (\text{A.8})$$

where $\Gamma_{\Lambda_k \cup \Lambda_K} = \pi_{\Lambda_k \leftarrow \Lambda_k \cup \Lambda_K}^{-1}[\alpha \xi_{\Lambda_k}] + \pi_{\Lambda_K \leftarrow \Lambda_k \cup \Lambda_K}^{-1}[\beta \zeta_{\Lambda_K}] \in \mathcal{F}_{\Lambda_k \cup \Lambda_K}$. Additionally, $\mathcal{L}_\Lambda^\sharp$ becomes a normed vector space by setting $\|\xi\| = \|\pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}]\| := \|\xi_{\Lambda_k}\|_{\text{op}}$, where $\|\cdot\|_{\text{op}}$ is the operator norm in \mathcal{F}_{Λ_k} . On this vector space, we can define the comb $\mathcal{C}_\Lambda^\sharp$ via $\mathcal{C}_\Lambda^\sharp[\xi] := \mathcal{C}_{\Lambda_k}[\xi_{\Lambda_k}]$, where $\xi = \pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}]$.

$\mathcal{C}_\Lambda^\sharp$ is well-defined: if there are two different operators $\xi_{\Lambda_k} \in \mathcal{F}_{\Lambda_k}$ and $\xi_{\Lambda_K} \in \mathcal{F}_{\Lambda_K}$, such that $\xi = \pi_{\Lambda_k}^{-1}[\xi_{\Lambda_k}] = \pi_{\Lambda_K}^{-1}[\xi_{\Lambda_K}]$, the consistency property ensures that $\mathcal{C}_\Lambda^\sharp[\xi]$ is unique; it is straightforward to see that $\pi_{\Lambda_k \leftarrow \Lambda_k \cup \Lambda_K}^{-1}[\xi_{\Lambda_k}] = \pi_{\Lambda_K \leftarrow \Lambda_k \cup \Lambda_K}^{-1}[\xi_{\Lambda_K}] := \xi_{\Lambda_k \cup \Lambda_K}$. Employing the consistency condition yields

$$\begin{aligned} \mathcal{C}_{\Lambda_k}[\xi_{\Lambda_k}] &= \mathcal{C}_{\Lambda_k \cup \Lambda_K}[\pi_{\Lambda_k \leftarrow \Lambda_k \cup \Lambda_K}^{-1}[\xi_{\Lambda_k}]] = \mathcal{C}_{\Lambda_k \cup \Lambda_K}[\xi_{\Lambda_k \cup \Lambda_K}] \\ &= \mathcal{C}_{\Lambda_k \cup \Lambda_K}[\pi_{\Lambda_K \leftarrow \Lambda_k \cup \Lambda_K}^{-1}(\xi_{\Lambda_K})] = \mathcal{C}_{\Lambda_K}[\xi_{\Lambda_K}], \end{aligned} \quad (\text{A.9})$$

and consequently $\mathcal{C}_\Lambda^\sharp[\xi]$ is independent of the representation of ξ .

The operator $\mathcal{C}_\Lambda^\sharp$ is bounded, because every \mathcal{C}_{Λ_k} is bounded. It is also linear; due to the linearity of \mathcal{C}_{Λ_k} and the linearity of the inverse projection operators, we have

$$\mathcal{C}_\Lambda^\sharp[\alpha \xi + \beta \eta] = \alpha \mathcal{C}_\Lambda^\sharp[\xi] + \beta \mathcal{C}_\Lambda^\sharp[\eta] \quad (\text{A.10})$$

for all $\alpha, \beta \in \mathbb{R}$ and $\eta, \xi \in \mathcal{F}_\Lambda^\sharp$. Any linear bounded transformation from a normed vector space X to a normed complete vector space Y can be uniquely extended to a linear transformation from the completion \bar{X} of X to Y [253]. Consequently, there exists a unique map $\bar{\mathcal{C}}_\Lambda^\sharp$ defined

on the completion $\overline{\mathcal{F}}_\Lambda^\sharp$ of $\mathcal{F}_\Lambda^\sharp$ (this completion is sometimes called *quasilocal algebra* in the literature [18, 254]) that has – by construction – the family $\{\mathcal{C}_{\Lambda_k}\}_{\Lambda_k \subseteq \Lambda}$ as ‘marginals’. This concludes the proof. \square

It is important to note that $\overline{\mathcal{F}}_\Lambda^\sharp$ does *not* coincide with \mathcal{F}_Λ (they coincide iff Λ is finite [252]). Consequently, there might be different combs \mathcal{C}_Λ defined on \mathcal{F}_Λ with coinciding action on all elements of $\overline{\mathcal{F}}_\Lambda^\sharp$. This, however, is not problematic. On the one hand, $\overline{\mathcal{F}}_\Lambda^\sharp$ ‘is in a way more important than’ \mathcal{F}_Λ because its elements arise from the ones of $\mathcal{F}(\mathcal{H}_\alpha)$ ‘by extension and algebraical and topological processes’ [252]. On the other hand – just like for the KET [71] – the different possible combs on \mathcal{F}_Λ all lead to the same measurement statistics on any experimentally accessible set of times, so this non-uniqueness is not accessible/detectable in practice.

A.3 PROOF GET \Rightarrow KET

Here, we show that the GET contains the KET as a corollary. In detail, we show that any family of classical joint probability distributions can be mapped onto a family of quantum combs that satisfies the generalized consistency condition. The GET then guarantees that there is an underlying classical comb $\mathcal{C}_\Lambda^{\text{cl}}$, and thus also an underlying classical process \mathbb{P}_Λ .

For the proof, we exploit the structural requirements a classical comb has to satisfy. To make these requirements evident, we employ the Choi-Jamiołkowski isomorphism, denoting the combs by \mathcal{C}_{Λ_k} , and the CP maps they act on by $M_{x_{j_k}} \in \mathcal{B}(\mathcal{H}_\alpha^{\text{i}} \otimes \mathcal{H}_\alpha^{\text{o}})$.² Then, the action of a comb is expressed via

$$\mathcal{C}_{\Lambda_k}[M_{x_{j_k}}, \dots, M_{x_{j_1}}] = \text{tr} \left[\left(M_{x_{j_k}}^{\text{T}} \otimes \dots \otimes M_{x_{j_1}}^{\text{T}} \right) \mathcal{C}_{\Lambda_k} \right], \quad (\text{A.11})$$

where, as usual, \cdot^{T} denotes the transpose in the reference basis. In this notation, letting a comb act on identity maps amounts to projecting it on maximally entangled states, *i.e.*,

$$\mathcal{C}_{\Lambda_k}^{\text{i}\Lambda_k} = \text{tr}_{\Lambda_k \setminus \Lambda_k} \left[\left(\mathbb{1}_{\Lambda_k} \otimes_{\alpha \in \Lambda_k \setminus \Lambda_k} \Phi_\alpha^+ \right) \mathcal{C}_{\Lambda_k} \right], \quad (\text{A.12})$$

where $\text{tr}_{\Lambda_k \setminus \Lambda_k}$ signifies a trace over the Hilbert spaces corresponding to times $t_\alpha \in \Lambda_k \setminus \Lambda_k$ and $\mathbb{1}_{\Lambda_k}$ is the identity matrix on the remaining Hilbert space $\mathcal{H}_{\Lambda_k} := \otimes_{\alpha \in \Lambda_k} (\mathcal{H}_\alpha^{\text{i}} \otimes \mathcal{H}_\alpha^{\text{o}})$.

A classical family of joint probability distributions \mathbb{P}_{Λ_k} can be represented by a family of classical combs

$$\mathcal{C}_{\Lambda_k}^{\text{cl}} = \sum_{j_k, \dots, j_1} \mathbb{P}_{\Lambda_k}(x_{j_k}, t_k; \dots; x_{j_1}, t_1) \mathbb{1}_k^{\text{i}} \otimes |x_{j_k}\rangle \langle x_{j_k}| \otimes \dots \otimes \mathbb{1}_1^{\text{i}} \otimes |x_{j_1}\rangle \langle x_{j_1}|, \quad (\text{A.13})$$

where $\mathbb{1}_\alpha^{\text{i}} \in \mathcal{B}(\mathcal{H}_\alpha^{\text{i}})$ are identity matrices and $|x_{j_\alpha}\rangle \langle x_{j_\alpha}| \in \mathcal{B}(\mathcal{H}_\alpha^{\text{o}})$ are orthogonal pure states corresponding to the measurement outcomes x_{j_α} . The classical combs defined by (A.13) are diagonal in the reference basis and correctly reproduce the probabilities given by \mathbb{P}_{Λ_k} ; indeed, it is easy to see that we have

$$\text{tr} \left[\left(\mathbb{P}_{x_{j_k}}^{\text{T}} \otimes \dots \otimes \mathbb{P}_{x_{j_1}}^{\text{T}} \right) \mathcal{C}_{\Lambda_k}^{\text{cl}} \right] = \mathbb{P}_{\Lambda_k}(x_{j_k}, t_k; \dots; x_{j_1}, t_1), \quad (\text{A.14})$$

² As classical processes form a subset of quantum processes, here, we do not lose generality by explicitly employing the CJL.

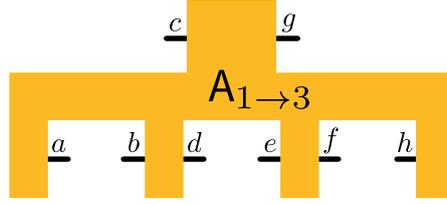


Figure A.1: Graphical representation of a $1 \rightarrow 3$ adapter $A_{1 \rightarrow 3}$. This figure is a copy of Fig. 3.9 in the main text, and shown here for ease of reference.

where $P_{x_{j_\alpha}} = |x_{j_\alpha}\rangle\langle x_{j_\alpha}| \otimes |x_{j_\alpha}\rangle\langle x_{j_\alpha}|$ is the CP map that corresponds to a projection onto $|x_{j_\alpha}\rangle\langle x_{j_\alpha}|$.

As \mathcal{I} and the completely dephasing map \mathcal{D} have the same effect on this classical comb, *i.e.*,

$$C_{\Lambda_k}^{\text{cl}} \star_{\alpha \in \Lambda_{k'}} \Phi_{\alpha}^+ = C_{\Lambda_k}^{\text{cl}} \star_{\alpha \in \Lambda_{k'}} D_{\alpha} \quad \forall \Lambda_{k'} \subseteq \Lambda_k, \quad (\text{A.15})$$

the consistency property of the joint probability distributions induces a consistency property of the family of classical combs $\{C_{\Lambda_k}^{\text{cl}}\}$ constructed via to Eq. (A.13). Then, according to the GET, there exists a classical comb C_{Λ}^{cl} that has all the finite combs $C_{\Lambda_k}^{\text{cl}}$ as ‘marginals’, *i.e.*, $C_{\Lambda_k}^{\text{cl}} = C_{\Lambda}^{\text{cl}}|_{\Lambda_k}$ for all finite $\Lambda_k \subseteq \Lambda$. This implies the existence of a joint probability distribution \mathbb{P}_{Λ} that has all finite \mathbb{P}_{Λ_k} as marginals, which proves the original KET as a corollary of the GET. \square

A.4 PROPERTIES OF $1 \rightarrow 3$ ADAPTERS

As mentioned in the main text, the ordering of the wires g and c of the $1 \rightarrow 3$ adapter with respect to the remaining wires can be chosen freely (see Fig. A.1). Here, we choose it such that a and b can influence measurement statistics at g , and c can influence d (and obviously g), and everything that comes after d . With this choice, we can represent the adapter $A_{1 \rightarrow 3}$ in a more explicit way in terms of a quantum network (see Fig. A.2), consisting of a quantum state ρ , two deterministic quantum maps L and K and the deterministic effect $\mathbb{1}$. The arrangement of these elements makes the the causal relation between events manifest.

From Fig. A.2, we can derive the corresponding $1 \rightarrow 3$ adapter:

$$A_{1 \rightarrow 3} = \rho_{a'a} \star L_{dd'ca'b} \star K_{gf'fd'e} \star \mathbb{1}_{f'h}, \quad (\text{A.16})$$

where, for better orientation, we have labeled all elements by the Hilbert spaces they are defined on, *e.g.*, $\rho_{a'a} \in \mathcal{B}(\mathcal{H}_{a'} \otimes \mathcal{H}_a)$. To avoid confusion, it is important to note that these subscripts do *not* correspond to matrix elements, and we will not address individual matrix elements at any point in this section. Now, if the resulting $A_{1 \rightarrow 3}$ satisfies Eq. (3.21), *i.e.*, if it is a universal $1 \rightarrow 3$ adapter, then the elements of this deterministic network have to satisfy certain conditions. In detail, we have

$$\begin{aligned} A_{1 \rightarrow 3} \star \Phi_{ba}^+ \star \Phi_{hf}^+ &= \rho_{a'a} \star L_{dd'ca'b} \star K_{gf'fd'e} \star \mathbb{1}_{f'h} \star \Phi_{ba}^+ \star \Phi_{hf}^+ \\ &= \rho_{a'a} \star L_{dd'ca'b} \star \text{tr}_{ff'}(K_{gf'fd'e}) := \tilde{L}_{dd'c} \star \tilde{K}_{ged} \\ &\stackrel{!}{=} \Phi_{dc}^+ \otimes \Phi_{ge}^+, \end{aligned} \quad (\text{A.17})$$

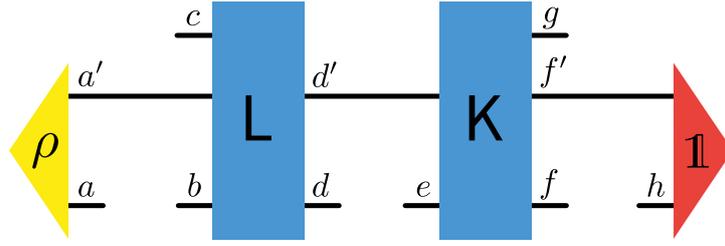


Figure A.2: Possible deterministic quantum network corresponding to a $1 \rightarrow 3$ adapter. L and K are deterministic quantum maps, ρ is a quantum state, and $\mathbb{1}$ is the deterministic effect corresponding to the trace operation. For better orientation, all wires, including the ones that are not open, are labeled.

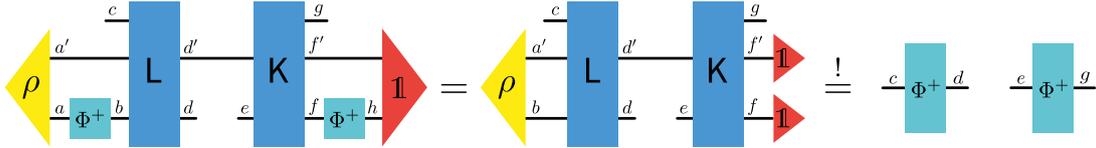


Figure A.3: Graphical representation of Eq. (A.17).

where we have used $\mathbb{1}_{f'h} \star \Phi_{hf}^+ = \mathbb{1}_{f'} \otimes \text{tr}_h \Phi_{hf}^+ = \mathbb{1}_{ff'}$, and the fact that $\rho_{a'a} \star \Phi_{ba}^+$ is simply a relabeling of Hilbert spaces, as Φ_{ba}^+ corresponds to the identity channel (see Fig. A.3).

Writing Eq. (A.17) explicitly, we obtain

$$\text{tr}_{d'} \left[\left(\mathbb{1}_{ge} \otimes \tilde{\mathcal{L}}_{dd'c}^{\text{T}} \right) \left(\mathbb{1}_{cd} \otimes \tilde{\mathcal{K}}_{ged'} \right) \right] = \Phi_{dc}^+ \otimes \Phi_{ge}^+ \quad (\text{A.18})$$

To make this requirement on $\tilde{\mathcal{L}}_{dd'c}$ and $\tilde{\mathcal{K}}_{ged'}$ more concrete, we express both maps in terms of a basis, *i.e.*,

$$\tilde{\mathcal{L}}_{dd'c} = \sum_{ij} r^{(ij)} \rho_{dc}^{(i)} \otimes \eta_{d'}^{(j)} \quad \text{and} \quad \tilde{\mathcal{K}}_{ged'} = \sum_{pq} s^{(pq)} \zeta_{ge}^{(p)} \otimes \Delta_{d'}^{(q)*}, \quad (\text{A.19})$$

where³ $r^{(ij)}, s^{(pq)} \in \mathbb{R}$, $\{\rho_{dc}^{(i)}\}$ is a basis of $\mathcal{B}(\mathcal{H}_d \otimes \mathcal{H}_c)$, $\{\eta_{d'}^{(j)}\}$ is a basis of $\mathcal{B}(\mathcal{H}_{d'})$, $\{\zeta_{ge}^{(p)}\}$ is a basis of $\mathcal{B}(\mathcal{H}_g \otimes \mathcal{H}_e)$, and $\{\Delta_{d'}^{(q)}\}$ is the dual set to $\{\eta_{d'}^{(j)}\}$, *i.e.*, $\text{tr} \left(\Delta_{d'}^{(q)\dagger} \eta_{d'}^{(j)} \right) = \delta_{qj}$ (see App. B.1 for an explicit construction of dual sets). Additionally, let both bases be such that $\rho_{dc}^{(1)} = \Phi_{dc}^+$ and $\zeta_{ge}^{(1)} = \Phi_{ge}^+$. With this representation, Eq. (A.18) reads

$$\begin{aligned} \text{tr}_{d'} \left[\left(\mathbb{1}_{ge} \otimes \tilde{\mathcal{L}}_{dd'c}^{\text{T}} \right) \left(\mathbb{1}_{cd} \otimes \tilde{\mathcal{K}}_{ged'} \right) \right] &= \sum_{ijpq} r^{(ij)} s^{(pq)} \text{tr} \left(\Delta_{d'}^{(q)\dagger} \eta_{d'}^{(j)} \right) \rho_{dc}^{(i)} \otimes \zeta_{ge}^{(p)} \\ &= \sum_j r^{(1j)} s^{(1j)} \Phi_{dc}^+ \otimes \Phi_{ge}^+ + \sum_{\substack{i=2 \\ p=2}} \sum_j r^{(ij)} s^{(pj)} \rho_{dc}^{(i)} \otimes \zeta_{ge}^{(p)}. \end{aligned} \quad (\text{A.20})$$

In order for Eq. (A.18) to hold, the second term in Eq. (A.20) has to vanish, which is equivalent to $\sum_j r^{(ij)} s^{(pj)} = 0$ for all $i, p \geq 2$. While this requirement is in general not easy to fulfill in conjunction with the remaining requirements imposed on $r^{(ij)}$ and $s^{(pq)}$ – both $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{K}}$ have to be positive and trace preserving – it is, for example sufficient to choose the parameters such that if $r^{(ij)} \neq 0$ for some $i \geq 2$, then $s^{(pj)} = 0$ and vice versa. While this is a stricter requirement, it allows for an easy construction of $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{K}}$, such that they are positive, trace preserving, and Eq. (A.18) is satisfied. Importantly, neither $\tilde{\mathcal{L}}$ nor $\tilde{\mathcal{K}}$ have to be of product form, which implies

³ We use superscripts here, to not create the impression that $r^{(ij)}$ and $s^{(pq)}$ are operators on Hilbert spaces, respectively.

that there are non-trivial $1 \rightarrow 3$ adapters of the causal order we discussed in this section. In a similar way to the construction above, one could go from \tilde{L} to L and ρ , and from \tilde{K} to K .

B

OPEN QUANTUM SYSTEM DYNAMICS

B.1 DUAL MATRICES

In this appendix, for comprehensiveness, we provide the proof for the existence of a set of dual matrices for any set of linearly independent matrices $\{A_j\}$ and provide an explicit construction. Note, that this proof is a slight generalization of the original one, presented in [68] for the case of Hermitian matrices A_j .

LEMMA B.1: *For any set of linearly independent matrices $\{A_j\}$, there exists a dual set $\{\Delta_k\}$ satisfying $\text{tr}[\Delta_k^\dagger A_j] = \delta_{jk}$.*

Proof. Write $A_j = \sum_\ell g_{j\ell} \Gamma_\ell$, where $g_{j\ell}$ are complex numbers and $\{\Gamma_\ell\}$ form a Hermitian, self-dual linearly independent basis of the space A_j is defined on, satisfying $\text{tr}[\Gamma_m \Gamma_n] = 2\delta_{mn}$ [60]. Since the matrices $\{A_j\}$ constitute a linearly independent set, the columns of matrix $G = \sum_{ij} g_{ij} |j\rangle\langle i|$ are linearly independent vectors, which means that G has an inverse. Let the matrix $F^\dagger = G^{-1}$, then $GF^\dagger = \mathbb{1}$, implying that the columns of F^* are orthonormal to the columns of G . We define $\Delta_k = \frac{1}{2} \sum_\ell f_{k\ell} \Gamma_\ell$, where $f_{k\ell} = \langle k|F|l\rangle$ are elements of F . \square

Our definition of dual matrices differs from the one in [68] by an adjoint to make the relation to the scalar product, *i.e.*, the Hilbert-Schmidt product, explicit. In general, the dual matrices are not all positive, even if the basis $\{A_j\}$ only consists of positive matrices. However, for the case where all basis matrices A_j are Hermitian, we have $\Delta_k^\dagger = \Delta_k$, as in this case all f_{kl} are real. Furthermore, the duals satisfy $\sum_k \Delta_k^\dagger = \sum_k \Delta_k^* = \mathbb{1}$ if all A_j are of unit trace. We have

$$\text{tr} \left(\sum_k \Delta_k^\dagger A \right) = \sum_{k,\ell} r_\ell \text{tr} \left(\Delta_k^\dagger A_\ell \right) = \sum_\ell r_\ell = \text{tr}(A) \quad \forall A, \quad (\text{B.1})$$

where we have used $A = \sum_\ell r_\ell A_\ell$. The only matrix M that satisfies $\text{tr}(MA) = \text{tr}(A)$ for all A is the identity matrix.

B.2 PROOF THAT OCP DIVISIBILITY \neq ICP DIVISIBILITY

Here, we give an alternative, continuous example of oCP-divisible dynamics that are not iCP-divisible. Let the system and the environment both be qubits, and let the initial environment

state at time u be maximally mixed, i.e., $\eta_u = \mathbb{1}/2$. In what follows, without loss of generality, we choose $u = 0$. The system-environment dynamics is given by a partial swap

$$U_{v|0} = \exp(-i\Omega v \mathbf{S}) = \cos(\Omega v) \mathbb{1}_4 - i \sin(\Omega v) \mathbf{S} \quad (\text{B.2})$$

where $\mathbf{S} |ij\rangle = |ji\rangle$ is the action of the swap operator between system and environment. For this dynamics, the system state at time v is given by $\rho_v = \cos^2(\Omega v) \rho_0 + \sin^2(\Omega v) \mathbb{1}_2/2$, where ρ_0 is the system state at time $u = 0$. Denoting by $\mathcal{R}_{\mathbb{1}}$ the point map that replaces every state by $\mathbb{1}/2$, we see that

$$\mathcal{L}_{v|0} = \cos^2(\Omega v) \mathcal{I} + \sin^2(\Omega v) \mathcal{R}_{\mathbb{1}}, \quad (\text{B.3})$$

which is invertible for $\Omega v \in [0, \pi/2)$. We have

$$\begin{aligned} \mathcal{K}_{w|v} &= \mathcal{L}_{w|0} \circ \mathcal{L}_{v|0}^{-1} \\ &= \frac{\cos^2(\Omega w)}{\cos^2(\Omega v)} \mathcal{I} + \frac{\cos^2(\Omega v) - \cos^2(\Omega w)}{\cos^2(\Omega v)} \mathcal{R}_{\mathbb{1}} \end{aligned} \quad (\text{B.4})$$

which is CP for $w \geq v$ with $\Omega w \in [0, \pi/2)$, and, consequently, the dynamics is iCP-divisible in this interval. However, it is not oCP-divisible. Between time $u = 0$ and v , the environment is partially swapped with the initial state ρ_0 . Subsequently, after the system state is discarded and freshly prepared at time v , between time v and w , the system is partially swapped again with the environment state, which depends on the state of the system at $u = 0$. Consequently, there is conditional signaling from time $u = 0$ to time w , and the process is not oCP-divisible.

B.3 CONDITIONAL NON-SIGNALING PROCESSES THAT ARE NOT OCP-DIVISIBLE

As mentioned in Sec. 4.8.1, conditional non-signaling is necessary for oCP divisibility to hold, but not sufficient. To see this, consider the following example: Let $\eta_{e_u e_v} = \frac{1}{4} (\mathbb{1}_{e_u e_v} + \sigma_{e_u}^{(x)} \otimes \sigma_{e_v}^{(z)})$ be a correlated two-qubit environment state with $e_u e_v := e$, where $\{\sigma^{(x)}, \sigma^{(y)}, \sigma^{(z)}\}$ are the Pauli matrices. Initially, i.e., at time u , the one-qubit system is uncorrelated with the environment. Let the system-environment dynamics between time u and time v be given by the swap \mathbf{S}_{se_u} . If the system-environment dynamics from v to w only acts non-trivially on s and e_v , then there is no conditional signaling between u and w . Nonetheless, the process is not necessarily oCP-divisible.

For example, if the unitary evolution $\mathcal{U}_{w|v}$ between v and w is given by the unitary matrix $U_{w|v} = \frac{1}{\sqrt{3}} (\mathbb{1}_{se_v} + i\sigma_s^{(y)} \otimes \mathbb{1}_{e_v} + i\sigma_s^{(x)} \otimes \sigma_{e_v}^{(z)})$, with trivial action on e_u . With this, the final system state at time v , without intervention at at time v is $\tilde{\rho} = \frac{1}{2} \mathbb{1}_s + \frac{1}{3} \sigma_s^{(y)}$, independently of the input state at u . Consequently, the action of the overall map $\mathcal{L}_{w|u}$ can be written as $\mathcal{L}_{w|u}[\rho_u] = \text{tr}(\rho_u) \tilde{\rho}$. On the other hand, the map $\mathcal{L}_{v|u}$ simply replaces the system state at u with $\text{tr}_{e_v}(\eta_{e_u e_v}) = \frac{1}{2} \mathbb{1}_{e_u}$, which means, that for oCP divisibility to hold, the map $\mathcal{L}_{w|v}$ would have to be of the form $\mathcal{L}_{w|v}[\rho_v] = \text{tr}(\rho_v) \tilde{\rho}$. However, it is easy to check, that for an input state

$$\rho_v = \frac{1}{2} (\mathbb{1}_s + a\sigma_s^{(x)} + b\sigma_s^{(y)} + c\sigma_s^{(z)}), \quad (\text{B.5})$$

at time v , the corresponding output state at time w is given by

$$\mathcal{L}_{w|v}[\rho_v] = \frac{1}{2}\mathbb{1}_s + \frac{1}{6}(a-2b)\sigma_s^{(x)} + \frac{b}{2}\sigma_s^{(y)} + \frac{1}{6}(2a-c)\sigma_s^{(z)}, \quad (\text{B.6})$$

and the process is thus not oCP-divisible.

B.4 STRUCTURAL PROPERTIES OF PROCESSES WITH FINITE MARKOV ORDER

If a quantum process is of Markov order $|M| = \ell$, then there exists an instrument \mathcal{J}_M , such that for every CP map M_{x_M} of the instrument, the past and future statistics are independent. As we have seen in the main text, this implies, that for all history and future instruments \mathcal{J}_H and \mathcal{J}_F we have

$$\begin{aligned} & \frac{\mathbb{T}_{FMH} \star M_{x_F} \star M_{x_M} \star M_{x_H}}{\mathbb{T}_{MH} \star M_{x_M} \star M_{x_H}} \\ &= \frac{\sum_{x_H} \mathbb{T}_{FMH} \star M_{x_F} \star M_{x_M} \star M_{x_H}}{\sum_{x_H} \mathbb{T}_{MH} \star M_{x_M} \star M_{x_H}} := \frac{\mathbb{T}_{FMH} \star M_{x_F} \star M_{x_M} \star M_H}{\mathbb{T}_{MH} \star M_{x_M} \star M_H}, \end{aligned} \quad (\text{B.7})$$

where $M_H = \sum_{x_H} M_{x_H}$ is the CPTP map (or, more precisely, the deterministic comb, for the case of a temporally correlated instrument) corresponding to the instrument \mathcal{J}_H . By causality, we know that $\mathbb{T}_{MH} := \sum_{x_F} \mathbb{T}_{FMH} \star M_{x_F}$ is independent of the instrument \mathcal{J}_F . Eq. (B.7) has to hold for all instruments \mathcal{J}_H . As we can vary the CP maps M_{x_H} without changing the overall CPTP map M_H , this implies that

$$\frac{\mathbb{T}_{FMH} \star M_{x_F} \star M_{x_M} \star M_{x_H}}{\mathbb{T}_{MH} \star M_{x_M} \star M_{x_H}} = \frac{\mathbb{T}_{FMH} \star M_{x_F} \star M_{x_M} \star M_{x'_H}}{\mathbb{T}_{MH} \star M_{x_M} \star M_{x'_H}}, \quad (\text{B.8})$$

for all CP maps $M_{x_H}, M_{x'_H}$. As the above equality also has to hold for all M_{x_F} , we see that

$$\frac{(\mathbb{T}_{FMH} \star M_{x_M}) \star M_{x_H}}{\mathbb{T}_{MH} \star M_{x_M} \star M_{x_H}} = \frac{(\mathbb{T}_{FMH} \star M_{x_M}) \star M_{x'_H}}{\mathbb{T}_{MH} \star M_{x_M} \star M_{x'_H}}. \quad (\text{B.9})$$

Put differently, this implies that

$$(\mathbb{T}_{FMH} \star M_{x_M}) \star M_{x_H} \propto (\mathbb{T}_{FMH} \star M_{x_M}) \star M_{x'_H}, \quad (\text{B.10})$$

and as $(\mathbb{T}_{FMH} \star M_{x_M}) \star M_{x_H}$ is a matrix, the only way for this proportionality to hold for all possible CP maps M_{x_H} , is if $\mathbb{T}_{FMH} \star M_{x_M}$ is of product form, *i.e.*,

$$\mathbb{T}_{FMH} \star M_{x_M} = \mathbb{T}_F^{(x_M)} \otimes \tilde{\mathbb{T}}_H^{(x_M)}, \quad (\text{B.11})$$

where $\mathbb{T}_F^{(x_M)}$ is by construction a causally ordered process tensor on the times F , while the structure of $\tilde{\mathbb{T}}_H^{(x_M)}$ is *a priori* not clear. In general, $\tilde{\mathbb{T}}_H^{(x_M)}$ constructed via contraction of a process tensor with a CP operation M_{x_M} (and tracing over the times F) does not have to be a properly causally ordered object (see Ch. 6). If $\tilde{\mathbb{T}}_H^{(x_M)}$ is not causally ordered, then the probability to obtain outcomes on the memory block depends on the instrument that was used at times H ; we have

$$\mathbb{P}(x_M | \mathcal{J}_M, \mathcal{J}_H) = \tilde{\mathbb{T}}_H^{(x_M)} \star M_H \quad (\text{B.12})$$

which in general depends on M_H . This dependence on past instruments is in line with the breakdown of Kolmogorov conditions that we discussed in Ch. 3. However, when summing over the possible outcomes x_M , we must have $\sum_{x_M} \mathbb{P}(x_M | \mathcal{J}_M, \mathcal{J}_H) = 1$ for all possible instruments \mathcal{J}_H , which implies

$$\sum_{x_M} \tilde{\mathbb{T}}_H^{(x_M)} \star M_H = 1 \quad \forall M_H. \quad (\text{B.13})$$

Consequently, the positive matrices $\mathbb{T}_H^{(x_M)}$ have to add up to a deterministic comb $\mathbb{T}_H^{(\mathcal{J}_M)} = \sum_{x_M} \tilde{\mathbb{T}}_H^{(x_M)}$ when summed over the possible outcomes x_M , which means that the set $\{\tilde{\mathbb{T}}_H^{(x_M)}\}$ forms a tester; every CP operation M_{x_M} corresponding to an outcome of an instrument that blocks the memory between F and H leaves the process tensor in a product of a deterministic comb $\mathbb{T}_F^{(x_M)}$ on the future F , and an element $\tilde{\mathbb{T}}_H^{(x_M)}$ of a tester on the history. This mathematical structure coincides nicely with physical intuition, since, when conditioning on an outcome of the memory blocking instrument, the statistics of future and history have to be independent.

PROCESS TENSORS AND LIMITED RESOURCES

C.1 BASIS OF UNITARY MAPS ACTING ON A QUBIT

Here, we construct a basis of the space of unitary maps for the case of qubits, *i.e.*, $d = 2$. Importantly, this does not amount to constructing a basis of the set of unitary matrices. There are d^2 unitary matrices that span the space of unitary matrices [255, 256], while the space of unitary maps is $(d^2 - 1)^2 + 1$ -dimensional. This is analogous to the case of pure states and density matrices. Every vector corresponding to a pure state of a d -dimensional system can be represented as a linear combination of d basis vectors, while one needs d^2 projectors to represent every projector of a d -dimensional system.

Any unitary matrix $V \in \text{SU}(2)$ can be expressed in terms of Pauli matrices $\{\sigma_k\}$ (where $k \in \{x, y, z\}$) in the following form:

$$V_{(\vartheta, \vec{a})} = \cos\left(\frac{\vartheta}{2}\right) \mathbb{1} - i \sin\left(\frac{\vartheta}{2}\right) \sum_{\ell=1}^3 a_\ell \sigma_\ell, \quad (\text{C.1})$$

where $|\vec{a}| = 1$. Hence, a generic unitary 1-qubit map is of the form

$$\begin{aligned} \mathcal{V}_{(\vartheta, \vec{a})}[\rho] = & c_{\vartheta/2}^2 \underbrace{\mathbb{1}\rho\mathbb{1}}_{(I)} + s_{\vartheta/2}^2 \sum_{k=1}^3 a_k^2 \underbrace{(\sigma_k \rho \sigma_k)}_{(II)} + c_{\vartheta/2} s_{\vartheta/2} \sum_{k=1}^3 a_k \underbrace{i(\mathbb{1} \rho \sigma_k - \sigma_k \rho \mathbb{1})}_{(III)} \\ & + s_{\vartheta/2}^2 \sum_{k < \ell}^3 a_k a_\ell \underbrace{(\sigma_k \rho \sigma_\ell + \sigma_\ell \rho \sigma_k)}_{(IV)}, \end{aligned} \quad (\text{C.2})$$

where $s_{\vartheta/2} := \sin(\vartheta/2)$ and $c_{\vartheta/2} := \cos(\vartheta/2)$. The term (I) in Eq. (C.2) can be accounted for by choosing $V_0 = \mathbb{1}$. Following the analogous derivation for the case of projective maps in [144], we set

$$V_{(k, \pm)} = \frac{1}{\sqrt{2}} (\mathbb{1} \pm i\sigma_k). \quad (\text{C.3})$$

With the six matrices given in Eq. (C.3) both (II) and (III) can be obtained:

$$\begin{aligned} \sigma_k \rho \sigma_k &= 2 \left(\mathcal{V}_{(k,+)} + \mathcal{V}_{(k,-)} \right) [\rho] - \mathcal{V}_0[\rho] \\ \text{and } i(\mathbb{1} \rho \sigma_k - \sigma_k \rho \mathbb{1}) &= 2 \left(\mathcal{V}_{(k,-)} - \mathcal{V}_{(k,+)} \right) [\rho]. \end{aligned} \quad (\text{C.4})$$

The three remaining terms (IV) can be obtained with the three additional unitary matrices

$$V_{(k+\ell+1,+)} = \frac{1}{\sqrt{2}} \left(\mathbb{1} + \frac{i}{\sqrt{2}} \sigma_k + \frac{i}{\sqrt{2}} \sigma_\ell \right), \quad (\text{C.5})$$

with $k < \ell$. We have

$$\begin{aligned} \sigma_k \rho \sigma_\ell + \sigma_\ell \rho \sigma_k = & 2 \left\{ 2 \mathcal{V}_{(k+\ell+1,+)} - (1 + \sqrt{2}) \left(\mathcal{V}_{(k,+)} + \mathcal{V}_{(\ell,+)} \right) \right. \\ & \left. - (1 - \sqrt{2}) \left(\mathcal{V}_{(k,-)} + \mathcal{V}_{(\ell,-)} \right) \right\} [\rho]. \end{aligned} \quad (\text{C.6})$$

Hence, any unitary map $\mathcal{V}[\rho] = V\rho V^\dagger$ acting on a qubit can be represented as a linear combination of the ten unitary maps $\{\mathcal{V}_0, \mathcal{V}_{(k,\pm)}, \mathcal{V}_{(k+\ell+1,+)}\}$. A one-step process tensor constructed based on this set of operations can meaningfully be applied to any completely positive map that lies in its linear span, which, in this case, is the set of all one-qubit unital maps.

C.2 BASIS OF PROJECTIVE MAPS

As was the case for unitary maps and unitary matrices, constructing a basis of projective maps is not equivalent to constructing a basis of the space of projectors. The action of any rank-1 projective map \mathcal{P} acting on a d -dimensional state ρ can be written as:

$$\mathcal{P}[\rho] = \Pi\rho\Pi, \quad (\text{C.7})$$

where $\Pi = \sum_{k,\ell=1}^d c_k c_\ell^* |k\rangle \langle \ell|$ is a (d -dimensional) pure state and $\{|k\rangle\}_{k=1}^d$ is a set of d orthogonal states. The Choi matrix of the map \mathcal{P} has the form $\Pi \otimes \Pi^\dagger$. Any map $\mathsf{N} \in W_P$ can be represented as

$$\mathsf{N} = \sum_{\nu} b_{\nu} \Pi_{\nu} \otimes \Pi_{\nu}^{\dagger} = \sum_{\nu} b_{\nu} \sum_{\substack{k,\ell=1 \\ k',\ell'=1}}^d c_k^{(\nu)} c_\ell^{(\nu)*} c_{k'}^{(\nu)*} c_{\ell'}^{(\nu)} |kk'\rangle \langle \ell\ell'|. \quad (\text{C.8})$$

The elements of any $\mathsf{N} \in W_P$ with respect to the basis $|kk'\rangle \langle \ell\ell'|$ possess the following symmetry properties:

$$(1) \mathsf{N}_{kk';\ell\ell'} = \mathsf{N}_{\ell'\ell;k'k} \quad (2) \mathsf{N}_{kk';\ell\ell'} = \mathsf{N}_{k\ell;k'\ell'} \quad (3) \mathsf{N}_{kk';\ell\ell'}^* = \mathsf{N}_{k'\ell';\ell\ell}. \quad (\text{C.9})$$

By counting the number of remaining independent entries in the matrix N , one can deduce that the vector space of matrices with the properties given by Eq. (C.9) is $\frac{1}{4}d^2(d+1)^2$ -dimensional. In principle, it remains to be shown that W_P actually coincides with this vector space, as the matrix N given in Eq. (C.8) could have further symmetries than the ones accounted for in Eq. (C.9). However, for the qubit case, a set of $\frac{1}{4}d^2(d+1)^2 = 9$ pure states $\{\Pi_{\mu}\}$ that yields linearly independent projective maps has been constructed in Ref. [144]:

$$\Pi_{(\omega,\pm)} = \frac{1}{2}(\mathbb{1} \pm \sigma_{\omega}) \quad \text{and} \quad \Pi_{(\varepsilon+\kappa+1)} = \frac{1}{2} \left(\mathbb{1} + \frac{1}{\sqrt{2}}\sigma_{\varepsilon} + \frac{1}{\sqrt{2}}\sigma_{\kappa} \right), \quad (\text{C.10})$$

where $\omega, \varepsilon, \kappa \in \{1, 2, 3\}$, $\varepsilon < \nu$, and $\{\sigma_1, \sigma_2, \sigma_3\} \equiv \{\sigma_x, \sigma_y, \sigma_z\}$. For other low-dimensional cases, it is possible to numerically find $\frac{1}{4}d^2(d+1)^2$ linearly independent projective maps. Here, we have shown that the set of projective maps is at most $\frac{1}{4}d^2(d+1)^2$ -dimensional, and we shall leave the statement that this upper bound is tight as a conjecture.

C.3 DERIVATION OF INTERMEDIATE RESTRICTED PROCESS TENSORS

Here, we show how to obtain intermediate restricted process tensors in the case where $\text{Span}(R)$ contains an informationally complete set of measurements. We have already seen that a restricted process tensor can be reconstructed, once the output state for any performable sequence of operations is known. If the set of performable operations allows for an informationally complete measurement at each time, given $\mathbb{T}_{\Lambda_K}^R$, we can deduce the state of the system at *any* time $t_\alpha \in \Lambda_K$ for any sequence of performable previous operations. Thus, we can reconstruct any intermediary restricted process tensor on times $\Lambda_k \subseteq \Lambda_K$. To keep notational overhead to a minimum, we shall show this explicitly for the derivation of $\mathbb{T}_{\{t_4, t_2\}}^R$ from $\mathbb{T}_{5:1}^R$, the case already used in the main text as a guiding example. The general case then follows straightforwardly.

From Eq. (5.9), we know that given $\mathbb{T}_{5:1}^R$ we can compute the correct restricted process tensor $\mathbb{T}_{\{t_5, t_4, t_2\}}^R$ defined on times $\{t_5, t_4, t_2\}$ by inserting identity maps at times $\{t_1, t_3\}$. Now, let $\{M_{x_4} \in W^R\}_{x_4=1}^{d^2}$ be a set of experimentally performable CP maps that corresponds to an informationally complete measurement at t_4 . We can calculate the probability $\mathbb{P}(x_4, x_2)$ to obtain an outcome x_4 corresponding to M_{x_4} and an outcome corresponding to the operation $M_{x_2} \in R$ at t_2 via

$$\mathbb{P}(x_4, x_2) = \mathbb{T}_{\{t_5, t_4, t_2\}}^R \star \mathbb{1}_5^0 \star M_{x_4} \star M_{x_2}, \quad (\text{C.11})$$

where we suppress the dependence of the probabilities on the respective instruments that are used to interrogate the system, and it is understood that outcome x_α corresponds to the map M_{x_α} . Each of the CP maps $M_{x_4} \in \mathcal{B}(\mathcal{H}_4^1 \otimes \mathcal{H}_4^0)$ has a corresponding POVM element $E_{x_4} = \text{tr}_{4^1} M_{x_4}$. With this, we can compute the state $\rho'_{4|x_2}$ of the system at time t_4 , given that the operation M_{x_2} was performed at t_2 :

$$\rho'_{4|x_2} = \sum_{x_4} \mathbb{P}(x_4, x_2) \mathbf{e}_{x_4}^*, \quad (\text{C.12})$$

where $\{\mathbf{e}_{x_4}\}_{x_4=1}^{d^2}$ is the dual set to $\{E_{x_4}\}_{x_4=1}^{d^2}$. Indeed, the states given by the above equation yield the correct probabilities for every element of the informationally complete measurement at time t_4 . If the operations M_{x_2} form a basis of $\text{Span}(R)$, then Eq. (C.12) yields the output states for a full basis, and as such allows for the reconstruction of $\mathbb{T}_{\{t_4, t_2\}}^R$ via

$$\mathbb{T}_{\{t_4, t_2\}}^R = \sum_{x_2} \rho'_{4|x_2} \otimes \mathbf{m}_{x_2}^*, \quad (\text{C.13})$$

where $\{\mathbf{m}_{x_2}\}_{x_2=1}^{d^2}$ is the dual set to $\{M_{x_2}\}_{x_2=1}^{d^2}$. The generalization to more times follows in a straightforward manner.

While the algebraic derivation of intermediate restricted process tensors appears somewhat cumbersome, the underlying intuitive idea is very clear: If the set of available operations allows for an informationally complete set of measurements, the the restricted process tensor $\mathbb{T}_{N+1:1}^R$ contains enough information to deduce the state of the system at every time $t_\alpha \in \{t_{N+1}, \dots, t_1\}$, and intermediate restricted process tensors with an open final output wire can be derived.

C.4 NDGD AND CLASSICALITY

To prove theorem 5.1, we first note that measurements in the classical basis commute with the completely dephasing map, *i.e.*, for every projective measurement \mathcal{P}_{x_α} (say, at time t_α) in the classical basis, with outcome x_α , we have

$$\mathcal{P}_{x_\alpha} = \mathcal{D}_\alpha \circ \mathcal{P}_{x_\alpha} = \mathcal{P}_{x_\alpha} \circ \mathcal{D}_\alpha. \quad (\text{C.14})$$

Now, following the considerations in the main text, the probability to measure the outcomes $\{x_1, \dots, x_N\}$ at times $\{t_1, \dots, t_N\}$ is given by

$$\mathbb{P}(x_1, \dots, x_N) = p_{x_1} |x_1\rangle\langle x_1| \star \mathbf{G}_1 \star \mathcal{P}_{x_2} \star \dots \star \mathcal{P}_{x_{N-1}} \star \mathbf{G}_{N-1} \star \mathcal{P}_{x_N} \star \mathbf{1}_N^{se_N}, \quad (\text{C.15})$$

where p_{x_1} is the probability to prepare state $|x_1\rangle\langle x_1|$ at time t_1 , $\mathcal{P}_{x_\alpha} = |x_\alpha\rangle\langle x_\alpha| \otimes |x_\alpha\rangle\langle x_\alpha|$ is the map corresponding to the projective measurement in the classical basis that yields outcome x_α at time t_α , and the initial state of the environment is absorbed into \mathbf{G}_1 . Here, we run into a slight notational asymmetry considering the first time t_1 , where we conduct a preparation – denoted by $|x_1\rangle\langle x_1|$ – instead of a projective measurement – denoted by $|x_1\rangle\langle x_1| \otimes |x_1\rangle\langle x_1|$ – which stems from the fact that we considered the initial system state to be an experimental choice, instead of part of the uncontrollable underlying process. While being notationally cleaner, the latter would have made the connection to the results for Markovian dynamics less direct, which is why we opted to keep the initial state an experimental choice. Naturally, this has no bearing on the correctness of the results, and either way, NDGD dynamics lead to classical statistics.

Using the property that the dynamics is NDGD, we can now show that averaging over a time step is the same as not performing an operation from the point of view of a classical observer. To this end, we shall consider the average over a time $t_\alpha \leq t_N$. The corresponding relevant part of Eq. (C.15) is

$$\begin{aligned} & \dots \star \mathcal{P}_{x_{\alpha-1}} \star \mathbf{G}_{\alpha-1} \star \mathcal{D}_\alpha \star \mathbf{G}_\alpha \star \mathcal{P}_{x_\alpha} \star \dots \\ &= \dots \star \mathcal{P}_{x_{\alpha-1}} \star \mathcal{D}_{\alpha-1} \star \mathbf{G}_{\alpha-1} \star \mathcal{D}_\alpha \star \mathbf{G}_\alpha \star \mathcal{D}_{\alpha+1} \star \mathcal{P}_{x_{\alpha+1}} \star \dots \\ &= \dots \star \mathcal{P}_{x_{\alpha-1}} \star \mathbf{G}_{\alpha-1} \star \Phi_\alpha^+ \star \mathbf{G}_\alpha \star \mathcal{P}_{x_\alpha} \star \dots, \end{aligned} \quad (\text{C.16})$$

where in the second line we employed Eq. (C.14), and in the last line we used that the dynamics is NDGD, *i.e.*, it satisfies Eq. (5.53). As we can successively replace all occurring completely dephasing maps by identity maps in the same manner, dynamics that are NDGD display classical statistics when probed by measurements in the classical basis.

On the other hand, it is relatively straightforward to construct examples of dynamics that are not NDGD but display classical dynamics. For instance, one can consider the following example, which is similar in spirit to Ex. 5.1 (see Fig. C.1 for a graphical representation): Let the system be a qubit, and let the initial environment be in a maximally entangled two-qubit state Φ_{ee}^+ . The first dynamics \mathbf{G}_1 from t_1 to t_2 swaps the system with one half (denoted by e) of the environment state and then discards it. The dynamics \mathbf{G}_2 from t_2 to t_3 yields a system-environment state $\mathbf{1}_s/2 \otimes |0\rangle\langle 0|$ if the se' input state is $\Phi_{se'}^+$, and $\mathbf{1}_s/2 \otimes |1\rangle\langle 1|$ otherwise. Consequently, when the completely dephasing map is applied at t_2 , the system-environment state at t_3 is $\mathbf{1}_s/2 \otimes \mathbf{1}_e/2$, while it is equal to $\mathbf{1}_s/2 \otimes |1\rangle\langle 1|$ if the identity map

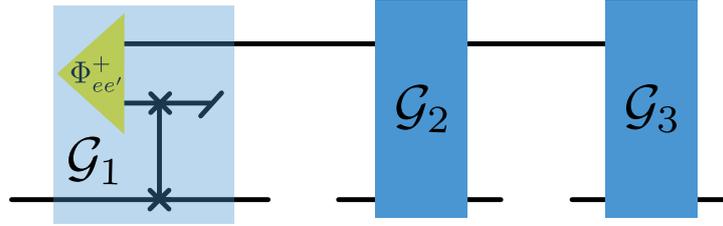


Figure C.1: Non-NDGD dynamics that leads to classical statistics. The first map \mathcal{G}_1 (blue transparent box) swaps the system with one half of a maximally entangled state and then discarded. The subsequent CPTP map \mathcal{G}_2 maps Φ^+ and $\mathbb{1}/4$ onto two different system-environment states with the same reduced system state $\rho_{t_3} = \mathbb{1}/2$. The final CPTP map \mathcal{G}_3 is such that it induces a unital dynamics on the system. Consequently, the system state at t_2 , t_3 , and t_4 is maximally mixed independent of whether the completely dephasing, or the identity map was implemented at t_2 and t_3 .

was implemented, and as such, the dynamics is not NDGD. However, the system state is always maximally mixed, independent of whether \mathcal{D}_2 or \mathcal{I}_2 was implemented. To make the example non-trivial, we add a third dynamics \mathcal{G}_3 from t_3 to t_4 . We choose this dynamics such that it induces a unital dynamics on the level of the system, independent of the environment state at t_3 . For example, this happens when the corresponding system-environment Hamiltonian is of product form, *i.e.*, $H_{se} = H_s \otimes H_e$, independent of the explicit form of the respective terms [7]. With this final dynamics, the system state at each of the times t_2 , t_3 , and t_4 is maximally mixed, and the resulting statistics satisfy Kolmogorov conditions, *i.e.*, they are classical.

D

CAUSALLY INDEFINITE PROCESSES

D.1 ALLOWED TERMS IN THE PROCESS MATRIX $W^{B_2 B_1 A_2 A_1}$

For completeness, here, we provide the generalized Pauli terms that cannot appear in a process matrix that respects local causality. Originally, they were derived in [54].

Process matrices must respect local causality. This requirement is expressed explicitly by the requirement

$$\text{tr}[(M^A \otimes M^B)W^{AB^T}] = 1 \quad \forall \text{ CPTP maps } M^A, M^B. \quad (\text{D.1})$$

The process matrix W^{AB} is positive (and Hermitian). Consequently, it can be written in the form [54]

$$W^{B^1 B^0 A^1 A^0} = \sum_{\alpha\beta\gamma\epsilon=0} w_{\alpha\beta\gamma\epsilon} \Gamma_\alpha^{B^1} \otimes \Gamma_\beta^{B^0} \otimes \Gamma_\gamma^{A^1} \otimes \Gamma_\epsilon^{A^0} \quad (\text{D.2})$$

where the matrices $\{\Gamma_a^{X^z}\}_{a=0}^{d_{X^z}^2-1}$ are generalized Pauli matrices, *i.e.*, they are traceless (except for $\Gamma_0^{X^z} = \mathbb{1}_{X^z}$) and $\text{tr}(\Gamma_a^{X^z} \Gamma_b^{X^z}) = d_{X^z} \delta_{ab}$. The prefactor w_{0000} is equal to $\frac{1}{d_{A_1 B_1}}$ for correct normalization. Not all positive matrices $W^{A_2 A_1 B_2 B_1}$ of the form (D.2) satisfy the requirement (D.1) for local causality; in order for (D.1) to hold for all CPTP maps M^B and M^A , W^{AB} can only contain Pauli terms that do not appear in $(M^A)^T \otimes (M^B)^T$ (except for $\mathbb{1}_{A_2 A_1 B_2 B_1}$). Otherwise, it would always be possible to find two valid CPTP maps, such that (D.1) is violated [54].

Using the trace property $\text{tr}_{X^1}(M^{X^1 X^0}) = \mathbb{1}_{X^0}$ of CPTP maps, we can explicitly write down conditions that define the terms that can appear in the decomposition (D.2). In a concise notation, we have

$$\begin{aligned} \text{tr}(\Gamma_\alpha^{X^1} W^T) &= 0, & \text{tr}[(\Gamma_\alpha^{X^1} \otimes \Gamma_\beta^{Y^1}) W^T] &= 0 & \text{tr}[(\Gamma_\alpha^{X^1} \otimes \Gamma_\beta^{X^0}) W^T] &= 0, \\ \text{tr}[(\Gamma_\alpha^{X^1} \otimes \Gamma_\beta^{X^0} \otimes \Gamma_\gamma^{Y^1}) W^T] &= 0, & \text{and} & & \text{tr}[(\Gamma_\alpha^{X^1} \otimes \Gamma_\beta^{X^0} \otimes \Gamma_\gamma^{Y^1} \otimes \Gamma_\epsilon^{Y^0}) W^T] &= 0, \end{aligned} \quad (\text{D.3})$$

where we have omitted the respective identity matrices, $W := W^{B^1 B^0 A^1 A^0}$, $\alpha, \beta, \gamma, \epsilon \geq 1$, $X, Y \in \{B, A\}$ and $X \neq Y$ when they both appear in the same equation. For simplicity of notation, following the convention of [54], we label terms in the decomposition (D.2) of the form $\Gamma_\alpha^{A^1} \otimes \mathbb{1}_{A^0 B^1 B^0}$ ($\alpha \geq 1$) by A^1 , terms of the form $\mathbb{1}_{B^1 B^0} \otimes \Gamma_\alpha^{A^1} \otimes \Gamma_\beta^{A^0}$ ($\alpha, \beta \geq 1$) by $A^1 A^0$, etc.. In this notation, for example, the second equation in (D.3) states that terms of the form $A^1 A^0$ and $B^1 B^0$ cannot appear in a valid process matrix.

BIBLIOGRAPHY

- [1] S. MILZ, M. S. KIM, F. A. POLLOCK, and K. MODI, [arXiv:1901.05223](#) (2019). Accepted for publication in *Phys. Rev. Lett.*
- [2] S. MILZ, F. A. POLLOCK, and K. MODI, *Phys. Rev. A* **98**, 012108 (2018).
- [3] S. MILZ, F. A. POLLOCK, T. P. LE, G. CHIRIBELLA, and K. MODI, *New J. Phys.* **20**, 033033 (2018).
- [4] S. MILZ, F. SAKULDEE, F. A. POLLOCK, and K. MODI, [arXiv:1712.02589](#) (2017).
- [5] P. TARANTO, S. MILZ, F. A. POLLOCK, and K. MODI, *Phys. Rev. A* **99**, 042108 (2019).
- [6] P. TARANTO, F. A. POLLOCK, S. MILZ, M. TOMAMICHEL, and K. MODI, *Phys. Rev. Lett.* **122**, 140401 (2019).
- [7] F. SAKULDEE, S. MILZ, F. A. POLLOCK, and K. MODI, *J. Phys. A* **51**, 414014 (2018).
- [8] S. MILZ, F. A. POLLOCK, and K. MODI, *Open Sys. Info. Dyn.* 1740016 (2017).
- [9] D. SMYTH, *Australia from the Sea : A Thirteenth Book of Paintings, Poetry and Prose*, Smyth, Dacre, Toorak, Vic. (2006).
- [10] S. M. LANE, *Categories for the Working Mathematician*, Springer-Verlag, New York (1978).
- [11] G. M. D'ARIANO, G. CHIRIBELLA, and P. PERINOTTI, *Quantum Theory from First Principles: An Informational Approach*, Cambridge University Press (2017).
- [12] B. COECKE and A. KISSINGER, *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning*, Cambridge University Press (2017).
- [13] C. J. WOOD, J. D. BIAMONTE, and D. G. CORY, *Quant. Inf. Comp.* **15**, 759–811 (2015).
- [14] J. C. BRIDGEMAN and C. T. CHUBB, *J. Phys. A* **50**, 223001 (2017).
- [15] G. CHIRIBELLA, G. M. D'ARIANO, and P. PERINOTTI, *EPL* **83**, 30004 (2008).
- [16] G. CHIRIBELLA, G. M. D'ARIANO, and P. PERINOTTI, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [17] G. CHIRIBELLA, G. M. D'ARIANO, and P. PERINOTTI, *Phys. Rev. A* **80**, 022339 (2009).
- [18] D. KRETSCHMANN and R. F. WERNER, *Phys. Rev. A* **72**, 062323 (2005).
- [19] M. A. NIELSEN and I. L. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).
- [20] E. B. DAVIES and J. T. LEWIS, *Comm. Math. Phys.* **17**, 239–260 (1970).
- [21] K. KRAUS, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Springer, Berlin, Heidelberg (1983).
- [22] A. GLEASON, *J. Math. Mech.* **6**, 885–893 4 (1957).

- [23] P. BUSCH, *Phys. Rev. Lett.* **91**, 120403 (2003).
- [24] C. M. CAVES, C. A. FUCHS, K. K. MANNE, and J. M. RENES, *Found. Phys.* **34**, 193–209 (2004).
- [25] P. PECHUKAS, *Phys. Rev. Lett.* **73**, 1060–1062 (1994).
- [26] R. ALICKI, *Phys. Rev. Lett.* **75**, 3020–3020 16 (1995).
- [27] P. PECHUKAS, *Phys. Rev. Lett.* **75**, 3021–3021 16 (1995).
- [28] F. MASILLO, G. SCOLARICI, and L. SOLOMBRINO, *J. Math. Phys.* **52**, 012101 (2011).
- [29] F. BUSCEMI, *Phys. Rev. Lett.* **113**, 140502 (2014).
- [30] G. ARGENTIERI, F. BENATTI, R. FLOREANINI, and M. PEZZUTTO, *EPL* **107**, 50007 (2014).
- [31] K. MODI, *Sci. Rep.* **2**, 581 (2012).
- [32] F. A. POLLOCK, C. RODRÍGUEZ-ROSARIO, T. FRAUENHEIM, M. PATERNOSTRO, and K. MODI, *Phys. Rev. A* **97**, 012127 (2018).
- [33] I. BENGTSOON and K. ŻYCZKOWSKI, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press (2007).
- [34] B. MIELNIK, *Phys. Lett. A* **289**, 1–8 (2001).
- [35] C. H. BENNETT, G. BRASSARD, C. CRÉPEAU, R. JOZSA, A. PERES, and W. K. WOOTTERS, *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
- [36] L. VAIDMAN, *Phys. Rev. A* **49**, 1473–1476 (1994).
- [37] R. PENROSE, *Applications of negative dimensional tensors*, In: D. J. A. WELSH (ed.), *Combinatorial mathematics and its applications*, Academic Press (1971).
- [38] P. SELINGER, *Electron. Notes Theor. Comput. Sci.* **170**, 139–163 (2007).
- [39] V. BERGHOLM and J. D. BIAMONTE, *J. Phys. A* **44**, 245304 (2011).
- [40] C. J. WOOD, *Initialization and characterization of open quantum systems*, PhD Thesis, University of Waterloo (2015), UWspace.
- [41] M. WILDE, *Quantum Information Theory*, Cambridge University Press (2013).
- [42] O. ORESHKOV and C. GIARMATZI, *New J. Phys.* **18**, 093020 (2016).
- [43] S. SHRAPNEL and F. COSTA, *Quantum* **2**, 63 (2018).
- [44] D. BECKMAN, D. GOTTESMAN, M. A. NIELSEN, and J. PRESKILL, *Phys. Rev. A* **64**, 052309 (2001).
- [45] T. EGGELING, D. SCHLINGEMANN, and R. F. WERNER, *EPL* **57**, 782 (2002).
- [46] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *Phys. Rev. Lett.* **101**, 180501 (2008).
- [47] G. CHIRIBELLA and D. EBLER, *New J. Phys.* **18**, 093053 (2016).
- [48] G. CHIRIBELLA, *Entropy* **20**, 358 (2018).
- [49] J. PEARL, *Causality: Models, Reasoning and Inference*, Cambridge University Press (2009).
- [50] F. COSTA and S. SHRAPNEL, *New J. Phys.* **18**, 063032 (2016).

- [51] J.-M. A. ALLEN, J. BARRETT, D. C. HORSMAN, C. M. LEE, and R. W. SPEKKENS, *Phys. Rev. X* **7**, 031021 (2017).
- [52] O. ORESHKOV and N. J. CERF, *New J. Phys.* **18**, 073037 (2016).
- [53] G. CHIRIBELLA, G. M. D'ARIANO, P. PERINOTTI, and B. VALIRON, *Phys. Rev. A* **88**, 022318 2 (2013).
- [54] O. ORESHKOV, F. COSTA, and Č. BRUKNER, *Nat. Commun.* **3**, 1092 (2012).
- [55] M. JIANG, S. LUO, and S. FU, *Phys. Rev. A* **87**, 022310 (2013).
- [56] M. D. CHOI, *Linear Algebra Appl.* **10**, 285 (1975).
- [57] A. JAMIOŁKOWSKI, *Rep. Math. Phys.* **3**, 275–278 (1972).
- [58] K. KRAUS, *Ann. Phys.* **64**, 311–335 (1971).
- [59] F. VERSTRAETE and H. VERSCHELDE, arXiv:0202124 (2002).
- [60] M. S. BYRD and N. KHANEJA, *Phys. Rev. A* **68**, 062322 (2003).
- [61] E. B. DAVIES, *Quantum theory of open systems*, Academic Press, London (1976).
- [62] G. LINDBLAD, *Comm. Math. Phys.* **65**, 281–294 (1979).
- [63] P. PERINOTTI, *Causal Structures and the Classification of Higher Order Quantum Computations*, In: R. RENNER, and S. STUPAR (eds.), *Time in Physics*, Springer International Publishing (2017).
- [64] A. BISIO and P. PERINOTTI, *Proc. Royal Soc. A* **475**, 20180706 (2019).
- [65] H.-P. BREUER, E.-M. LAINE, J. PIILO, and B. VACCHINI, *Rev. Mod. Phys.* **88**, 021002 2 (2016).
- [66] L. LI, M. J. W. HALL, and H. M. WISEMAN, *Phys. Rep.* **759**, 1–51 (2018).
- [67] K. MODI, *Open Syst. Inf. Dyn.* **18**, 253–260 (2011).
- [68] K. MODI, C. A. RODRÍGUEZ-ROSARIO, and A. ASPURU-GUZIŁ, *Phys. Rev. A* **86**, 064102 (2012).
- [69] A. N. KOLMOGOROV, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin (1933) [*Foundations of the Theory of Probability*, Chelsea, New York, (1956)].
- [70] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Wiley, New York (1971).
- [71] H.-P. BREUER and F. PETRUCCIONE, *The Theory of Open Quantum Systems*, Oxford University Press (2007).
- [72] T. TAO, *An introduction to measure theory*, American Mathematical Society (2011).
- [73] A. J. LEGGETT and A. GARG, *Phys. Rev. Lett.* **54**, 857–860 (1985).
- [74] A. J. LEGGETT, *Rep. Prog. Phys.* **71**, 022001 (2008).
- [75] C. EMARY, N. LAMBERT, and F. NORI, *Rep. Prog. Phys.* **77**, 016001 (2014).
- [76] O. J. E. MARONEY, *Stud. Hist. Philos. Sci. B* **58**, 41–53 (2017).
- [77] N. GILBERT, *Agent-Based Models*, SAGE Publications Inc., Los Angeles (2007).

- [78] G. PARRAVICINI and A. ZECCA, *Rep. Math. Phys.* **12**, 423–424 (1977).
- [79] I. L. CHUANG and M. A. NIELSEN, *J. Mod. Opt.* **44**, 2455–2467 (1997).
- [80] A. SMIRNE, D. EGLOFF, M. G. DÍAZ, M. B. PLENIO, and S. F. HUELGA, *Quantum Sci. Technol.* **4**, 01LT01 (2019).
- [81] Z. SCHUSS, *Theory and Applications of Stochastic Processes: An Analytical Approach*, Springer, New York (2009).
- [82] M. LIAO, *Applied Stochastic Processes*, Chapman and Hall/CRC, Boca Raton (2013).
- [83] J. M. JAUCH and C. PIRON, *Helv. Phys. Acta* **40**, 559–570 (1967).
- [84] A. S. HOLEVO, *Probabilistic and statistical aspects of quantum mechanics*, North Holland, Amsterdam (1982).
- [85] R. TUMULKA, *Lett. Math. Phys.* **84**, 41–46 (2008).
- [86] L. ACCARDI, A. FRIGERIO, and J. T. LEWIS, *Publ. Rest. Inst. Math. Sci.* **18**, 97–133 (1982).
- [87] S. SHRAPNEL, F. COSTA, and G. MILBURN, *New J. Phys.* **20**, 053010 (2018).
- [88] M. P. LÉVY, *Am. J. Math.* **62**, 487–550 (1940).
- [89] N. WIENER, A. SIEGEL, B. RANKIN, and W. T. MARTIN (Eds.), *Differential Space, Quantum Systems, and Prediction*, The MIT Press, Cambridge (MA) (1966).
- [90] Z. CIESIELSKI, *Lectures on Brownian motion, heat conduction and potential theory*, Aarhus Universitet, Matematisk Institut (1966).
- [91] R. BHATTACHARYA and E. C. WAYMIRE, *A Basic Course in Probability Theory*, Springer, New York, NY (2017).
- [92] L. ACCARDI, *Adv. Math.* **20**, 329–366 (1976).
- [93] J. F. POYATOS, J. I. CIRAC, and P. ZOLLER, *Phys. Rev. Lett.* **78**, 390–393 (1997).
- [94] W. F. STINESPRING, *Proc. Amer. Math. Soc.* **6**, 211–216 2 (1955).
- [95] M. ARAÚJO, C. BRANCIARD, F. COSTA, A. FEIX, C. GIARMATZI, and Č. BRUKNER, *New J. Phys.* **17**, 102001 (2015).
- [96] R. W. SPEKKENS, *Phys. Rev. A* **71**, 052108 (2005).
- [97] E. G. CAVALCANTI, *Phys. Rev. X* **8**, 021018 (2018).
- [98] F. CARUSO, V. GIOVANNETTI, C. LUPO, and S. MANCINI, *Rev. Mod. Phys.* **86**, 1203–1259 (2014).
- [99] C. PORTMANN, C. MATT, U. MAURER, R. RENNER, and B. TACKMANN, *IEEE Trans. Inf. Theory* **63**, 3277–3305 (2017).
- [100] L. HARDY, arXiv:1608.06940 (2016).
- [101] L. HARDY, *Phil. Trans. R. Soc. A* **370**, 3385–3417 (2012).
- [102] J. COTLER, C.-M. JIAN, X.-L. QI, and F. WILCZEK, *J. High Energy Phys.* **2018**, 93 (2018).
- [103] N. BARNETT and J. P. CRUTCHFIELD, *J. Stat. Phys.* **161**, 404–451 (2015).
- [104] J. THOMPSON, A. J. P. GARNER, V. VEDRAL, and M. GU, *npj Quantum Inf.* **3**, 6 (2017).

- [105] F. A. POLLOCK, C. RODRÍGUEZ-ROSARIO, T. FRAUENHEIM, M. PATERNOSTRO, and K. MODI, *Phys. Rev. Lett.* **120**, 040405 (2018).
- [106] E. HAAPASALO, T. HEINOSAARI, and Y. KURAMOCHI, *J. Phys. A* **49**, 33LT01 (2016).
- [107] N. VAN KAMPEN, *Braz. J. Phys.* **28**, (1998).
- [108] V. GIOVANNETTI, S. MONTANGERO, and R. FAZIO, *Phys. Rev. Lett.* **101**, 180503 (2008).
- [109] R. N. C. PFEIFER, G. EVENBLY, and G. VIDAL, *Phys. Rev. A* **79**, 040301 (2009).
- [110] S. SINGH, *Phys. Rev. D* **97**, 026012 (2018).
- [111] R. SILVA, Y. GURYANOVA, A. J. SHORT, P. SKRZYPCZYK, N. BRUNNER, and S. POPESCU, *New J. Phys.* **19**, 103022 (2017).
- [112] K. MODI and E. C. G. SUDARSHAN, *Phys. Rev. A* **81**, 052119 (2010).
- [113] S. VINJANAMPATHY and K. MODI, *Phys. Rev. A* **92**, 052310 5 (2015).
- [114] R. ALICKI and K. LENDI, *Quantum dynamical semi-groups and applications*, Springer, Berlin (1987).
- [115] V. A. FRANKE, *Theor. Math. Phys.* **27**, 406–413 (1976).
- [116] V. GORINI, A. KOSSAKOWSKI, and E. C. G. SUDARSHAN, *J. Math. Phys.* **17**, 821–825 (1976).
- [117] S. NAKAJIMA, *Prog. Theor. Phys.* **20**, 948–959 (1958).
- [118] R. ZWANZIG, *J. Chem. Phys.* **33**, 1338–1341 (1960).
- [119] S. H. MYRSKOG, J. K. FOX, M. W. MITCHELL, and A. M. STEINBERG, *Phys. Rev. A* **72**, 013615 (2005).
- [120] M. HOWARD, J. TWAMLEY, C. WITTMANN, T. GAEBEL, F. JELEZKO, and J. WRACHTRUP, *New J. Phys.* **8**, 33 (2006).
- [121] J. M. CHOW, J. M. GAMBETTA, L. TORNBERG, J. KOCH, L. S. BISHOP, A. A. HOUCK, B. R. JOHNSON, L. FRUNZIO, S. M. GIRVIN, and R. J. SCHOELKOPF, *Phys. Rev. Lett.* **102**, 090502 (2009).
- [122] K. MODI, *A theoretical analysis of experimental open quantum dynamics*, PhD Thesis, University of Texas, Austin (2009), arXiv:0903.2724.
- [123] R. ALICKI, *Phys. Rev. Lett.* **75**, 3020–3020 (1995).
- [124] T. F. JORDAN, A. SHAJI, and E. C. G. SUDARSHAN, *Phys. Rev. A* **70**, 052110 (2004).
- [125] C. A. RODRÍGUEZ-ROSARIO, K. MODI, and A. ASPURU-GUZIĆ, *Phys. Rev. A* **81**, 012313 (2010).
- [126] P. ŠTELMACHOVIČ and V. BUŽEK, *Phys. Rev. A* **64**, 062106 (2001).
- [127] C. ARENZ, R. HILLIER, M. FRAAS, and D. BURGARTH, *Phys. Rev. A* **92**, 022102 (2015).
- [128] G. A. PAZ-SILVA, M. J. W. HALL, and H. M. WISEMAN, arXiv:1810.12540 (2018).
- [129] M. M. WOLF, J. EISERT, T. S. CUBITT, and J. I. CIRAC, *Phys. Rev. Lett.* **101**, 150402 15 (2008).

- [130] H. BREUER, E.-M. LAINE, and J. PIILO, *Phys. Rev. Lett.* **103**, 210401 21 (2009).
- [131] Á. RIVAS, S. F. HUELGA, and M. B. PLENIO, *Phys. Rev. Lett.* **105**, 050403 (2010).
- [132] X. LU, X. WANG, and C. P. SUN, *Phys. Rev. A* **82**, 042103 4 (2010).
- [133] S. C. HOU, X. X. YI, S. X. YU, and C. H. OH, *Phys. Rev. A* **83**, 062115 6 (2011).
- [134] A. R. USHA DEVI, A. K. RAJAGOPAL, and SUDHA, *Phys. Rev. A* **83**, 022109 2 (2011).
- [135] A. R. USHA DEVI, A. K. RAJAGOPAL, S. SHENOY, and R. W. RENDELL, *J. Quant. Info. Sci.* **2**, 47–54 (2012).
- [136] L. MAZZOLA, C. A. RODRÍGUEZ-ROSARIO, K. MODI, and M. PATERNOSTRO, *Phys. Rev. A* **86**, 010102 (2012).
- [137] C. A. RODRÍGUEZ-ROSARIO, K. MODI, L. MAZZOLA, and A. ASPURU-GUZIŁ, *EPL* **99**, 20010 (2012).
- [138] S. LUO, S. FU, and H. SONG, *Phys. Rev. A* **86**, 044101 4 (2012).
- [139] B. BYLICKA, D. CHRUSCIŃSKI, and S. MANISCALCO, *Sci. Rep.* **4**, 5720 (2014).
- [140] D. CHRUSCIŃSKI and S. MANISCALCO, *Phys. Rev. Lett.* **112**, 120404 12 (2014).
- [141] B. BYLICKA, M. JOHANSSON, and A. ACÍN, *Phys. Rev. Lett.* **118**, 120501 12 (2017).
- [142] Z. HE, H. ZENG, Y. LI, Q. WANG, and C. YAO, *Phys. Rev. A* **96**, 022106 2 (2017).
- [143] D. CHRUSCIŃSKI, Á. RIVAS, and E. STØRMER, *Phys. Rev. Lett.* **121**, 080407 8 (2018).
- [144] A.-M. KUAH, K. MODI, C. A. RODRÍGUEZ-ROSARIO, and E. C. G. SUDARSHAN, *Phys. Rev. A* **76**, 042113 (2007).
- [145] M. RINGBAUER, C. J. WOOD, K. MODI, A. GILCHRIST, A. G. WHITE, and A. FEDRIZZI, *Phys. Rev. Lett.* **114**, 090402 (2015).
- [146] D. SCHMID, K. RIED, and R. W. SPEKKENS, *arXiv:1806.02381* (2018).
- [147] C. GIARMATZI and F. COSTA, *arXiv:1811.03722* (2018).
- [148] S. SHRAPNEL, F. COSTA, and G. MILBURN, *Int. J. Quantum Inf.* **16**, 1840010 (2018).
- [149] C. ALTAFINI and F. TICOZZI, *IEEE Trans. Autom. Control* **57**, 1898–1917 (2012).
- [150] M. SHAPIRO and P. BRUMER, *Principles of the Quantum Control of Molecular Processes*, John Wiley & Sons (2003).
- [151] M. S. BLOK, C. BONATO, M. L. MARKHAM, D. J. TWITCHEN, V. V. DOBROVITSKI, and R. HANSON, *Nat. Phys.* **10**, 189–193 (2014).
- [152] L. VIOLA, E. KNILL, and S. LLOYD, *Phys. Rev. Lett.* **82**, 2417–2421 (1999).
- [153] L. HARDY, *arXiv:0101012* (2001).
- [154] B. DAKIĆ and Č. BRUKNER, *arXiv:0911.0695* (2009).
- [155] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *Phys. Rev. A* **81**, 062348 (2010).
- [156] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *Phys. Rev. A* **84**, 012311 (2011).
- [157] L. MASANES and M. P. MÜLLER, *New J. Phys.* **13**, 063001 (2011).

- [158] M. ARAÚJO, A. FEIX, M. NAVASCUÉS, and Č. BRUKNER, *Quantum* **1**, 10 (2017).
- [159] M. NEUMARK, *Izv. Akad. Nauk SSSR Ser. Mat.* **4**, 53–104 (1940).
- [160] A. PERES, *Quantum Theory: Concepts and Methods*, Springer Science & Business Media (1995).
- [161] V. PAULSEN, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press (2003).
- [162] I. KULL, P. A. GUÉRIN, and Č. BRUKNER, *arXiv:1807.09187* (2018).
- [163] G. THIJS, M. LESCOT, K. MARCHAL, S. ROMBAUTS, B. DE MOOR, P. ROUZÉ, and Y. MOREAU, *Bioinformatics* **17**, 1113–1122 (2001).
- [164] S. L. SALZBERG, A. L. DELCHER, S. KASIF, and O. WHITE, *Nucleic Acids Res.* **26**, 544–548 (1998).
- [165] M. ROSVALL, A. V. ESQUIVEL, A. LANCICHINETTI, J. D. WEST, and R. LAMBIOTTE, *Nat. Commun.* **5**, 4630 (2014).
- [166] H.-P. BREUER, E.-M. LAINE, and J. PILO, *Phys. Rev. Lett.* **103**, 210401 (2009).
- [167] S. C. HOU, X. X. YI, S. X. YU, and C. H. OH, *Phys. Rev. A* **83**, 062115 (2011).
- [168] E.-M. LAINE, J. PILO, and H.-P. BREUER, *EPL* **92**, 60010 (2010).
- [169] M. GESSNER and H.-P. BREUER, *Phys. Rev. Lett.* **107**, 180402 (2011).
- [170] Á. RIVAS, S. F. HUELGA, and M. B. PLENIO, *Rep. Prog. Phys.* **77**, 094001 (2014).
- [171] J. RAU, *Phys. Rev.* **129**, 1880–1888 (1963).
- [172] M. ZIMAN, P. ŠTELMACHOVIČ, V. BUŽEK, M. HILLERY, V. SCARANI, and N. GISIN, *Phys. Rev. A* **65**, 042105 4 (2002).
- [173] T. A. BRUN, *Am. J. Phys.* **70**, 719–737 (2002).
- [174] V. SCARANI, M. ZIMAN, P. ŠTELMACHOVIČ, N. GISIN, and V. BUŽEK, *Phys. Rev. Lett.* **88**, 097905 9 (2002).
- [175] T. RYBÁR, S. N. FILIPPOV, M. ZIMAN, and V. BUŽEK, *J. Phys. B* **45**, 154006 (2012).
- [176] J. PRESKILL, *Quantum* **2**, 79 (2018).
- [177] M. M. WOLF and J. I. CIRAC, *Commun. Math. Phys.* **279**, 147–168 (2008).
- [178] G. LINDBLAD, *Comm. Math. Phys.* **48**, 119–130 (1976).
- [179] B. VACCHINI, A. SMIRNE, E.-M. LAINE, J. PILO, and H.-P. BREUER, *New J. Phys.* **13**, 093004 (2011).
- [180] J. BAE and D. CHRUSCIŃSKI, *Phys. Rev. Lett.* **117**, 050403 5 (2016).
- [181] S. WISSMANN, H.-P. BREUER, and B. VACCHINI, *Phys. Rev. A* **92**, 042108 (2015).
- [182] M. PIANI, M. HORODECKI, P. HORODECKI, and R. HORODECKI, *Phys. Rev. A* **74**, 012305 (2006).
- [183] G. CHIRIBELLA, *Phys. Rev. A* **86**, 040301 (2012).
- [184] J. KOFLER and Č. BRUKNER, *Phys. Rev. A* **87**, 052115 (2013).

- [185] M. ZIMAN, P. ŠTELMACHOVIV C, and V. BUŽEK, *Open Syst. Inf. Dyn.* **12**, 81–91 (2005).
- [186] N. K. BERNARDES, A. R. R. CARVALHO, C. H. MONKEN, and M. F. SANTOS, *Phys. Rev. A* **90**, 032111 3 (2014).
- [187] P. TARANTO, *Memory Effects in Quantum Processes*, Masters Thesis, Monash University (2019).
- [188] D. GROSS, Y.-K. LIU, S. T. FLAMMIA, S. BECKER, and J. EISERT, *Phys. Rev. Lett.* **105**, 150401 (2010).
- [189] M. CRAMER, M. B. PLENIO, S. T. FLAMMIA, R. SOMMA, D. GROSS, S. D. BARTLETT, O. LANDON-CARDINAL, D. POULIN, and Y.-K. LIU, *Nat. Commun.* **1**, 149 (2010).
- [190] M. RECK, A. ZEILINGER, H. J. BERNSTEIN, and P. BERTANI, *Phys. Rev. Lett.* **73**, 58–61 (1994).
- [191] THE IBM QUANTUM EXPERIENCE, <http://www.research.ibm.com/quantum>.
- [192] F. VATAN, V. P. ROYCHOWDHURY, and M. P. ANANTRAM, *IEEE Trans. Inf. Theory* **45**, 1703–1708 (1999).
- [193] O. ORESHKOV and T. A. BRUN, *Phys. Rev. A* **76**, 022318 (2007).
- [194] G. CHIRIBELLA, M. DALL’ARNO, G. M. D’ARIANO, C. MACCHIAVELLO, and P. PERINOTTI, *Phys. Rev. A* **83**, 052305 (2011).
- [195] M. ZIMAN, M. PLESCH, and V. BUŽEK, *Found. Phys.* **36**, 127–156 (2006).
- [196] M. ZIMAN, M. PLESCH, and V. BUŽEK, *Eur. Phys. J. D* **32**, 215–222 (2005).
- [197] E. POON, *Linear Algebra Appl.* **468**, 122–144 (2015).
- [198] C. B. MENDL and M. M. WOLF, *Commun. Math. Phys.* **289**, 1057–1086 (2009).
- [199] M. GESSNER and H.-P. BREUER, *Phys. Rev. A* **87**, 042107 (2013).
- [200] M. GESSNER, M. RAMM, T. PRUTTIVARASIN, A. BUCHLEITNER, H.-P. BREUER, and H. HÄFFNER, *Nat. Phys.* **10**, 105–109 (2014).
- [201] T. YU and J. H. EBERLY, *Quantum Info. Comput.* **7**, 459–468 (2007).
- [202] M. ALI, A. R. P. RAU, and G. ALBER, *Phys. Rev. A* **81**, 042105 (2010).
- [203] L. VIOLA and S. LLOYD, *Phys. Rev. A* **58**, 2733–2744 (1998).
- [204] J. ÅBERG, *Phys. Rev. Lett.* **113**, 150402 (2014).
- [205] A. STRELTSOV, G. ADESSO, and M. B. PLENIO, *Rev. Mod. Phys.* **89**, 041003 (2017).
- [206] A. SMIRNE, Private Communication.
- [207] H. OLLIVIER and W. H. ZUREK, *Phys. Rev. Lett.* **88**, 017901 (2001).
- [208] L. HENDERSON and V. VEDRAL, *J. Phys. A* **34**, 6899 (2001).
- [209] A. DATTA, [arXiv:1003.5256](https://arxiv.org/abs/1003.5256) (2010).
- [210] K. MODI, A. BRODUTCH, H. CABLE, T. PATEREK, and V. VEDRAL, *Rev. Mod. Phys.* **84**, 1655–1707 (2012).

- [211] M. G. DÍAZ, K. FANG, X. WANG, M. ROSATI, M. SKOTINIOTIS, J. CALSAMIGLIA, and A. WINTER, *Quantum* **2**, 100 (2018).
- [212] T. BAUMGRATZ, M. CRAMER, and M. B. PLENIO, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [213] A. WINTER and D. YANG, *Phys. Rev. Lett.* **116**, 120404 (2016).
- [214] B. YADIN, J. MA, D. GIROLAMI, M. GU, and V. VEDRAL, *Phys. Rev. X* **6**, 041028 (2016).
- [215] L. M. PROCOPIO, A. MOQANAKI, M. ARAÚJO, F. COSTA, I. A. CALAFELL, E. G. DOWD, D. R. HAMEL, L. A. ROZEMA, Č. BRUKNER, and P. WALTHER, *Nat. Commun.* **6**, 8913 (2015).
- [216] G. RUBINO, L. A. ROZEMA, A. FEIX, M. ARAÚJO, J. M. ZEUNER, L. M. PROCOPIO, Č. BRUKNER, and P. WALTHER, *Sci. Adv.* **3**, e1602589 (2017).
- [217] K. GOSWAMI, C. GIARMATZI, M. KEWMING, F. COSTA, C. BRANCIARD, J. ROMERO, and A. G. WHITE, *Phys. Rev. Lett.* **121**, 090503 (2018).
- [218] L. HARDY, *arXiv:0509120* (2005).
- [219] L. HARDY, *J. Phys. A* **40**, 3081–3099 (2007).
- [220] Č. BRUKNER, *New J. Phys.* **17**, 083034 (2015).
- [221] C. BRANCIARD, M. ARAÚJO, A. FEIX, F. COSTA, and Č. BRUKNER, *New J. Phys.* **18**, 013008 (2016).
- [222] A. A. ABBOTT, J. WECHS, D. HORSMAN, M. MHALLA, and C. BRANCIARD, *arXiv:1810.09826* (2018).
- [223] P. A. GUÉRIN, G. RUBINO, and Č. BRUKNER, *Phys. Rev. A* **99**, 062317 (2019).
- [224] T. MORIMAE, *Phys. Rev. A* **90**, 010101 1 (2014).
- [225] M. ARAÚJO, C. BRANCIARD, F. COSTA, A. FEIX, C. GIARMATZI, and Č. BRUKNER, *New J. Phys.* **17**, 102001 (2015).
- [226] S. SALEK, D. EBLER, and G. CHIRIBELLA, *arXiv:1809.06655* (2018).
- [227] G. CHIRIBELLA and H. KRISTJÁNSSON, *Proc. Royal Soc. A* **475**, 20180903 (2019).
- [228] R. F. WERNER, *Phys. Rev. A* **40**, 4277–4281 (1989).
- [229] C. T. M. HO, F. COSTA, C. GIARMATZI, and T. C. RALPH, *arXiv:1804.05498* (2018).
- [230] C. H. BENNETT, *Simulated Time Travel, Teleportation Without Communication, and How to Conduct a Romance with Someone who has Fallen into a Black Hole* (QUPON, 2005).
- [231] S. LLOYD, L. MACCONE, R. GARCIA-PATRON, V. GIOVANNETTI, and Y. SHIKANO, *Phys. Rev. D* **84**, (2011).
- [232] S. LLOYD, L. MACCONE, R. GARCIA-PATRON, *et al.*, *Phys. Rev. Lett.* **106**, (2011).
- [233] D. GENKINA, G. CHIRIBELLA, and L. HARDY, *Phys. Rev. A* **85**, 022330 2 (2012).
- [234] A. A. ABBOTT, C. GIARMATZI, F. COSTA, and C. BRANCIARD, *Phys. Rev. A* **94**, 032131 (2016).
- [235] J. WECHS, A. A. ABBOTT, and C. BRANCIARD, *New J. Phys.* (2018).

- [236] H. BARNUM, C. A. FUCHS, J. M. RENES, and A. WILCE, arXiv:0507108 (2005).
- [237] G. GUTOSKI, *Quantum Inf. Comput.* **9**, 739–764 (2009).
- [238] G. M. D’ARIANO, *Philos. Trans. Royal Soc. A* **376**, 20170313 (2018).
- [239] J. BARRETT, *Phys. Rev. A* **65**, 042302 (2002).
- [240] A. FEIX, M. ARAÚJO, and Č. BRUKNER, *New J. Phys.* **18**, 083040 (2016).
- [241] I. E. NESTEROV, A. NEMIROVSKII, and Y. NESTEROV, *Interior-Point Polynomial Algorithms in Convex Programming*, Society for Industrial & Applied Mathematics, Philadelphia (1987).
- [242] J. WATROUS, *Lecture notes in Quantum Information Theory* (2011).
- [243] M. GRANT and S. BOYD, *CVX: Matlab Software for Disciplined Convex Programming*, version 2.1 (2014).
- [244] M. GRANT and S. BOYD, *Graph implementations for nonsmooth convex programs*, In: V. BLONDEL, S. BOYD, and H. KIMURA (eds.), *Recent Advances in Learning and Control*, Springer London (2008).
- [245] C. H. BENNETT, A. W. HARROW, D. W. LEUNG, and J. A. SMOLIN, *IEEE Trans. Inf. Theory* **49**, 1895–1911 (2003).
- [246] F. COSTA, M. RINGBAUER, M. E. GOGGIN, A. G. WHITE, and A. FEDRIZZI, *Phys. Rev. A* **98**, 012328 (2018).
- [247] D. JURKSCHAT, *Causality in Quantum Processes*, Honours Thesis, Monash University (2018).
- [248] Ä. BAUMELER, A. FEIX, and S. WOLF, *Phys. Rev. A* **90**, 042106 (2014).
- [249] M. ARAÚJO, P. A. GUÉRIN, and Ä. BAUMELER, *Phys. Rev. A* **96**, 052315 (2017).
- [250] T. BAUMGRATZ, D. GROSS, M. CRAMER, and M. B. PLENIO, *Phys. Rev. Lett.* **111**, 020401 (2013).
- [251] T. J. OSBORNE, arXiv:1901.06124 (2019).
- [252] J. v. NEUMANN, *Comp. Math.* **6**, 1–77 (1939).
- [253] M. REED and B. SIMON, *Functional Analysis*, Academic Press, New York (1980).
- [254] R. HAAG and D. KASTLER, *J. Math. Phys.* **5**, 848–861 (1964).
- [255] R. F. WERNER, *J. Phys. A* **34**, 7081 (2001).
- [256] J. OPPENHEIM and B. REZNIK, *Phys. Rev. A* **70**, 022312 (2004).