

MEMORY EFFECTS IN QUANTUM PROCESSES

PHILIP TARANTO

DR. KAVAN MODI & DR. FELIX A. POLLOCK



MONASH
University

Master of Science (Research)
School of Physics and Astronomy
Monash University

May 6, 2019

To my family,
for their unwavering support.

ABSTRACT

Understanding temporal processes and their correlations in time is of paramount importance for the development of near-term technologies that operate under realistic conditions. Capturing the complete multi-time statistics defining a stochastic process lies at the heart of any proper treatment of memory effects. This is well-understood in classical theory, where a hierarchy of joint probability distributions completely characterises the process at hand. However, attempting to generalise this notion to quantum mechanics is problematic: observing realisations of a quantum process necessarily disturbs it, breaking an implicit, and crucial, assumption in the classical setting. This issue can be overcome by separating the experimental interventions from the underlying process, enabling an unambiguous description of the process itself and accounting for all possible multi-time correlations for any choice of interrogating instruments.

In this thesis, using a novel framework for the characterisation of quantum stochastic processes, we first solve the long standing question of unambiguously describing the *memory length* of a quantum processes. This is achieved by constructing a *quantum Markov order* condition that naturally generalises its classical counterpart for the quantification of finite-length memory effects. As measurements are inherently invasive in quantum mechanics, one has no choice but to define Markov order with respect to the interrogating instruments that are used to probe the process at hand: different memory effects are exhibited depending on how one addresses the system, in contrast to the standard classical setting. We then fully characterise the structural constraints imposed on quantum processes with finite Markov order, shedding light on a variety of memory effects that can arise through various examples. Lastly, we introduce an instrument-specific notion of *memory strength* that allows for a meaningful quantification of the temporal correlations between the history and the future of a process for a given choice of experimental intervention.

These findings are directly relevant to both characterising and exploiting memory effects that persist for a finite duration. In particular, immediate applications range from developing efficient compression and recovery schemes for the description of quantum processes with memory to designing coherent control protocols that efficiently perform information-theoretic tasks, amongst a plethora of others.

PUBLICATIONS

Many of the ideas and figures contained in this thesis have appeared in the following publications: [1, 2].

- [1] P. TARANTO, F. A. POLLOCK, S. MILZ, M. TOMAMICHEL, and K. MODI, *Phys. Rev. Lett.* **122**, 140401 (2019).
- [2] P. TARANTO, S. MILZ, F. A. POLLOCK, and K. MODI, *Phys. Rev. A* **99**, 042108 (2019).

Memory is not an instrument for surveying the past but its theatre
— Walter Benjamin.

ACKNOWLEDGEMENTS

First of all, I must thank my supervisors, Kavan Modi and Felix Pollock, for their support and guidance throughout my journey. Both of them have provided a perfectly balanced atmosphere of ongoing encouragement, diligent rigour and healthy scepticism that has undoubtedly helped me develop as a scientist and as a person. Kavan’s quick-witted counterexamples to my many outlandish claims saved me from pursuing countless dead-ends and taught me that often *the problem is your friend*. Felix’s attention to details and grammatical pedantry immensely improved my abilities in scientific communication and cured my unhealthy obsession with semi-colons.

Secondly, my deepest thanks goes out to my de facto supervisor, Simon Milz. It has been a pleasure to have learnt many fun facts about an undisclosed European nation thanks to you... oh, and a thing or two about physics. It’s been a wild ride so far and now we set sail for tomorrow!

I would like to express my gratitude to the School of Physics and Astronomy at Monash University. I couldn’t have asked for a more supportive environment in which to study both my undergraduate and masters degrees. In addition, I appreciate the financial support of the Australian Government Research Training Program (RTP) and the J. L. William Scholarships.

I especially thank the entire MONQIS group for their friendship, in particular Simon Milz, Francesco Campaioli, Josh Morris, Top Notoh and the honorary member, Cody Duncan. I am yet again thankful to Simon Milz for his thorough revisions of this thesis, Lee Miles for providing inspiration in times of need, and Ashley Bransgrove, Vanessa La Delfa and John Farrugia for their thoughtful comments on an early draft.

Lastly, thanks to my family and friends for their support throughout this journey.

Melbourne, 13 March, 2019.

Contents

CONTENTS [xiii](#)

LIST OF FIGURES [xv](#)

LIST OF TABLES [xv](#)

ACRONYMS [xvi](#)

I A KIND OF GLORY

- 1 INTRODUCTION [3](#)
 - 1.1 Reading Guide [7](#)

II NARRATIVE ON THE EDGE

- 2 CLASSICAL AND QUANTUM DYNAMICS WITH NOISE [13](#)
 - 2.1 Classical Stochastic Processes [15](#)
 - 2.1.1 Probability Spaces [15](#)
 - 2.1.2 Stochastic Processes [18](#)
 - 2.1.3 Modelling Stochastic Processes [22](#)
 - 2.1.4 Markovian Stochastic Processes [24](#)
 - 2.1.5 Non-Markovian Stochastic Processes [27](#)
 - 2.1.6 Classical Markov Order [29](#)
 - 2.2 Open Quantum Dynamics [33](#)
 - 2.2.1 Open Quantum Systems [35](#)
 - 2.2.2 Dynamical Maps [39](#)
 - 2.2.3 GKSL Equation [43](#)
 - 2.2.4 Tomographic Reconstruction of Quantum Channels [44](#)
 - 2.3 A Problem of Formalism [45](#)
 - 2.3.1 A Hierarchy of Notions of Non-Markovianity [48](#)
 - 2.3.2 Limitations of Traditional Approaches [49](#)
- 3 QUANTUM STOCHASTIC PROCESSES [53](#)
 - 3.1 Process Tensor Framework [54](#)
 - 3.1.1 Multi-time Quantum Experiments [55](#)
 - 3.1.2 Process Tensor [59](#)

CONTENTS

3.2	Representing Linear Maps	65
3.2.1	Choi-Jamiołkowski Isomorphism	65
3.2.2	Choi Representation of Process Tensor	68
3.2.3	Properties of the Process Tensor	70
3.3	Spatio-temporal Born Rule	72
3.4	Markovian Quantum Processes	77
 III UNSPOKEN WORDS		
4	MEMORY LENGTH	87
4.1	Quantum Markov Order	89
4.1.1	Instrument-specific Quantum Markov Order	90
4.1.2	Finite Memory Constraint on the Process Tensor	92
4.2	Relation to Classical Markov Order	97
4.2.1	Classical Stochastic Processes with Interventions	97
4.3	Memory Length of a Generalised Collision Model	100
4.4	Chapter Summary	106
5	PROCESSES WITH FINITE MEMORY LENGTH	109
5.1	Structure of Quantum Processes with Finite Markov Order	110
5.1.1	Unitary History-Blocking Instrument Sequences	112
5.1.2	Informationally-Complete History-Blocking Sequences	114
5.2	Quantum Markov Order and Conditional Mutual Information	118
5.3	Chapter Summary	122
6	MEMORY STRENGTH	125
6.1	Quantifying Memory Strength	126
6.1.1	Instrument-specific Memory Strength	127
6.1.2	Memory Length for an Exactly Solvable Model	130
6.2	Chapter Summary	135
 IV ENVOI		
7	SUMMARY	139
7.1	Outlook	142
 V APPENDIX		
A	NOTATION SUMMARY	145
B	CLASSICAL AND QUANTUM DYNAMICS WITH NOISE	149
B.1	Classical Master Equation	149

B.2	Tomography of a Dynamical Map	150
C	MEMORY LENGTH	153
C.1	Quantum Markov Order Constraint on Process Tensor	153
C.2	Demanding Finite Quantum Markov Order for all Instruments Implies Markovianity	156
C.3	Classical Markov Order with Fuzzy Measurements	157
C.4	Memory Length of a Generalised Collision Model with Memory via Repeated System-Ancilla Interactions	159
C.5	Other Generalised Collision Models with Memory	161
D	PROCESSES WITH FINITE MEMORY LENGTH	165
D.1	Process with Finite Markov Order and Non-zero Quantum CMI	165
D.2	Finite Markov Order does not Imply Vanishing Quantum CMI	167
D.3	Fuzzy Orthogonal Projective Measurements on a Quantum Process	170
	BIBLIOGRAPHY	173

List of Figures

Figure 2.1	Continuous and discrete-time stochastic processes	19
Figure 2.2	Containment property for classical stochastic processes	21
Figure 2.3	Perturbed coin	25
Figure 2.4	Random walk conditioned on marbles drawn	28
Figure 2.5	Markov order as the natural notion of memory length	31
Figure 2.6	Breakdown of KET in a Stern-Gerlach experiment	34
Figure 2.7	Dynamical map	39
Figure 2.8	Complete-positivity	41
Figure 2.9	Stinespring dilation of a quantum channel	42
Figure 2.10	Tomographic reconstruction of a quantum channel	44
Figure 2.11	Dilation of a quantum process interrogated in time	46
Figure 2.12	Measuring the system conditions the environment and influences future dynamics	47
Figure 2.13	Initial correlation problem	51
Figure 3.1	Multi-time quantum experiment	56
Figure 3.2	Superchannel: resolution to the initial correlation problem	60
Figure 3.3	Process tensor: an operational description of quantum stochastic processes	61
Figure 3.4	Doing nothing vs. averaging over measurements	62
Figure 3.5	Consistency condition for the process tensor	63
Figure 3.6	CJI of a linear map	66
Figure 3.7	The unique trace-preserving effect	68
Figure 3.8	CJI for the process tensor	69
Figure 3.9	Complete-positivity for the superchannel	71
Figure 3.10	Graphical representation of a POVM	72
Figure 3.11	Graphical representation of an instrument	74
Figure 3.12	Graphical representation of a tester	76
Figure 3.13	Causal break	79

Figure 3.14	Markovian process tensor	81
Figure 4.1	Instrument-specific quantum Markov order	93
Figure 4.2	Generalised collision model with memory	101
Figure 4.3	Finite memory with respect to trash-and-prepare protocol	103
Figure 5.1	Finite memory for a unitary instrument sequence	113
Figure 5.2	Finite memory for an informationally-complete sequence	116
Figure 5.3	Process with non-vanishing quantum CMI	120
Figure 6.1	Instrument-specific memory strength	130
Figure 6.2	Abrupt transition between CP-divisible and non-divisible dynamics	132
Figure 6.3	Heatmap of non-Markovianity	133
Figure 6.4	Instrument-specific memory strength	134
Figure C.1	Instrument-dependence of classical Markov order with fuzzy measurements	158
Figure C.2	Generalised collision models with memory	162
Figure D.1	Process with finite quantum Markov order but non-vanishing quantum CMI	166
Figure D.2	Process with non-vanishing quantum CMI but finite Markov order with respect to (fuzzy) orthogonal projectors	171

List of Tables

Table A.1	Notational conventions	148
-----------	------------------------	-----

ACRONYMS

ME	Master equation
P	Positive
CP	Completely-positive
TP	Trace-preserving
CJI	Choi-Jamiołkowski isomorphism
CPTP	Completely-positive and trace-preserving
KET	Kolmogorov extension theorem
IC	Informationally-complete
POVM	Positive operator-valued measure
CMI	Conditional mutual information
GKSL	Gorini, Kossakowski, Sudarshan and Lindblad

Part I

A KIND OF GLORY

Sometimes a kind of glory lights up the mind of a man. It happens to nearly everyone. You can feel it growing or preparing like a fuse burning toward dynamite. It is a feeling in the stomach, a delight of the nerves, of the forearms. The skin tastes the air, and every deep-drawn breath is sweet.

— John Steinbeck, *East of Eden*.

INTRODUCTION

ALTHOUGH PHYSICAL LAWS ARE FUNDAMENTALLY LOCAL IN TIME, memory effects are ubiquitous in processes observed in nature [3–5]. We see such effects when we try to predict the weather or the stock market, describe transport processes at the microscopic level or understand the random motion of particles suspended in a fluid, to name but a few examples. Memory arises because, in reality, no system is truly isolated; our inability to capture interactions between the system of interest and its environment leads to dynamics that can exhibit complex temporal correlations. Through these interactions, information about the system’s past can be stored in the environment, which carries it forward to dictate the future evolution of the system.

In classical physics, should an experimenter be equipped with sufficient resources to track the evolution of *all* relevant degrees of freedom—including those of the environment—a deterministic, memoryless description of the evolution could, in principle, be derived. In practice, however, resource constraints quickly banish such lofty ambitions to the realm of the idyllic. *A priori*, one does not know the structure of system-environment interactions concerning complex phenomena in precise detail; even if one did, for sufficiently large environments, such a description rapidly defies feasibility with respect to the amount of data required to be stored and manipulated in a reasonable time on a reasonable computer.

From an operational perspective, it is desirable to understand properties of a stochastic process from information that can be inferred from probing the system of interest alone. Intuitively, in this sense a stochastic process refers to the joint probability distribution that expresses the likelihood of a quantity taking certain values: the probability for a certain stock to have price x_1 and x_2 and x_3 on three consecutive days, for instance. Fluctuations of the stock price on any given day can directly influence the rest of the market (*i. e.*, its environment), which in turn impacts the price of the original stock at a later time.

Such memory effects are encoded as correlations in the joint probability distribution over the relevant timesteps and can manifest themselves as genuinely multi-time correlations. In the simplest non-trivial memoryless scenario, a process can exhibit only two-point correlations: the probability distribution of tomorrow’s stock price only depends upon today’s price, and not any further back in the history. However, more generally, when memory effects are at play, all multi-time correlations must be considered to unambiguously describe the process, as joint effects between sequences of events can play a significant role in the future evolution.

When attempting to generalise this understanding of stochastic processes to the quantum realm, perhaps unsurprisingly, the intricacies involved come to light. Just like in the classical realm, any realistic phenomenon must be described within the theory of *open systems* to account for the stochasticity that arises due to our subjective ignorance of the degrees of freedom of the environment [6–8]. In this formalism—as is the case for classical processes—dynamical phenomena are described solely in terms of accessible quantities that are derivable from the system of interest alone. The open systems framework has enjoyed tremendous recent success, translating fundamental theories into real-world predictions, and has led to a multitude of tools and techniques for manipulating quantum processes which have fostered many technological advances.¹ However, in stark distinction to the classical setting, here we must also deal with the fact that additional randomness arises at an elementary level through the very nature of *measurement* at the microscopic scale.

In quantum mechanics, measurements are fundamentally invasive, which seemingly leads to incompatibility of measurement statistics observed in time [12, 13]. Indeed, such invasiveness leads directly to a breakdown of the *Kolmogorov extension theorem* [14], which importantly links the operational description of a classical stochastic process in terms of joint statistics measured over a set of times to an underlying continuous-time mechanism [15]. Whenever measurements can directly influence the state of the system, it becomes seemingly impossible to define the process independently of the experimental interventions. This scenario is pertinent to both classical causal modelling [16], where an experimenter intervenes with the system in order to deduce causal relations between events, and quantum theory more generally [14, 17–19]. In any such theory, the hierarchy of joint probability distributions alone does not tell the entire story regarding the

¹ See, for instance, early work regarding the development of error-correcting codes for resilient quantum computing and the dynamical decoupling protocol for effective quantum control in Refs. [9, 10] and [11] respectively.

underlying process; thus, we are seemingly at a loss in answering *what we really mean when we talk about a quantum stochastic process*.

The lack of equivalence between the accessible and the underlying physical descriptions of quantum stochastic processes has irked the open systems and quantum information theory communities for some time, leading to nonequivalent definitions of key concepts pertinent to open quantum evolution; the most relevant to our present interests being that of *memorylessness*. Classically, a memoryless process is one for which the statistics observed at any point in time only depend on the most recent state of the system. Satisfaction of this condition has profound implications that importantly lead to a significantly simplified description of the process; indeed, the study of memoryless processes forms an entire branch of mathematics, and the evolution of many systems is frequently approximated to be memoryless, because of both the relative ease with which they can be analytically and computationally manipulated and the experimental evidence supporting this simplification [3–5, 20, 21].

However, the complications discussed above regarding a multi-time description for general quantum processes make it challenging to define the process independently of the interventions applied by the experimenter. Without such a description, there is no meaningful way to take intermediate measurements into account and check for conditional independence of the future evolution from the history. Thus, until recently, there has been no unique condition to define memorylessness in quantum mechanics. Nonetheless, many traditional approaches of the open systems formalism focus on the time-evolving state of the system of interest, which can provide valid witnesses for the presence of memory effects [22–24]. Whilst of immense practical importance, such descriptions do not serve to fully characterise the process, as they are limited in scope to capturing two-time correlations, specifically those between that of the state prepared at some point in time and the subsequent measurement statistics deduced at any later time. As is also true in the classical setting [25], when memory effects play a non-negligible role, all multi-time correlations must be considered.

Various frameworks circumvent the crucial problem of formalism that arises due to the invasiveness of measurements in quantum mechanics by actively taking measurements and *controllable* manipulations of the system of interest into account, enabling the separation of the underlying, *uncontrollable* process *per se* and the influence enacted by the experimenter. These were initially introduced by Lindblad [26] and Accardi, Frigerio and Lewis [27]. Modern variants of similar formalisms have been applied to study general quantum circuit architectures [28–31], foundational aspects of causal-

ity in quantum theory [32–37], quantum causal modelling [17–19, 38], quantum theory in spacetime [39–41], quantum game theory [42, 43], generalised communication protocols [44], non-equilibrium quantum thermodynamics [45, 46], and quantum processes with memory [1, 2, 14, 47–55].

Indeed, this school of thought has led to the development of a generalised Kolmogorov extension theorem that holds in any generalised probabilistic theory—including quantum theory—crucially giving rise to an operational definition of quantum stochastic processes [14, 27]. Most relevant to our present purposes, the breakthrough result of Ref. [49] provides an unambiguous criterion for a quantum process to be memoryless, thereby unifying all previous approaches. With this comprehensive mathematical language that captures all possible multi-time correlations at hand, one can properly describe quantum stochastic processes with memory independently of the experimenter and accurately understand important phenomena in which temporal correlations play a significant role, such as, *e. g.*, the emission spectra of quantum dots [56] and the vibrational modes of interacting fluids [57].

The discussion so far has centred on the existing literature; now we move to focus on the developments of this thesis. While the previously discussed frameworks are perfectly tailored to unambiguously define memory effects in quantum mechanics—and describe processes that display them—a thorough analysis of their structure, length, and strength is still missing. In particular, the concept of Markov order, which is regularly invoked in the study and simulation of classical stochastic processes with finite-length memory effects, has not been generalised to, or analysed in the context of, quantum mechanical processes. Such an investigation is of tremendous practical importance, as, although temporal correlations in complex phenomena are exhibited over various timescales, Markov order provides a natural notion of *memory length* that emerges in the context of statistical modelling, namely the amount of a system’s history that directly affects its future dynamics [4].

Classically, the concept of *Markov order*, ℓ , dictates that the statistics describing a system of interest at a given time only depend upon knowledge of its past ℓ observed states. Processes with finite Markov order therefore permit a significant reduction in modelling complexity: one must only estimate the conditional transition probabilities from the most recent set of observations, rather than the exponentially many more parameters for each additional timestep further back in the history [58–60]. Fortunately, many complex processes typically have an effectively finite-length memory, allowing for an efficient description that only considers the relevant portion of history necessary to

optimally predict the future [3–5]. In these cases, it is our understanding of the *memory strength* across a given duration of time that justifies the suitability of invoking finite-memory approximations

A thorough understanding and characterisation of memory effects in general stochastic processes is of crucial importance from both a theoretical and practical standpoint. Many theoretical developments regarding quantum systems have shown a superiority over their classical counterparts with respect to the successful enactment of certain information-processing tasks [11, 61–63]. Explicit understanding of memory effects will be of ever increasing importance and it is clear that future quantum technologies will need to embrace memory in order to display these advantages under realistic conditions [64].

The aforementioned modern frameworks pave the way for an unambiguous foray into the study of memory effects in quantum processes and provide the starting point for this thesis. There are, as we see it, a number of aspects regarding memory that must be considered for a holistic comprehension, namely the duration of time over which memory effects persist, the strength of their impact and the complexity of their simulation. The main concern and achievement of the present work regards extending the concept of finite memory length and the quantification of memory strength to quantum stochastic processes, and investigating the subsequent implications. Leveraging this vantage point, these concepts serve to provide a cohesive framework pertaining to the characterisation, simulation and exploitation of memory.

1.1 READING GUIDE

The outline of this thesis is as follows. In Part II, we present the story so far: the current understanding of physical processes in the presence of noise and memory effects therein. No new results are presented here; rather, the purpose of this part is to interrogate why a proper characterisation of memory effects in quantum processes has hitherto seemed an insurmountable task, and elucidate how recent theoretical advancements have provided a resolution, thereby suitably equipping us to overcome this concern. We begin in Chapter 2, where we first discuss classical stochastic processes in order to lay the foundations of some core concepts, before turning our attention to the quantum realm. Here, as we shall see, things become slightly more nuanced: although the standard approaches of the open systems paradigm are adequate to describe memoryless processes, they turn out to be fraught with problems in describing those with memory, as explored throughout the culminating discussion of this chapter. Chapter 3 formally presents the

mathematical formalism that circumvents the issues at hand by way of the process tensor framework. The first half of this thesis is thus deliberately pedagogical: many of these modern ideas have emerged in the context of a variety of disparate studies, but with a recently revamped understanding of quantum stochastic processes, we are now in a position to synthesise them into a cohesive story, as we attempt to undertake here.

With this unambiguous description of quantum stochastic processes at hand, we are then ready to present the novel results of this thesis, which are contained in Part III. We begin, in Chapter 4, by formalising the notion of memory length: we first generalise the concept of Markov order to open quantum processes, before exploring in detail the structure of processes with finite-length memory in Chapter 5. Many of the results presented here are developed through a number of illustrative examples that aim to build intuition rather than through general mathematical analysis, although the conclusions hold generically. Following this examination, in Chapter 6 we quantify the memory strength of a given process, analysing our proposed measures through application to an exactly solvable dynamics.

The main original contributions in this work are summarised as follows. In Chapter 4, we show that the introduced notion of quantum Markov order can be expressed as a constraint on the process tensor, provided in Eq. (4.4). We prove that only memoryless quantum processes can display finite Markov order with respect to all possible sequences of interrogations in Theorem 4.2. Nonetheless, it is sensible to consider the memory length in quantum processes with respect to specified choices of instruments; indeed, the classical notion of Markov order is one such special case. In Chapter 5, we provide the general description of a process with finite quantum Markov order in Theorem 5.1. Examples 5.1–5.3 highlight the implications for processes with finite memory length with respect to natural choices of manipulations applied to the system, including unitary transformations, measurements followed by independent re Preparations, and sharp, projective measurements. In light of the general structure deduced, we subsequently show that such processes do not necessarily have vanishing quantum conditional mutual information through Theorem 5.3. In Chapter 6, we introduce a number of measures for the memory strength of a process with finite quantum Markov order, particularly focusing on various important operational scenarios, such as when an experimenter does nothing to the system or, on the other extreme, actively tries to erase temporal correlations. The chapter culminates with an explicit study of the behaviour of this instrument-specific memory strength for a tuneable two-qubit model.

Lastly, in Part [IV](#), we conclude with a summary of our present work and a brief outlook. Note that in Appendix [A](#) we provide a synopsis of the notational conventions, as well as explicitly define a number of common functions and outline a conceptually important colour-coding scheme, that we employ throughout this thesis in order to aid the reader. Before proceeding, we wish to make the following disclaimer: to enhance legibility, we tend to cite references in which conceptual points of interest have been studied at their initial introduction, and refrain from repeatedly citing such references throughout subsequent discussion, except for the cases where direct results are pertinent or the citation is particularly poignant.

Part II

NARRATIVE ON THE EDGE

The nature of parties has been imperfectly studied. It is, however, generally understood that a party has a pathology, that it is a kind of individual and that it is likely to be a very perverse individual. And it is also generally understood that a party hardly ever goes the way it is planned or intended.

— John Steinbeck, *Cannery Row*.

CLASSICAL AND QUANTUM DYNAMICS WITH NOISE

ANY REALISTIC MODEL OF OBSERVED PHENOMENA must take into account the possibility of randomness. That we are often interested in quantities that are not deterministically predictable in advance, but rather exhibit some random variation, dictates that our descriptions be built from probabilistic models. Stochastic processes are ubiquitous in nature: they arise in a number of ways throughout the physical, biological and social sciences [3–5]. One way is through our *subjective ignorance*. Consider, for instance, a coin toss. Here, the classical laws of mechanics deem the underlying process to be deterministic—if an experimenter were to flip a coin in *exactly* the same fashion with *exactly* the same external conditions every time, they could predict each outcome with certainty.

However, realistically, this is not possible to achieve. Even if the experimenter could precisely apply the same force at the same point on the coin when initiating each flip, in each trial the coin is subject to different external factors that influence its trajectory, *e. g.*, those due to relentless collisions with molecules in the air. Although it is possible, in principle, to track enough variables to account for all of the additional degrees of freedom that the coin interacts with—thereby deterministically modelling the evolution—this is practically unfeasible due to sheer weight of numbers. Although the process is fundamentally deterministic, we are forced to deal with the fact that the outcomes observed look random to us due to our subjective ignorance, which is the root of all classical randomness.

Classical stochastic processes are thus characterised by a joint probability distribution over sequences of events in time. Crucially, from the correlations encoded in this multi-time distribution, all memory effects of the process can be deduced. For instance, a special case that has been extensively studied due to its particularly simple structure are *Markovian* or *memoryless* processes, in which only two-point correlations between events on adjacent timesteps are present. Here, the statistics observed at any point in time are completely determined by the most recently observed state. A more general

scenario where memory effects persist for a finite length of time is captured by the notion of *Markov order*, ℓ , in which the conditional statistics at any point in time only depend upon the most recent ℓ observations. In this case, multi-time effects can play a significant role, and the complexity of describing such processes grows exponentially in the length of the memory. Nonetheless, when the memory length is substantially less than the number of timesteps over which the process is defined, higher-order Markov models provide an alluring reduction in the computational resources required for accurate simulation [3–5].

Perhaps unsurprisingly, describing stochastic processes in quantum mechanics is somewhat more challenging. Here, in addition to randomness that arises from subjective ignorance, the outcomes of measurements are *fundamentally* random. Even in the static scenario, *i. e.*, when repeatedly measuring one and the same quantum state, we cannot describe the properties of any quantum system with certainty—the best we can do is to repeatedly perform experiments and make measurements to build up statistics that allow us to infer the quantum state. Just as in the classical case, our subjective ignorance can also lead to randomness, which necessitates treating realistic evolutions with the theory of open quantum systems [7, 8]. This formalism considers a joint system-environment state evolving in accordance with the laws of quantum mechanics, with the environmental degrees of freedom being regarded as inaccessible.

One way to understand correlations in time is via the dynamical map formalism. It bypasses the underlying details of the system-environment dynamics by expressing the effective mapping of the instantaneous quantum state of the system from some point in time to a later one. By preparing input states and performing subsequent measurements on the output states, the dynamical map prescribing the evolution between the two timesteps can be uniquely determined [8, 51, 65]. Importantly, any such description of a stochastic process, which only captures correlations between *pairs* of timesteps—as are frequently used in the experimental study of open quantum systems [22–24]—do not tell the whole story. Indeed, many non-Markovian processes can lead to two-point measurement statistics that could be incorrectly classified as Markovian [25, 49]. Specifically, when there is no memory in the process, two-point characterisations suffice to provide the correct description of the process at hand, but are insufficient otherwise.

As summarised by van Kampen [25]: “non-Markovian processes... cannot be considered merely as corrections to the class of Markov processes but require special treatment”. What is meant by this quote is that, in order to properly characterise processes with genuine memory effects, one must take into consideration all of the possible multi-time correlations, as two-point descriptions necessarily lack such information, which may

prove vital in the evolution. However, as discussed in Chapter 1, the invasiveness of measurements in quantum theory make it difficult to delineate between the active influence of the experimenter and the underlying process, which proves problematic in developing a multi-time description of quantum processes [14]. Since our aim is to understand memory effects in such processes, the first problem we must address in this thesis is an unambiguous understanding and operationally meaningful description of *quantum stochastic processes* that explicitly captures multi-time correlations.

In this chapter, we formally introduce classical stochastic processes and the formalism of open quantum systems in Sections 2.1 and 2.2 respectively. Stochastic processes are well-understood in classical theory, but not so in the quantum realm: there, for instance, a hierarchy of notions of non-Markovianity abound, which do not agree with each other in general. We present a brief literature review of such witnesses for non-Markovianity in order to highlight important gaps in the previous state of knowledge in Section 2.3. The incompatibility of the myriad of such definitions has been artificially reconciled by many through the belief that there exists no unique condition for Markovianity in the quantum realm. The overarching goal of this chapter is to explain why this is not true: we reiterate how the problems at hand arises from a breakdown of the standard formalism to properly describe quantum processes with memory, rather than a fundamentally irreconcilable problem. This motivation leads us naturally to an operationally meaningful framework for unambiguously describing quantum stochastic processes, which is the focus of Chapter 3.

2.1 CLASSICAL STOCHASTIC PROCESSES

We begin this section by presenting the key concepts needed to describe states of classical physical systems, before focusing on how these evolve over time. Most of the notions introduced here are explored within various contexts in a number of excellent textbooks, especially Refs. [3–5, 15, 20, 21].

2.1.1 Probability Spaces

The first primitive concept we need is that of a *sample space*, Ω , denoting the set of all possible outcomes for an experiment. This set can either be discrete—*e.g.*, the experiment of flipping two coins, each with outcomes heads (H) or tails (T), has the sample space $\Omega = \{\text{HH}, \text{HT}, \text{TH}, \text{TT}\}$ —or continuous—*e.g.*, the experiment of randomly

selecting a chord on a circle of given radius, which has the sample space $\Omega = \{a, b : a \text{ and } b \text{ are points on the circle}\}$.

Any subset of the sample space $\omega \subseteq \Omega$ can correspond to an *event*, which represents something that an experimenter can resolve or might be interested in. In the experiment of tossing two coins, the event that *at least one head occurs* is represented by $\omega = \{\text{HH}, \text{HT}, \text{TH}\}$. The collection of such events describes the overall *event space* of the experiment, \mathcal{G} , which constitutes a σ -algebra.

Definition 2.1 (σ -Algebra). Given some set Ω , let $\mathcal{P}(\Omega)$ represent its power set. A subset $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ is a σ -algebra if it satisfies:

1. \mathcal{G} contains the set Ω itself: $\Omega \in \mathcal{G}$.
2. \mathcal{G} is closed under complement: if $g \in \mathcal{G}$, then $(\Omega \setminus g) \in \mathcal{G}$.
3. \mathcal{G} is closed under countable unions: for $\{g_i\} \in \mathcal{G}$, their union $g = \cup_i g_i \in \mathcal{G}$.

These properties immediately imply that the emptyset is an element of \mathcal{G} and that \mathcal{G} is also closed under countable intersections (via De Morgan's law) [4].

A set Ω and a σ -algebra \mathcal{G} together constitute a *measurable space*, (Ω, \mathcal{G}) . The specific σ -algebra pertinent to an experiment is dictated by the events of interest, or, put differently, the questions deemed important by the experimenter or by what can be resolved. Regarding the example of tossing two coins and asking question: *did at least one head occur?*, the particular σ -algebra chosen is $\mathcal{G} = \{\emptyset, \text{TT}, \{\text{HH}, \text{HT}, \text{TH}\}, \Omega\}$, with the interpretation of individual events respectively being: *the experiment was not performed*, *no*, *yes*, and *the experiment was performed*.

The final necessary ingredient is the assignment of probabilities to events through a *probability measure*, $\mu : \mathcal{G} \rightarrow \mathbb{R}^+ \cup \{0\}$, which maps each event to non-negative real numbers in the following logical manner.

Definition 2.2 (Probability measure). A measure μ defined on a measurable space (Ω, \mathcal{G}) , where Ω is some set and $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ is a σ -algebra, is a probability measure if it satisfies:

1. The probability of any event is non-negative and real: $\mu(g) \in \mathbb{R}^+ \cup \{0\} \forall g \in \mathcal{G}$.
2. Some event occurs with certainty: $\mu(\Omega) = 1$.
3. The probability associated to any union of pairwise disjoint subsets of the event space is the sum of probabilities associated to each subset (σ -additivity):

$$\mu(\cup_i g_i) = \sum_i \mu(g_i).$$

These axioms immediately imply the following elementary properties:

1. *Probability of the Emptyset*: $\mu(\emptyset) = 0$.
2. *Monotonicity*: if $g_A \subseteq g_B \in \mathcal{G}$ then $\mu(g_A) \leq \mu(g_B)$.
3. *Boundedness*: $0 \leq \mu(g) \leq 1$ for all $g \in \mathcal{G}$.

Overall, the triple $(\Omega, \mathcal{G}, \mu)$ formally describe a *probability space*, providing the mathematical underpinning of any probabilistic theory [15].

In describing properties of a physical system, an experimenter is often interested in some function of the outcomes of an experiment rather than the outcomes themselves, such as average values. These scenarios are best encapsulated with the notion of a *random variable*, $Y : \Omega \rightarrow \Gamma$ (these are usually taken to be real-valued functions, *i. e.*, $\Gamma = \mathbb{R}$, as is assumed from now on), which labels possibly abstract events in a meaningful way, *i. e.*, such that

$$Y^{-1}(y) = g \in \mathcal{G} \quad \forall y \in \mathcal{Y}, \quad (2.1)$$

where \mathcal{Y} is a new σ -algebra generated by a collection of subsets of Γ . Since Ω is the sample space of a probability space, the new measurable space (Γ, \mathcal{Y}) naturally inherits the probability measure μ through a pushforward relation, which defines a *probability distribution* $\mathbb{P} : \mathcal{Y} \rightarrow [0, 1]$ via¹

$$\mathbb{P}(y) = \mu(Y^{-1}(y)). \quad (2.2)$$

The properties of the probability measure ensure that any such probability distribution is non-negative, normalised and additive over its σ -algebra. In summary, a random variable provides a logical relabelling, allowing us to begin with some abstract probability space $(\Omega, \mathcal{G}, \mu)$ and end up with a probability space $(\Gamma, \mathcal{Y}, \mathbb{P})$ that is potentially more amenable to further computational manipulation. Although random variables introduce versatility into the description of stochastic processes, we will ultimately only be concerned in probabilities of events and so the choice of random variable will not be of particular interest. The probability distribution \mathbb{P} , which is directly accessible via statistical measurements, defines what we will call a *state* of a classical system.

¹ Typically, such a distribution is labelled $\mathbb{P}_Y(y)$; in this thesis, we omit such unnecessary labels wherever possible for the sake of notational economy. Moreover, we are somewhat lax in often allowing the values x, y, \dots to represent either events or individual elements in the sample space, and using the same symbol \mathbb{P} to denote either a probability measure or a probability density.

2.1.2 Stochastic Processes

In describing *stochastic processes*, one can build time into the picture by way of a parameterised random variable. Just as a classical state is described by a probability distribution assigning probabilities to events, a classical stochastic process can be defined as a function allocating probabilities to sequences of events over time. This formulation is as follows. Firstly, upon defining a random variable, Y , an infinite collection of random variables automatically arises, namely any quantity $X := f(Y)$ that is some function of Y . Selecting a value $Y = y \in \mathcal{Y}$ gives a deterministic value for $X = f(y) =: x \in \mathcal{X}$; specification of an event in this way is referred to as a *realisation* or *sample*. The transformation law relating the probability distributions over two such random variables is

$$\mathbb{P}(x) = \int dy \delta(x - f(y)) \mathbb{P}(y), \quad (2.3)$$

where δ denotes the Dirac-delta distribution. One could then consider a function that involves Y and some auxiliary parameter, $t \in \mathbb{R}$, usually taken to denote time: $X_Y(t) := f(Y, t)$. Upon inserting for Y one of its possible values y , we obtain a deterministic function of time $X_{Y=y}(t) = f(y, t) =: x(t) \in (\mathcal{X}, \mathbb{R})$, representing a realisation of the process or sample trajectory. The stochastic process itself is, overall, regarded as the probability distribution over trajectories

$$\mathbb{P}(x(t)) = \int dy \delta(x(t) - f(y, t)) \mathbb{P}(y(t)). \quad (2.4)$$

So far, we have introduced classical stochastic processes on axiomatic grounds. We now consider their relation to what is typically measured in experimental procedures. In practice, it is often sensible to store information about the sample trajectories observed in a discrete manner by specifying a finite number $n \in \mathbb{N}$ of timesteps, denoted $\Lambda_n := \{t_1, \dots, t_n\}$, at which observations are made in order to build up a statistical description of the process.² The choice of these timesteps is, in principle, entirely up to the experimenter, although may be dictated by experimental capabilities. By repeatedly performing experiments, the experimenter yields statistics that describe the probability for the system of interest to take the value x_1 at t_1 , x_2 at t_2 , and so on, until x_n at t_n .

² When such a set Λ_n is taken to be an ordered sequence, we will indicate this with the subscript label $n : 1$. In some cases, we will find it convenient to consider unordered sets of timesteps of a given cardinality, which we will indicate with the label Λ_n .

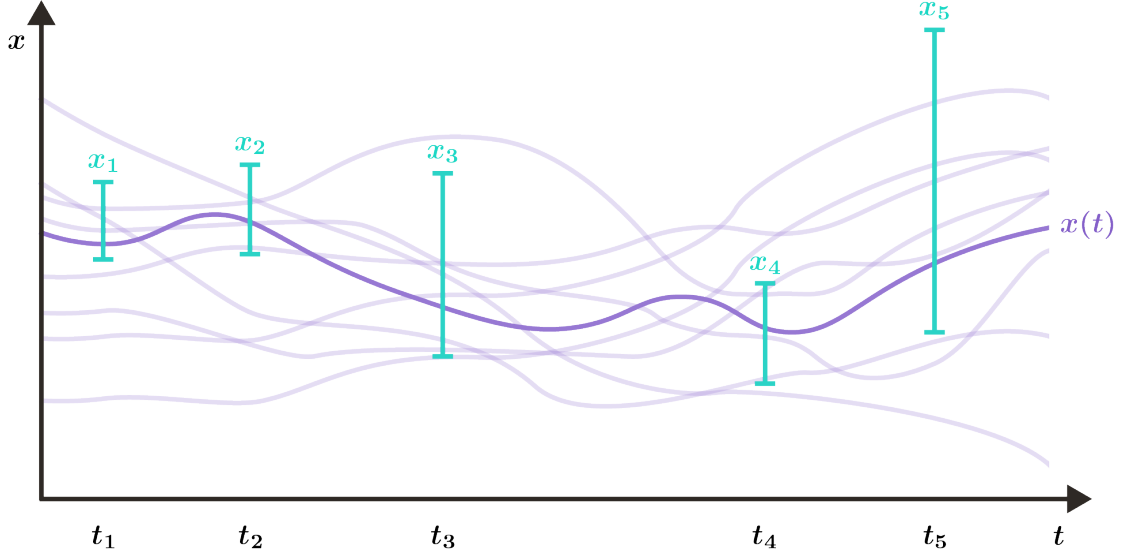


Figure 2.1: *Continuous and discrete-time stochastic processes.* In the continuous time picture, a stochastic process is characterised by a probability density $\mathbb{P}(x(t))$ over all possible sample trajectories $x(t)$ (purple). Alternatively, by measuring the system to be some regions x_1, \dots, x_5 at times t_1, \dots, t_5 (green), for instance, one deduces the multi-time statistics of the discrete-time joint probability distribution $\mathbb{P}_{5:1}(x_5, t_5; \dots, x_1, t_1)$ that characterises the process on the chosen timesteps. Only the bold trajectory can be measured to be in the regions depicted, and hence occurs with unit probability, although considering all possible different spatial measurement settings assigns non-zero probabilities for each trajectory to be realised.

This gives rise to a hierarchy of joint probability distributions written as

$$\begin{aligned}
 &\mathbb{P}_1(x_1, t_1) \\
 &\mathbb{P}_2(x_2, t_2; x_1, t_1) \\
 &\dots \\
 &\mathbb{P}_{n:1}(x_n, t_n; \dots; x_1, t_1).
 \end{aligned} \tag{2.5}$$

For example, $\mathbb{P}_{n:1}(x_n, t_n; \dots; x_1, t_1)$ might represent the probability to find a randomly moving particle in regions $\{x_1, \dots, x_n\}$ respectively at times $\{t_1, \dots, t_n\}$. We call a *joint probability distribution* over n timesteps an n -point distribution, since it contains the necessary information to calculate all n -point correlations. Insofar as the experimenter is concerned, the hierarchy of finite joint probability distributions above serves to characterise the stochastic process over the chosen timesteps.

It is not clear, *a priori*, that these two notions of stochastic processes are equivalent. This is a subtle point, but worth considering in some detail, as it will become crucial to our understanding of quantum stochastic processes. For the moment, we will distinguish between an underlying continuous stochastic process, as described by Eq. (2.4), and that

which is constructed on a finite number of timesteps, characterised by Eq. (2.5). Since the fundamental laws of physics are continuous in time, we are always implicitly assuming the existence of an underlying process that leads to the discrete-time statistics observed. In other words, we assume the existence of a continuous stochastic process that has all the finite ones that the experimenter measures as marginals. Beginning with Eq. (2.4) and specifying a finite set of timesteps, we can derive the hierarchy of joint distributions for all $t_j \in \Lambda_n$ as

$$\mathbb{P}_{j:1}(x_j, t_j; \dots; x_1, t_1) = \int dy \delta(x_j - f(y, t_j)) \dots \delta(x_1 - f(y, t_1)) \mathbb{P}(y). \quad (2.6)$$

It is easy to show that Eq. (2.6) implies that the n -point distribution $\mathbb{P}_{n:1}$ contains within it the correct descriptor of the process on any subset of times, which is deducible via *marginalisation*. To derive the statistics of the process on any subset $\Lambda_k \subseteq \Lambda_n$, one simply marginalises over the outcomes on the timesteps that are no longer of interest

$$\mathbb{P}_{\Lambda_k}(x_{\Lambda_k}, \Lambda_k) = \sum_{\Lambda_n \setminus \Lambda_k} \mathbb{P}_{\Lambda_n}(x_n, t_n; \dots; x_1, t_1). \quad (2.7)$$

Here the summation $\sum_{\Lambda_n \setminus \Lambda_k}$ runs over all realisations on the timesteps included in Λ_n but not Λ_k and x_{Λ_k} refers to the subset of possible outcomes corresponding to the times Λ_k . Thus, if a joint probability distribution arises from an underlying physical process, it necessarily satisfies this so-called *consistency* or *containment* property of Eq. (2.7), which allows one to derive the entire hierarchy of joint distributions in Eq. (2.5).

One of the pioneers of probability theory, Kolmogorov, was concerned with proving equivalence between the continuous and discrete descriptions of stochastic processes by establishing the alternate implication. That is, he endeavoured to understand the conditions a collection of finite joint probability distributions must satisfy in order for an underlying continuous process, with all of the finite ones as marginals, to exist. The *Kolmogorov extension theorem* (KET) says that satisfaction of the consistency condition outlined above implies the existence of such an underlying process, thereby proving equivalence between the descriptions given in Eqs. (2.4) and (2.5) [15]. Thus, practical motivations notwithstanding, the KET bridges the gap between the experimental reality we must face and solid mathematical underpinnings, importantly providing a definition of stochastic processes as the limit of families of finite probability distributions in time.

Intuitively, satisfaction of the consistency conditions means that once a process has been characterised on a set of timesteps, all behaviour on any subset of timesteps can be deduced by marginalising over the outcomes at the excessive times, as per Eq. (2.7) and depicted in Fig. 2.2. Thus, for a stochastic process over n timesteps, the n -point

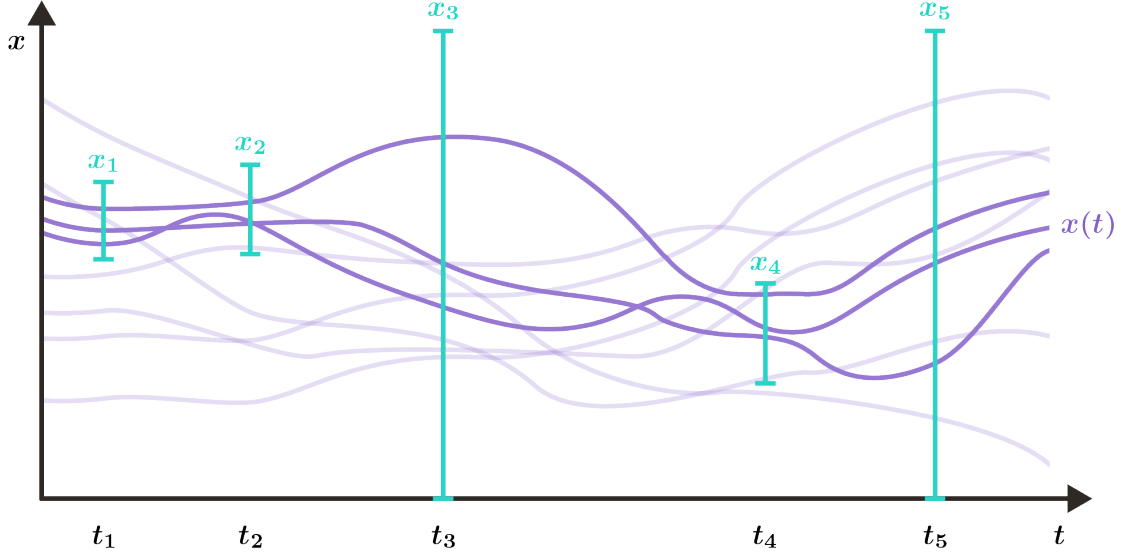


Figure 2.2: *Containment property for classical stochastic processes.* Given a joint probability distribution describing a classical stochastic process over some finite set of timesteps, *e.g.*, $\mathbb{P}_{5:1}(x_5, t_5; \dots, x_1, t_1)$, the correct description of the process on any subset of timesteps is calculated by marginalising over the outcomes at the excessive times. For example, the 3-step process over timesteps $\Lambda_3 = \{t_1, t_2, t_4\}$ is characterised by $\mathbb{P}_{\Lambda_3}(x_{\Lambda_3}, \Lambda_3) = \sum_{x_3 x_5} \mathbb{P}_{5:1}(x_5, t_5; \dots, x_1, t_1)$. Only the bold paths depicted have non-zero probabilities to be measured in the intervals x_{Λ_3} shown above, the weights of each of which are calculated from the original description by summing the five-step joint probabilities over all possible spatial intervals at t_3, t_5 , shown here as the extended vertical lines labelled x_3, x_5 .

joint probability distribution $\mathbb{P}_{n:1}$ completely characterises the process since it contains within it the entire hierarchy of Eq. (2.5). The fact that marginalisation is the correct way to obtain the contained description of a classical stochastic process is, loosely speaking, because there is no difference between having measured the statistics over all timesteps Λ_n and then discarding the observations we no longer care about, *i.e.*, summing over those on $\Lambda_n \setminus \Lambda_k$ (as per the r.h.s of Eq. (2.7)), and not having measured outcomes at those irrelevant timesteps anyway (as per the l.h.s).

To see this, consider tossing three fair coins consecutively, described by the joint distribution $\mathbb{P}_{3:1}(x_3, t_3; x_2, t_2; x_1, t_1) = \frac{1}{8}$ for each possible length-3 binary outcome sequence (*e.g.*, HHT). Given this distribution, if we are subsequently interested in describing the process at timesteps t_1 and t_3 only, we simply marginalise over the outcomes at t_2 , which gives: $\mathbb{P}_{3:1}(x_3, t_3; x_1, t_1) = \sum_{x_2} \mathbb{P}_{3:1}(x_3, t_3; x_2, t_2; x_1, t_1) = \frac{1}{4}$ for each length-2 binary outcome sequence. This procedure indeed provides the correct description of a process where two coins are tossed at times t_1 and t_3 . Marginalisation works because the state of the coin at time t_2 was *some* outcome, and it makes no difference whether we

average over all such possibilities or they are not even measured in the first place. The containment property, and indeed Eq. (2.6), implicitly assumes that we can measure realisations of outcomes throughout the process consistently without influencing the state of the system upon each interrogation. This is not fulfilled in more general stochastic frameworks such as classical causal modelling and quantum mechanics [14].

In summary, the connection between the continuous and discrete descriptions of a stochastic process is provided by the KET, which justifies our ability to work in either picture. For the purpose of this thesis, due to our operational perspective, we consider a classical stochastic process to be defined as follows.

Definition 2.3 (Classical stochastic process). A classical stochastic process over n timesteps is characterised by a joint probability distribution $\mathbb{P}_{n:1}(x_n, t_n; \dots; x_1, t_1)$ satisfying the containment property of Eq. (2.7) for all subsets of timesteps.

From now on, we drop the explicit labelling of the timesteps as arguments of the distribution and write $\mathbb{P}_{n:1}(x_n, \dots, x_1)$ to describe the process, with the subscripts indicating the timesteps on which the outcomes are observed. Additionally, although the sample space can be infinite, throughout this thesis we restrict our focus to the finite case.³

2.1.3 Modelling Stochastic Processes

Perhaps the most important reason for understanding stochastic processes from a practical perspective is to model them. Developing models that accurately simulate the statistics observed allows us to predict future behaviour of complex systems, *e. g.*, future stock market fluctuations or evolving population dynamics, amongst other applications [3–5]. While the joint probability distribution $\mathbb{P}_{n:1}(x_n, \dots, x_1)$ characterises the process at hand, containing all possible multi-time correlations between outcomes observed at different times, we are rarely, if ever, privy to such a detailed and resource-exhaustive description [3, 21]. Rather, what is typically feasible is to measure statistics describing the state of the system at each timestep, *i. e.*, the single-point marginals $\mathbb{P}_k(x_k) \forall t_k \in \Lambda_n$, and perhaps some lower order correlation terms (two- or three-point marginals), which describe how the state at some time is correlated with that at some others.

The complete description of the process contains significantly more information than can be deduced from such lower-order marginals; Indeed, any higher-order marginal

³ Almost all of the results presented in this chapter are extendable to the continuous case by replacing sums with integrals.

such as $\mathbb{P}_{\Lambda_k}(x_{\Lambda_k})$ for any proper subset $\Lambda_k \subset \Lambda_n$ remains insufficient to characterise a generic stochastic process [25]. Consequently, a computationally feasible *model* of a stochastic process does not generally specify the process at hand. Our assumption when modelling stochastic processes is that we can approximate the joint statistics from a manageable set of lower-order distributions. Put simply, the aim of modelling is to accurately reconstruct the joint distribution of the process from smaller ones [3, 4].

An important concept pertinent to modelling is that of *conditional probability distributions*. Consider a process characterised over n timesteps by $\mathbb{P}_{\Lambda_n}(x_{\Lambda_n})$ and suppose we observe a specific realisation over a subset of times $\Lambda_k \subseteq \Lambda_n$. Then

$$\mathbb{P}_{\Lambda_n \setminus \Lambda_k}(x_{\Lambda_n \setminus \Lambda_k} | x_{\Lambda_k}) = \frac{\mathbb{P}_{\Lambda_n}(x_{\Lambda_n})}{\mathbb{P}_{\Lambda_k}(x_{\Lambda_k})} \quad (2.8)$$

represents the conditional probability distribution over outcomes on the remaining timesteps $\Lambda_n \setminus \Lambda_k$ given that it took the values x_{Λ_k} on timesteps Λ_k . That we can calculate conditional distributions for stochastic processes in this way implicitly relies on the fact that $\mathbb{P}_{\Lambda_k}(x_{\Lambda_k})$ on the r.h.s is the correct descriptor of the process on timesteps Λ_k , *i. e.*, that the **KET** holds.

From Eq. (2.8), we can iteratively decompose any joint distribution as

$$\begin{aligned} \mathbb{P}_{n:1}(x_n, \dots, x_1) &= \mathbb{P}_n(x_n | x_{n-1}, \dots, x_1) \mathbb{P}_{n-1:1}(x_{n-1}, \dots, x_1) \\ &= \mathbb{P}_n(x_n | x_{n-1}, \dots, x_1) \mathbb{P}_{n-1}(x_{n-1} | x_{n-2}, \dots, x_1) \mathbb{P}_{n-2:1}(x_{n-2}, \dots, x_1) \\ &= \dots = \mathbb{P}_n(x_n | x_{n-1}, \dots, x_1) \dots \mathbb{P}_2(x_2 | x_1) \mathbb{P}_1(x_1). \end{aligned} \quad (2.9)$$

It is clear from this decomposition that we can build up the joint distribution from its constituent conditional distributions. Although a conditional distribution such as $\mathbb{P}_{\Lambda_n \setminus \Lambda_k}(x_{\Lambda_n \setminus \Lambda_k} | x_{\Lambda_k})$ is technically an $(n - k)$ -point joint probability distribution, we will refer to them as n -point *correlations*, since we assume the ability to calculate the k -point distribution $\mathbb{P}_{\Lambda_k}(x_{\Lambda_k})$ over the conditioning argument, which together provide the n -point distribution $\mathbb{P}_{\Lambda_n}(x_{\Lambda_n})$.

Such conditional distributions are often referred to as *transition probabilities*, as, *e. g.*, $\mathbb{P}_2(x_2 | x_1)$ represents the probability for the system to change its state from x_1 to x_2 . For a complete description of the process, one requires successively higher-order transition probabilities, *e. g.*, $\mathbb{P}_{3:1}(x_3, x_2, x_1) = \mathbb{P}_3(x_3 | x_2, x_1) \mathbb{P}_2(x_2 | x_1) \mathbb{P}_1(x_1)$, and so forth. Clearly, a complete description of a stochastic process contains a large amount of information. We can understand the complexity of the process in terms of the number of parameters needed to describe it. Suppose that a system of interest can take d distinct outcomes. For the initial distribution $\mathbb{P}_1(x_1)$, we must specify $d - 1$ numbers (*i. e.*, the

probability associated to each outcome, with the normalisation condition constraining one value); the joint distribution over two steps $\mathbb{P}_{2:1}(x_2, x_1) = \mathbb{P}_2(x_2|x_1)\mathbb{P}_1(x_1)$ requires $d^2 - 1$ specifications; and so on. In general, modelling an n -timestep evolution of a d -level classical system requires estimating $\mathcal{O}(d^n)$ transition probabilities. This exponential scaling in the number of timesteps quickly becomes intractable, hence the allure of developing accurate models built upon estimating lower-order transitions. Indeed, modelling means to fix the transition probabilities, which can be meaningfully done under the assumption that the memory is finite in length. We now consider the simplest case of such finite-memory processes.

2.1.4 Markovian Stochastic Processes

Example 2.1 (Perturbed Coin). Consider the toy classical process of a perturbed coin, depicted in Fig. 2.3. Here we have a coin resting on a piece of cardboard, which is being gently shaken at discrete times $t_k \in \Lambda_n$, resulting in a time-independent probability, $p > 1 - p$, for the coin to retain its previous orientation between each shake; with probability $1 - p$, the coin flips from H to T, or vice versa. The probability of the coin being in a particular state at arbitrary timestep t_k depends entirely on its most recent state, *i. e.*, the process is completely characterised by the family of two-point conditional distributions

$$\begin{aligned}\mathbb{P}_k(\text{H}_k|\text{H}_{k-1}) &= \mathbb{P}_k(\text{T}_k|\text{T}_{k-1}) = p \\ \mathbb{P}_k(\text{H}_k|\text{T}_{k-1}) &= \mathbb{P}_k(\text{T}_k|\text{H}_{k-1}) = 1 - p.\end{aligned}\tag{2.10}$$

Of course, temporal correlations between observations can be exhibited over various timescales; if one begins such a process with the coin facing H up, a few steps later it is more likely than not to be found in the same state. The crucial point is that, *once we know* the state at timestep t_k , we may as well discard any observations of previous states since they tell us no additional information.

The dependence of the future statistics on only the most recent outcome dramatically simplifies the complexity of any algorithm predicting its behaviour. This type of process is known as a *Markovian* or *memoryless* process, since the process itself stores no memory of historic outcomes; the only temporal correlations that can arise are mediated through the most recent state of the system. Markovian processes are formally defined as follows.

Definition 2.4 (Classical Markov process). A classical Markov process is described by a joint probability distribution $\mathbb{P}_{n:1}(x_n, \dots, x_1)$ whose conditional distribution at each timestep $t_k \in \Lambda_n$ only depends on the statistics of the most recent distribution at time t_{k-1} :

$$\mathbb{P}_k(x_k | x_{k-1}, \dots, x_1) = \mathbb{P}_k(x_k | x_{k-1}). \quad (2.11)$$

Note that for a Markovian process, the statistics depend *conditionally* on no more than the most recent timestep. A completely memoryless process is one in which the conditional statistics are independent of even the current state, such as a fair coin flip. Such processes are often said to be *super-Markovian* [4]; in this thesis, we generally drop this distinction and refer to both types of processes as Markovian or memoryless synonymously.

In contrast to the general case (see Eq. (2.9)), Markovian processes can be completely characterised with only an initial configuration $\mathbb{P}_1(x_1)$ and the collection of two-point transition probabilities $\mathbb{P}_k(x_k | x_{k-1})$, which specify the present state in terms of the most recent realisation of the system at each timestep. This is because for a Markovian process the overall joint distribution factorises as

$$\mathbb{P}_{n:1}(x_n, \dots, x_1) = \mathbb{P}_n(x_n | x_{n-1}) \dots \mathbb{P}_2(x_2 | x_1) \mathbb{P}_1(x_1). \quad (2.12)$$

Since each transition map is specified by $d^2 - d$ conditional probabilities and the initial single-point distribution by $d - 1$, we require at most $(n - 1)(d^2 - d) + d - 1$ parameters to describe a memoryless process, providing a significant reduction in complexity; this is perhaps the primary reason for the popularity of invoking the Markov assumption when modelling stochastic processes.

Equivalently, considering a discrete-time Markovian process on Λ_n and specifying an initial condition at t_1 , the state at an arbitrary later time $t_k \in \Lambda_n$ can be calculated via

$$\mathbb{P}_k = \mathcal{S}_{k,1} \mathbb{P}_1. \quad (2.13)$$

Here, $\mathcal{S}_{k,1}$ is a $d \times d$ matrix filled with conditional probabilities representing the likelihood of transition from each x_1 value to each x_k and \mathbb{P}_k is the d -dimensional vector

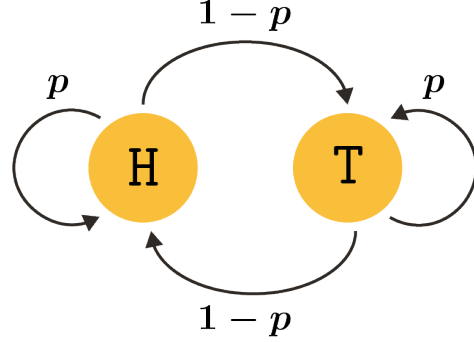


Figure 2.3: *Perturbed coin.* The perturbed coin is characterised by the transition probabilities prescribing the likelihood to retain its state or to flip at each shake.

representing the probability distribution at t_k . Comparing Eq. (2.13) with the entire description of the process up until the time specified, *i. e.*, Eq. (2.12) written up to t_k and marginalised over the outcomes on all timesteps except t_k , it is clear that transition matrix from some t_1 to t_k is simply a matrix multiplication of all of the intermediary transition matrices

$$\mathcal{S}_{k,1} = \mathcal{S}_{k,k-1}\mathcal{S}_{k-1,k-2}\dots\mathcal{S}_{2,1}. \quad (2.14)$$

Clearly, by grouping any sequence of matrix multiplications on the r.h.s of Eq. (2.14) into two stochastic matrices, we have satisfaction of the following *divisibility* property [4]

$$\mathcal{S}_{k,1} = \mathcal{S}_{k,j}\mathcal{S}_{j,1} \quad \forall t_k > t_j > t_1. \quad (2.15)$$

That the transition maps for a Markovian process are indeed conditional probability distributions dictates that their matrix representations have non-negative entries whose columns sum to unity, known as *stochastic matrices*. From an axiomatic perspective, stochastic matrices are of significant importance in classical physics, as their properties ensure that for any input probability distribution, the output of the map is always a valid probability distribution, thereby providing the most general unconditional transformation allowable in classical theory between two points in time. For reasons that will become clear, we refer to such processes that can be described by a divisible family of stochastic maps as (*positive*) *P-divisible* processes [66, 67].

As a brief aside, note that for a stationary *P-divisible* process, one can derive a closed form continuous-time equation of motion for the instantaneous state of the system, as is provided in Appendix B.1 for the sake of completeness. The classical *master equation* (ME) derived provides the most general form of an equation whose solution is guaranteed to provide a divisible single-parameter semi-group of stochastic maps [68].

We have seen already that a Markovian process lends itself to a complete description in terms of a divisible family of such two-point stochastic transition maps. However, it is well-known that *P-divisibility* is insufficient to classify a process as Markovian [67, 69–72]. The fact that non-Markovian processes can satisfy the *P-divisibility* criteria is emblematic of a deeper issue, namely that two-point information does not adequately capture the dynamics at hand. In other words, the core reason that the *P-divisibility* criteria does not imply Markovianity is because the former is based solely on two-point correlations which fail to capture multi-time effects and therefore do not provide full information about a generic process. Consequently, an experimental reconstruction of the two-point stochastic maps does not necessarily correspond to the actual (potentially multi-time)

conditional probabilities of the process. In other words, satisfaction of Eq. (2.15) does not imply Eq. (2.11): for a given process, one might be able to construct a stochastic map description satisfying Eq. (2.15) even when the process is non-Markovian.

The crucial point is that in such cases, in contradistinction to Markovian processes, the constituent divided portions of the stochastic maps *cannot* be identified with the actual conditional probabilities of the process; for a non-Markovian \mathbf{P} -divisible process, these sequences of stochastic maps are sensitive to initial conditions. The conditional probabilities in a non-Markovian process generally depend on multiple previous outcomes, which is necessarily overlooked by the two-point \mathbf{P} -divisibility criteria. We will return to this point with a more detailed discussion in Section 2.3; for now, the key message is that a proper description of non-Markovian processes must take into account all multi-time correlations and cannot simply be an extension to the study of Markovian processes as we know them.

2.1.5 Non-Markovian Stochastic Processes

From the discussion until now, we have seen that dealing with Markovian processes is much easier than with their non-Markovian counterparts, with the proper description of the latter requiring an exponentially-scaling amount of resources. This begs the question: *to what extent do we need to worry about a proper understanding of non-Markovian processes?* As noted by van Kampen: “non-Markov is the rule; Markov is the exception” [25]. Formally, in the space of all stochastic processes, the Markovian ones form a non-convex set of measure zero (with respect to any meaningful, non-singular measure); in other words, they are isolated special cases. The question of whether or not Markov processes exist in nature—and under which circumstances—is thus of significant interest. For instance, a random walk with the choice of each step being independently and identically distributed is a Markov process, so too is an experiment of picking coloured marbles out of a bag with replacement. However, as soon as non-negligible and realistic effects come into play, such as a direction bias for a random walker or the lack of replacement of marbles, such processes are almost always rendered non-Markovian.

To illustrate, consider the situation pictured in Fig. 2.4. Suppose an experimenter picks marbles from a bag whose initial distribution comprises a fraction r of red, b of blue and g of green coloured marbles, such that $r + b + g = 1$. Suppose that, according to whether the marble chosen at some timestep t_k is red, blue or green, a random walker will respectively step left ($s_k = -1$), stay put ($s_k = 0$) or move right ($s_k = 1$). The

position of the walker at t_n can be described by the random variable $X_n := \sum_{k=1}^{t_n} s_k$. If the marble chosen at each timestep is replaced before the next one picked out, then the random walk, $\mathbb{P}_{n:1}(x_n, \dots, x_1)$, is Markovian, since the probability for the walker to step in each direction at a given timestep remains unchanged by withdrawing any marble from the bag. If, on the other hand, the marbles are not replaced, then the probability for the walker to step in each direction changes with each marble withdrawal, since the relative proportions of the contents are modified. We cannot determine this change from the instantaneous position of the walker alone; instead, we require knowledge of the entire history up until the relevant point in time to determine the current composition of the bag, hence the probabilities for the walker to step in each direction.

Indeed, in many other physically realistic situations of interest, similar finite-sized effects lead to non-negligible memory effects in the process. For example, processes whose microscopic equations of motion are second order differential equations in time require two initial conditions (*i. e.*, position and velocity) to determine the statistics over future positions. Hence, processes considering the position of an initial distribution of particles can only be considered Markovian at sufficiently long timescales that the initial velocity is forgotten [4]. This is a special case of a Langevin-type equation; other continuous time Markovian processes are Wiener processes and Cauchy processes. Their extensive

study has shed light on the type of underlying dynamics that give rise to Markovian processes [4, 5]. Loosely speaking, the essential idea is that the system of interest couples weakly to a large environment: the weakness of the coupling ensures that the environment is relatively unperturbed by interactions with the system and the largeness of the environment ensures that from each timestep to the next, the system interacts with a fresh portion of the environment. Modelling and designing mesoscopic processes, where such assumptions are no longer satisfied and finite-sized effects come into significance, clearly requires an understanding of non-Markovian processes.

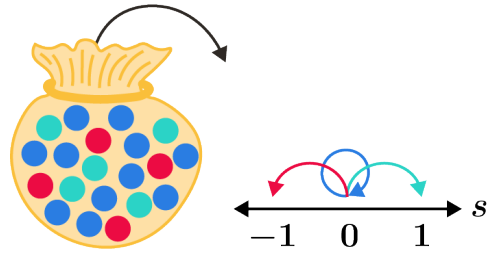


Figure 2.4: *Random walk conditioned on marbles drawn.* A bag is filled with a fraction of different coloured marbles: here we have $\frac{1}{5}$ red, $\frac{1}{2}$ blue and $\frac{1}{4}$ green. A marble is drawn at each time and depending on its colour, a walker steps left (red), stays put (blue) or steps right (green). If the marble is replaced after each step, the process is Markovian; otherwise it is not, since the extracted marbles disrupt the relative proportions of the bag.

2.1.6 *Classical Markov Order*

Although generic non-Markovian processes might require knowledge of their entire course of history to predict the future evolution, there is an important class of non-Markovian processes that are nonetheless feasible to model with a reasonable amount of resources: those with *finite memory length*. Formally, the natural way to account for such memory effects that are finite in duration is through the notion of *Markov order*, ℓ , which dictates that the statistics observed at any given time only depend upon knowledge of the past ℓ outcomes. For example, we can generalise the random walk process conditioned on marbles considered above to incorporate longer memory effects by holding out the marbles drawn for a certain number of timesteps before they are replaced. In this case, although the statistics of the next state depend conditionally upon the most recent sequence of outcomes, just like in the Markovian case, the salient point is that once we have knowledge of the most recent sequence of outcomes, we may as well discard observations from the prior history.

By grouping together timestep sequences of length ℓ , such non-Markovian processes can be tamed into ones that behave like Markovian processes, albeit on a larger state space extending over a period of time. Although processes with finite Markov order are close in spirit to Markovian processes, unlike their truly Markovian counterparts, multi-time correlations can play a crucial role in the future evolution of dynamics, giving rise to a potentially complex memory that is nonetheless limited in duration. To capture such behaviour, rather than begin with a straightforward extension of Markovian processes which do not account for such effects, we are forced to begin with a proper description of non-Markovian stochastic processes, *i. e.*, Eq. (2.5), and study the circumstances in which the process displays finite-length memory.

Intuitively, the concept of Markov order boils down to the following question: *is knowledge of a portion of the history of a process sufficient to predict the statistics of its future evolution?* In other words, given knowledge of the statistics over a sequence of ℓ timesteps, $\{t_{k-\ell}, \dots, t_{k-1}\}$, one can perfectly predict, in principle, the statistics to be expected at timestep t_k for such a process. The conditional distribution for arbitrary $t_k \in \Lambda_n$ of a process with Markov order- ℓ is therefore expressed as

$$\mathbb{P}_k(x_k | x_{k-1}, \dots, x_1) = \mathbb{P}_k(x_k | x_{k-1}, \dots, x_{k-\ell}). \quad (2.16)$$

In the interest of developing an economical notation for the remainder of this thesis, given an n -step stochastic process, we demarcate the timesteps into three intervals: the

future $F := \{t_n, \dots, t_k\}$, the memory $M := \{t_{k-1}, \dots, t_{k-\ell}\}$, and the earlier history $H := \{t_{k-\ell-1}, \dots, t_1\}$ (in principle, the history and future can extend to infinitely long times). The realisations observed over these timesteps x_j describing the system of interest are grouped together similarly as $\{x_F, x_M, x_H\}$. The Markov order of the process is defined in terms of the conditional statistics of outcomes as follows.

Definition 2.5 (Classical Markov order). A classical stochastic process has Markov order- ℓ if the conditional probability for any outcomes x_F at or beyond any time $t_k \in \Lambda_n$ depends only on the realisations x_M over the previous ℓ timesteps, and not on those x_H of the earlier history

$$\mathbb{P}_F(x_F|x_M, x_H) = \mathbb{P}_F(x_F|x_M). \quad (2.17)$$

As special cases, $\ell = 1$ corresponds to a Markovian process (see Def. 2.4) and $\ell = 0$ a super-Markovian (completely random) process.

The property of Markov order- ℓ constrains the underlying joint probability distribution characterising the process, from which the above conditional distributions arise. To highlight the important reduction of complexity in the description of a process with finite Markov order, compare the following decomposition with the general expression of Eq. (2.9)

$$\mathbb{P}_{n:1}(x_n, \dots, x_1) = \prod_{j=\ell+1}^n \mathbb{P}_j(x_j|x_{j-1} \dots, x_{j-\ell}) \mathbb{P}_{\ell:1}(x_\ell, \dots, x_1). \quad (2.18)$$

To reiterate, ℓ determines the number of timesteps over which one must observe states in order to optimally predict, in principle, the next state, thereby providing a natural and fundamental timescale for *memory length* in stochastic processes. This property is of tremendous practical importance, as processes with finite Markov order can be effectively reduced to Markovian processes upon a suitable grouping of timesteps, allowing for efficient simulation [25].

An alternative but equivalent way of expressing the notion of Markov order is that any statistics that an experimenter might deduce over the history and the future timesteps are *conditionally independent* with respect to an intermediate sequence of realised values

$$\mathbb{P}_{FH}(x_F, x_H|x_M) = \mathbb{P}_F(x_F|x_M) \mathbb{P}_H(x_H|x_M). \quad (2.19)$$

Importantly, it is knowledge of the outcomes on the memory block that renders the future and history conditionally independent. To prelude the discussion to follow in our extension to the quantum realm, an alternative way to think of this is as a sequence

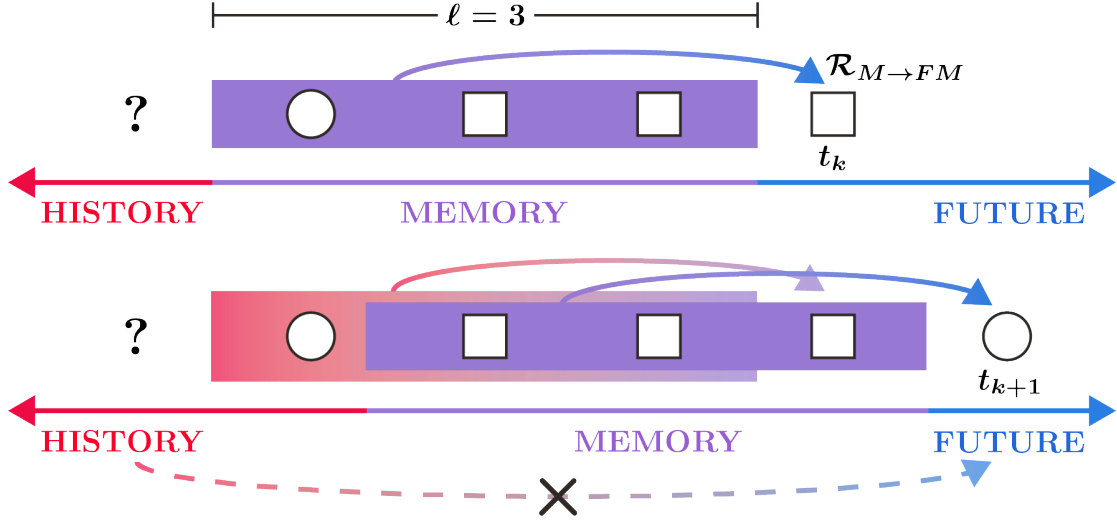


Figure 2.5: *Markov order as the natural notion of memory length.* Knowledge of the $\ell = 3$ states in the memory block is sufficient to predict the probabilities of future states by way of the recovery map, $\mathcal{R}_{M \rightarrow FM}$. In particular, no information about the prior history is required to determine these probabilities (indicated by the question mark). This property is independent of which timestep is being considered; holding equally well for steps t_k and t_{k+1} . At every step, any influence the history (beyond ℓ timesteps ago) has on the future must be mediated through the memory blocks. Importantly, states observed in the memory can jointly influence the future statistics. Nonetheless, conditional on the statistics realised in the most recent block, there can be no correlations between history and future, as indicated by the faded, dashed arrow.

of interventions (for instance, measurements) on the system that serves to block any possible historic influence on the future dynamics for each outcome realised.

Importantly, while Markov order- ℓ means that the state of the process at any time only depends conditionally upon the previous ℓ states, it does not imply an absolute separation of the timesteps into blocks of memory and irrelevant history. In other words, the probability distribution $\mathbb{P}_{FMH}(x_F, x_M, x_H)$ factorises conditionally, but we do not necessarily have $\mathbb{P}_{FMH}(x_F, x_M, x_H) = \mathbb{P}_F(x_F)\mathbb{P}_M(x_M)\mathbb{P}_H(x_H)$, or $\mathbb{P}_{FMH}(x_F, x_M, x_H) = \mathbb{P}_{FH}(x_F, x_H)\mathbb{P}_M(x_M)$. Instead, the memory blocks corresponding to different timesteps overlap, allowing for the existence of unconditional correlations between timesteps with a separation greater than ℓ in general. These are themselves often referred to as memory, however, these temporal correlations are always mediated through overlapping memory blocks as show in Fig. 2.5.

Lastly, a classical stochastic process with Markov order- ℓ can be equivalently characterised by the following two statements. Firstly, from an operational perspective, the significance of finite Markov order is best encapsulated through the existence of a so-called *recovery* map $\mathcal{R}_{M \rightarrow FM}$, which acts only on M to give the correct future

statistics: $\mathbb{P}_{FMH} = \mathcal{R}_{M \rightarrow FM}(\mathbb{P}_{MH})$. This map can be directly used to simulate future dynamics, and the complexity of any predictive model is fundamentally upper-bounded by the length of the block M on which it acts (as well as by the number of possible values for each x_j). Secondly, an entropic characterisation that is convenient to check in practice states that the classical *conditional mutual information* (**CMI**) vanishes: $I^{\text{cl}}(F : H|M) := H(\mathbb{P}_{FM}) + H(\mathbb{P}_{MH}) - H(\mathbb{P}_{FMH}) - H(\mathbb{P}_M) = 0$, where $H(\mathbb{P}) := -\sum_x \mathbb{P}(x) \log \mathbb{P}(x)$ is the Shannon entropy. The equivalence between these statements is trivial: satisfaction of Eq. (2.17) implies the distribution factorises as $\mathbb{P}_{FMH}(x_F, x_M, x_H) = \mathbb{P}_F(x_F|x_M)\mathbb{P}_{MH}(x_M, x_H)$; the recovery map $\mathcal{R}_{M \rightarrow FM}$ can then be chosen to act as multiplication by the higher-order stochastic transition map $\mathbb{P}_F(x_F|x_M)$. Equivalence to vanishing classical **CMI** is obvious by writing the **CMI** as a relative entropy between probability distributions (Kullback-Liebler divergence) as follows $I^{\text{cl}}(F : H|M) = \mathcal{D}^{\text{cl}}(\mathbb{P}_{FH|M} \| \mathbb{P}_{F|M} \mathbb{P}_{H|M})$, where $\mathcal{D}^{\text{cl}}(\mathbb{P}|\mathbb{Q}) := -\sum_x \mathbb{P}(x) \log \frac{\mathbb{P}(x)}{\mathbb{Q}(x)}$. The relative entropy vanishes iff the arguments are identical. Thus, in the classical setting, vanishing **CMI** is equivalent to finite Markov order.

We now briefly summarise the key points of this section that we will examine in detail when attempting to understand quantum stochastic processes. We have seen that classical stochastic processes can be characterised completely by a joint probability distribution over a discrete set of timesteps satisfying a natural consistency condition, with the existence of an underlying continuous-time physical process guaranteed by the **KET**. This allows for unambiguous calculation of conditional statistics, which are crucial to defining memoryless processes by way of the Markov condition. This special class of processes are the ones in which two-point information is sufficient for their characterisation. However, such two-point descriptions are destined to overlook multi-time memory effects and any criteria based upon them is not suitable for addressing non-Markovian processes. A particular example of this point is that divisibility does not imply Markovianity. Following this, we considered some properties of a non-trivial but practically important subset of non-Markovian processes which exhibit finite-length memory, namely those with finite Markov order. Our goal in the next section is to consider the extensions of these ideas into the quantum setting, where we will see that a number of subtleties must be addressed for a meaningful description of quantum stochastic processes.

2.2 OPEN QUANTUM DYNAMICS

In the classical setting, the finite-length memory approximation underpins the success of the often-invoked order- ℓ Markov models, which make use of information from only the past ℓ states to predict the next. However, even in the simplest non-trivial case of memoryless dynamics (*i. e.*, $\ell = 1$), the study of stochastic processes is vastly different in the quantum realm than its classical counterpart. This is mainly because, in quantum mechanics, one must necessarily disturb the system in order to observe realisations of the process, breaking an implicit assumption of the classical setting. In quantum mechanics, there is a continuous family of possible non-commuting observables that could be measured, and the choice of measurement at one point in time (or even whether to measure at all) can directly affect the future statistics [14, 51, 65, 73–75]. This is in stark contrast to the consistency conditions satisfied by the joint probability distributions corresponding to classical stochastic processes.

To highlight this issue explicitly, consider the following quantum experiment which is represented schematically in Fig. 2.6 (we follow the example presented in Ref. [14]). Begin with a spin- $\frac{1}{2}$ particle initially prepared in an equal superposition $\frac{1}{\sqrt{2}}(|\downarrow\rangle + |\uparrow\rangle)$ in the z -direction. Suppose an experimenter were to set up a sequence of Stern-Gerlach apparatus that allows them to measure the spin orientation at successive timesteps in any direction of their choosing. Consider the case where they measure the system at times $\{t_1, t_2, t_3\}$ respectively in the z -, x -, and z -directions, whose outcomes are represented by $\{\downarrow, \uparrow\}$ for z -direction measurements and $\{\rightarrow, \leftarrow\}$ for x -direction measurements. We assume that the quantum system undergoes trivial dynamics in between measurements. The first measurement in the z -direction has equal probability of $\frac{1}{2}$ to yield the result \downarrow or \uparrow ; in either case, once the outcome is observed, the post-measurement state is either $|\downarrow\rangle$ or $|\uparrow\rangle$ respectively. These states can be expressed as an equal superposition in the x -basis: $|\downarrow\rangle = \frac{1}{\sqrt{2}}(|\rightarrow\rangle + |\leftarrow\rangle)$ and $|\uparrow\rangle = \frac{1}{\sqrt{2}}(|\rightarrow\rangle - |\leftarrow\rangle)$. Thus, the spin measurement in the x -direction at the second timestep again yields each possible outcome \rightarrow or \leftarrow with equal probability, and the subsequent state is either $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\downarrow\rangle + |\uparrow\rangle)$ or $|\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\downarrow\rangle - |\uparrow\rangle)$ accordingly. The statistics of the final measurement are then identical to the first one. Thus, the probability to measure any sequence of outcomes in this case is uniformly distributed over the possibilities, *e. g.*, $\mathbb{P}_{3:1}(\downarrow_3, \leftarrow_2, \downarrow_1) = \mathbb{P}_{3:1}(\downarrow_3, \rightarrow_2, \downarrow_1) = \frac{1}{8}$.

On the other hand, consider an alternative experiment where the experimenter does *not* perform the x -direction measurement at the second timestep. Then, suppose for concreteness that the outcome \downarrow was observed at the first timestep, which occurs with

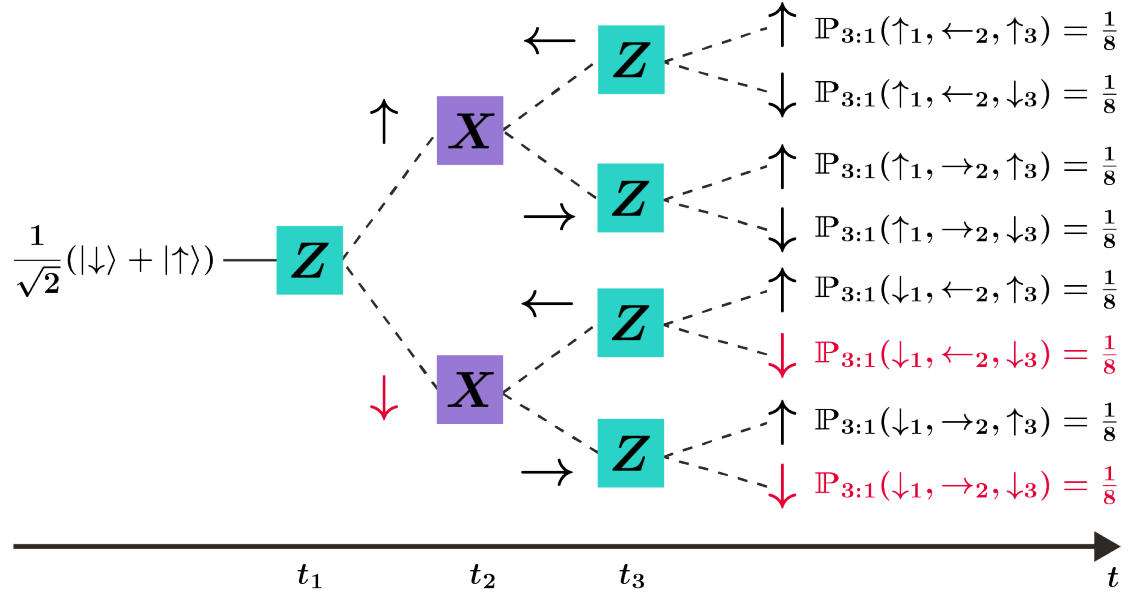


Figure 2.6: Breakdown of *KET* in a Stern-Gerlach experiment. An initial state prepared in an even superposition is subject to three sequential measurements, in the Z, X, Z directions. The joint statistics for each possible sequence of outcomes are equal to $\frac{1}{8}$ (shown on the right). However, marginalising over the outcomes observed at the second timestep does not provide the correct probabilities that are predicted by theory in the case where *no* measurement is performed there, highlighting the breakdown of the *KET*. If the first outcome is \downarrow and no intervention is made at t_2 , the measurement at t_3 yields \downarrow with certainty. This gives $\mathbb{P}_{3:1}(\downarrow_3, \downarrow_1) = \frac{1}{2}$, in contradiction with the marginalised statistics computed as the sum of probabilities displayed in red.

probability $\frac{1}{2}$ and leaves the system in the $|\downarrow\rangle$ state. When subsequently measured in the z -direction at the third timestep, without any intermediary measurement at t_2 , the outcome \downarrow is realised with certainty; thus we have $\mathbb{P}_{3:1}(\downarrow_3, \downarrow_1) = \frac{1}{2}$. This is not equal to a marginalisation over the possible outcomes of the second timestep, which gives $\mathbb{P}_{3:1}(\downarrow_3, \leftarrow_2, \downarrow_1) + \mathbb{P}_{3:1}(\downarrow_3, \rightarrow_2, \downarrow_1) = \frac{1}{4}$. The statistics of the process over the first and third timestep *cannot* be deduced by simply marginalising the 3-point distribution $\mathbb{P}_{3:1}$ over the outcomes of the second timestep; the standard *KET* clearly does not hold on the level of statistics measured in quantum processes.

The *KET* breaks down here on the level of measured statistics because implicitly assumes that there is only one method of probing the degrees of freedom of interest to observe outcomes, and that this probing does not actively change the state of the system. Neither of these assumptions are fulfilled in quantum theory (or, *e.g.*, more general classical stochastic theories with interventions—including causal modelling—which we briefly return our attention to in Section 4.2) [14, 15, 20]. Loosely speaking, it is the fact that for quantum processes, there *is* a difference between averaging over

all possible measurement outcomes and not having made a measurement at all, which essentially leads to violation of the KET [14]. Without a consistent method for deriving the contained descriptors of a process from that of its description defined over more timesteps, there is inherent ambiguity in what we mean when we talk about a quantum stochastic process. This leads directly to problems in defining conditional probabilities, and thus an unambiguous classification of memory in quantum processes.

This problem has irked the open systems and quantum information communities for some time, leading to various incompatible descriptions of quantum stochastic processes [76]. Conventional approaches attempt to sidestep this issue by describing properties of the process in terms of the time-evolving density operator of the system of interest, inherently failing to capture memory effects that only appear in multi-time correlations [22–24]; whilst others impose constraints on general system-environment interactions to facilitate a specific mechanism for memory transfer throughout the dynamics [77–79]. In either case, both types of approach fail to yield a comprehensive framework of quantum stochastic processes, and the corresponding definitions of memory are necessary but not sufficient to characterise Markovian processes [49].

As is evident in the classical case, one can only make limited assertions about a stochastic process from, *e.g.*, functions evaluated in terms of correlations between the state of the system at any two timesteps, as is provided by the solutions of generalised master equations. Subsequently, for completeness, we provide a traditional account of the open quantum systems approach to describing quantum stochastic processes. Along the way, we will highlight some critical shortcomings in order to motivate the necessity for the more general and operationally meaningful *process tensor* formalism to properly describe quantum processes [50], which we introduce in Chapter 3. Many of the notions presently introduced can be found in a number of excellent textbooks on the subject, for instance Refs. [6–8, 80–82].

2.2.1 Open Quantum Systems

The open systems paradigm acknowledges the fact that it is generally impossible to isolate a system of interest from its surroundings, and, as such, we must consider the effects of the external environment on the system [6, 7]. As we are typically unable to track the total evolution of the system and its environment—either by way of experimental limitation or lack of computational resources—the aim is to understand how the system of interest evolves dynamically. While the overall system and environment

evolves unitarily according to the standard Schrödinger equation, the evolution on the level of the system alone need not, due to the environmental influence.

We begin with a brief description of closed quantum dynamics which leads to the open setting by eventually restricting our consideration to a subsystem of interest. States of a quantum system, S , are described by bounded linear operators on an associated Hilbert space $\rho \in \text{BL}(\mathcal{H}^S)$ satisfying the following three properties

$$\rho = \rho^\dagger, \quad \rho \geq 0 \quad \text{and} \quad \text{tr}[\rho] = 1. \quad (2.20)$$

This *density operator* description is nicely tailored for our present purposes as it allows for the preparation of classical mixtures of pure quantum states which are pertinent to any stochastic theory, *e. g.*, a stochastic preparation procedure of a spin- $\frac{1}{2}$ particle where the $|\uparrow\rangle$ state is prepared with probability p and the $|\downarrow\rangle$ state is prepared with probability $1 - p$. Suppose we have a device preparing a system in one of a number of pure states $\{|\psi^{(i)}\rangle\}$ with probabilities $\{p_i\}$. The density operator of such an ensemble encodes all of the physically meaningful information about the system and is constructed as follows

$$\rho = \sum_i p_i |\psi^{(i)}\rangle \langle \psi^{(i)}|. \quad (2.21)$$

Pure states correspond to extremal states that cannot be written as a convex sum as above. In general, a density operator decomposition in terms of an ensemble of pure states is non-unique. The properties of Eq. (2.20) provide an intrinsic characterisation that holds independently of the ensemble interpretation, guaranteeing that its spectrum always represents a valid probability distribution. The Hermiticity condition ensures the eigenvalues are real; the positivity condition ensures they are all non-negative, thus corresponding to probabilities; and the trace condition stipulates an overall normalisation of probabilities. Importantly, density operators form a convex set: consider an ensemble of density operators $\{\rho^{(i)}\} \in \text{BL}(\mathcal{H}^S)$ distributed with probabilities $\{p_i\}$. Then

$$\rho = \sum_i p_i \rho^{(i)} \quad (2.22)$$

is a valid representation of a quantum state.

In the density operator picture, the closed unitary evolution described by the Schrödinger equation translates into the *von Neumann* equation

$$\frac{\partial \rho_t}{\partial t} = -i[H_t, \rho_t], \quad (2.23)$$

where $[A, B] := AB - BA$ represents the commutator and we set $\hbar = 1$. Integrating the von Neumann equation admits a formal solution in terms of a unitary map, $\mathcal{U}_{t,0}$, that expresses the time-evolution of the system

$$\rho_t = \mathcal{U}_{t,0} \rho_0 := U_{t,0} \rho_0 U_{t,0}^\dagger, \quad (2.24)$$

where ρ_0 is the initial system state and the family of unitary operators are generated by a Hamiltonian $H_{t'}$

$$U_{t:0} = \mathcal{T} \left\{ \exp \left(-i \int_0^t dt' H_{t'} \right) \right\}, \quad (2.25)$$

where \mathcal{T} represents the time-ordering operator.

The theory of open quantum systems concerns the evolution of a system, S , and some uncontrollable and inaccessible environment, E . The environment can describe any degrees of freedom that develop in time with the state of the system, *e.g.*, the system of interest could be a spin- $\frac{1}{2}$ particle and its environment an infinite field of bosonic modes. The joint Hilbert space of the system and environment is the tensor product of each subspace, $\mathcal{H}^{SE} = \mathcal{H}^S \otimes \mathcal{H}^E$, with the joint state space described accordingly as the subset of $\text{BL}(\mathcal{H}^{SE})$ satisfying Eq. (2.20). The tensor product operation thus provides a natural way to represent composite systems: given some quantum states $\rho^S \in \text{BL}(\mathcal{H}^S)$ and $\tau^E \in \text{BL}(\mathcal{H}^E)$, the composition $\rho^S \otimes \tau^E \in \text{BL}(\mathcal{H}^{SE})$ represents the joint system-environment state. On the other hand, there are elements $\rho^{SE} \in \text{BL}(\mathcal{H}^{SE})$ of the joint state space that *cannot* be written as a tensor product of individual constituent quantum states, *e.g.*, convex mixtures $\sum_i p_i \rho^S \otimes \tau^E$ of product states, which are known as *separable* states; or those states that are not of separable form, known as *entangled* states. In either case, given knowledge of the joint state of some composite system, ρ^{SE} , we can deduce a description of the subsystems that adequately describes the local properties of each subsystem, *i.e.*, provides the correct statistics for any local measurement, via the partial trace operation

$$\rho^S = \text{tr}_E [\rho^{SE}] \in \text{BL}(\mathcal{H}^S) \quad \text{and} \quad \rho^E = \text{tr}_S [\rho^{SE}] \in \text{BL}(\mathcal{H}^E). \quad (2.26)$$

The open dynamics framework considers the entire system-environment to evolve according to Eq. (2.23), and the dynamics of the system alone is deduced by taking the partial trace over the environment

$$\frac{\partial \rho_t^S}{\partial t} = -i \text{tr}_E [H^{SE}, \rho_t^{SE}], \quad (2.27)$$

which admits the solution

$$\rho_t^S = \text{tr}_E [\mathcal{U}_{t:0}^{SE} \rho_0^{SE}] = \text{tr}_E [U_{t:0}^{SE} \rho_0^{SE} U_{t:0}^{SE\dagger}]. \quad (2.28)$$

In contradistinction to Eq. (2.24), this is not a closed-form equation for the system state, as it depends on the dynamics of the joint state.

In addition to the evolution of quantum states, an important concept in quantum theory is that of *measurement*. A measurement on a quantum system is described by

a collection of operators, $\mathcal{J} = \{K^{(x)}\} \in \text{BL}(\mathcal{H}^S)$, such that $\sum_x K^{(x)\dagger} K^{(x)} = \mathbb{1}^S$. In anticipation of later developments, we refer to the entire collection as a *measurement instrument*: the index x refers to the possible measurement outcomes that can be observed given that an experimenter interrogates the system with the instrument \mathcal{J} . If the pre-measurement state of the system is ρ , then the probability that outcome x occurs is given by

$$\mathbb{P}(x|\mathcal{J}) = \text{tr} \left[K^{(x)\dagger} K^{(x)} \rho \right]. \quad (2.29)$$

Upon recording outcome x when using the instrument \mathcal{J} to interrogate the system, its state undergoes the transformation

$$\rho \xrightarrow{x} K^{(x)} \rho K^{(x)\dagger}. \quad (2.30)$$

As a special case of this general notion of quantum measurement is when the post-measurement state is not of interest, but merely the statistics associated to measurement outcomes are. In this case, it is sufficient to consider only collections of operators defined via $\Pi^{(x)\text{T}} := K^{(x)\dagger} K^{(x)} \in \text{BL}(\mathcal{H}^S)$, where the transpose $(\bullet)^{\text{T}}$ is added to the standard definition to better align with later notation. By definition, such a collection satisfies $\sum_x \Pi^{(x)\text{T}} = \mathbb{1}$ and contains only Hermitian, positive semidefinite operators, leading to the standard *Born rule*

$$\mathbb{P}(x|\mathcal{J}) = \text{tr} \left[\Pi^{(x)\text{T}} \rho \right]. \quad (2.31)$$

Overall, any such set $\{\Pi^{(x)}\}$ is known as a *positive-operator valued measure* (**POVM**), with each constituent operator referred to as a **POVM-element**. The linear functional induced on quantum states to yield probabilities via the Born rule, *i. e.*, $\mathcal{E}^{(x)}(\bullet) := \text{tr} \left[\Pi^{(x)\text{T}} \bullet \right]$, is known as an *effect*. A **POVM** is a special case of a measurement instrument that contains the necessary information to determine the observed statistics of any measurement on a normalised quantum state, but is insufficient to deduce the post-measurement state. Another special case considers projective measurements, where each operator associated to a measurement instrument is a projector that is orthogonal to every other, *i. e.*, $\{\Pi^{(x)}\}$ such that $\Pi^{(x)} = \Pi^{(x)\dagger}$ and $\Pi^{(x)} \Pi^{(x')} = \delta_{xx'} \Pi^{(x)}$.

2.2.2 Dynamical Maps

To reiterate an earlier point, what we often desire when studying a physical evolution is simply a description of how to map quantum states from one point in time to another. In the open dynamics framework, where we assume only access to the system, we would like such a map to act only on the space of the system, in analogy to the action of the unitary map for closed dynamics in Eq. (2.24). However, due to potentially dissipative system-environment interactions in the open setting, such a map is no longer restricted to being unitary. We first aim to understand the properties such a map must have in order to represent a physical evolution within quantum theory from an abstract perspective, before providing a connection to the standard open systems framework.

In short, we aim to understand the physically allowable quantum transformations between two points in time, *i. e.*, the set of maps $\mathcal{C} : \text{BL}(\mathcal{H}^S) \rightarrow \text{BL}(\mathcal{H}^S)$ that take an arbitrary initial quantum state $\rho \in \text{BL}(\mathcal{H}^S)$ to a valid output quantum state $\sigma \in \text{BL}(\mathcal{H}^S)$, schematically depicted in Fig. 2.7, as follows⁴

$$\sigma = \mathcal{C}(\rho). \quad (2.32)$$



Figure 2.7: *Dynamical map.* A dynamical map, \mathcal{C} (yellow), takes input states ρ to output states σ (green). Any such transformation in quantum theory must be linear, completely-positive and trace-preserving.

In the classical setting, the requirement that a map takes arbitrary input probability distributions to valid output distributions serves to constrain their structure to be stochastic maps. Similarly, the analogous demand in quantum theory imposes structural constraints on the allowable maps. Any meaningful transformation must preserve the key properties of the density operator, *i. e.*, it must preserve trace, Hermiticity and positivity, in order to ensure that σ indeed represents a valid quantum state. In addition, its action must preserve convex mixtures of states

$$\mathcal{C}\left(\sum_i p_i \rho^{(i)}\right) = \sum_i p_i \mathcal{C}\left(\rho^{(i)}\right) = \sum_i p_i \sigma^{(i)}. \quad (2.33)$$

This is not a requirement that stems from the linearity of quantum mechanics; rather, it follows from the *linearity of mixing* principle that must be satisfied by any statistical theory. The importance of this principle can be highlighted through the following

⁴ The present restriction of our attention to maps taking states of some input system to those of *the same* system, *i. e.*, living on the same Hilbert space, as its output is by no means necessary and will be relinquished in the forthcoming chapter.

Gedankenexperiment: suppose Alice prepares a quantum system in either state ρ_1 or ρ_2 , which she sends to Bob, with the transmission represented by the map \mathcal{C} . Bob then performs state tomography [8, 83] to determine which state Alice sent him, with the entire protocol repeated many times. Suppose first that Alice only sends ρ_1 on Monday and ρ_2 on Tuesday. Bob will then conclude that he receives $\sigma_1 = \mathcal{C}(\rho_1)$ on Monday and $\sigma_2 = \mathcal{C}(\rho_2)$ on Tuesday. Now suppose instead that Alice sends these two states at random, with probabilities p and $1 - p$ respectively. Without knowledge of which state was sent in each run, Bob would conclude that he receives the average state $\sigma = \mathcal{C}(\rho)$, where Alice's preparation corresponds to the average state $\rho = p\rho_1 + (1 - p)\rho_2$. Consider now the scenario where Alice reveals to Bob which state she sent in which run; surely now Bob concludes that he received the states σ_1 or σ_2 whenever Alice sent him ρ_1 or ρ_2 respectively. On the other hand, averaging over the trials would amount to Bob receiving σ . Thus, it must be the case that the transformation map acts linearly: $\sigma = p\sigma_1 + (1 - p)\sigma_2 = p\mathcal{C}(\rho_1) + (1 - p)\mathcal{C}(\rho_2)$.

In analogy to the classical case, trace-preservation of the density operator corresponds straightforwardly to the classical notion that the output of a stochastic map is normalised; likewise, Hermiticity and positivity preservation echo the demand that stochastic maps take probability distributions to probability distributions. However, a distinct departure in the quantum setting arises due to the existence of entangled states within the theory, which enforces us to strengthen the notion of positivity, which means that the output state of a map is always positive semidefinite, *i. e.*, $\sigma = \mathcal{C}(\rho) \geq 0 \forall \rho$, to the stricter one of *complete-positivity*.

Consider the situation where the initial state of some bipartite system $\rho^{AB} \in \text{BL}(\mathcal{H}^{AB})$ is represented by an entangled density operator

$$\rho^{AB} \neq \sum_i p_i \alpha^{(i)} \otimes \beta^{(i)}, \quad (2.34)$$

where $\{\alpha^{(i)}\} \in \text{BL}(\mathcal{H}^A)$ and $\{\beta^{(i)}\} \in \text{BL}(\mathcal{H}^B)$ are density operators describing the states of the subsystems A and B respectively and $\{p_i\}$ are probabilities. Suppose these subsystems are sufficiently well-separated and undergo separate evolutions; one can easily construct examples for which the post-evolution joint state is not a positive semidefinite operator. For example, if A undergoes a trivial evolution $\mathcal{I}^A(\bullet)$ whilst B is subject to the transposition map, *i. e.*, and $\mathcal{T}^B(\bullet) := (\bullet)^T$. Both of these maps are positive, *i. e.*, $\mathcal{I}^A(\rho^A) \geq 0 \forall \rho^A \in \text{BL}(\mathcal{H}^A)$ and $\mathcal{T}^B(\rho^B) \geq 0 \forall \rho^B \in \text{BL}(\mathcal{H}^B)$. However, when applied jointly to an entangled state $\rho^{AB} \in \text{BL}(\mathcal{H}^{AB})$, a negative operator, representing no physical state, can result $\mathcal{I}^A \otimes \mathcal{T}^B(\rho^{AB}) < 0$.

The natural condition required to describe meaningful quantum evolutions of composite systems is that their local implementation leads to a valid joint output state in the presence of any innocuous ancillary system (see Fig. 2.8). Formally, when a map acting on some system, \mathcal{C}^S , is extended to act trivially on an arbitrarily-sized ancillary space, \mathcal{H}^R , via the identity map \mathcal{I}_d^R acting on d -dimensional quantum states, its action on an arbitrary joint state $\rho^{SR} \in \text{BL}(\mathcal{H}^{SR})$ must lead to a positive semidefinite output

$$\sigma^{SR} = \mathcal{C}^S \otimes \mathcal{I}_d^R(\rho^{SR}) \geq 0 \quad \forall d \in \mathbb{N}. \quad (2.35)$$

We refer to maps satisfying the above condition (which also implies Hermiticity preservation [35]) as *completely-positive* (CP) maps.⁵

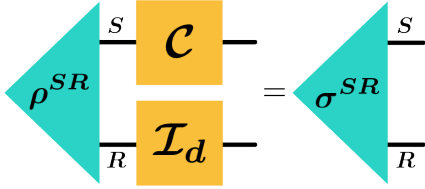


Figure 2.8: *Complete-positivity.* To represent a valid evolution in quantum theory, a dynamical map \mathcal{C} must be completely-positive. Consider a joint system, SR , with R of dimension d , in an arbitrary state ρ^{SR} . Complete-positivity of \mathcal{C} means that when it is implemented locally on S , whilst R evolves trivially under \mathcal{I}_d , the joint output is a valid quantum state $\sigma^{SR} \in \text{BL}(\mathcal{H}^{SR})$, for all $d \in \mathbb{N}$.

Lastly, the trace-preservation requirement, stemming from the fact that the probabilities of any measurement on the output state sum to one, is simply $\text{tr}[\sigma] = \text{tr}[\mathcal{C}(\rho)] = \text{tr}[\rho] \quad \forall \rho \in \text{BL}(\mathcal{H}^S)$. An alternative way to consider trace-preservation is that the transformation \mathcal{C} taking ρ to σ occurs overall with certainty. Indeed, one can envisage situations where a process only occurs conditionally with some probability, such as that on a quantum state upon recording a measurement outcome, in which case the trace-preservation condition must be slightly modified to capture such scenarios (although the dynamics must remain CP). We will focus on such maps in detail within the broader framework introduced in the coming chapter.

To summarise, the most general form of maps describing overall deterministic transformations in quantum theory are those that are both CP and *trace-preserving* (TP). Due to their importance in the study of open dynamics, these maps are referred to synonymously as CPTP maps, *dynamical maps* or *quantum channels* [6, 80]. Since the set of density operators is convex, and CPTP maps act linearly, the space of CPTP maps is also a convex set. So much for the mathematical properties of the allowable maps taking input states to output states in open quantum theory. We now make the connection with an underlying joint system-environment evolution.

⁵ We note that although maps that are positive but not completely-positive, such as the partial transpose considered above, are useful for witnessing entanglement in quantum states [84, 85], for the purposes of describing quantum evolution in a meaningful way, we require the associated map to be CP.

A major result concerning the axiomatic considerations above and the vantage point of open quantum systems dynamics is Stinespring's dilation theorem [86], which guarantees that any CPTP map on the system can be thought of as arising from some underlying unitary dynamics of the system with some environment. That is, we can represent any CPTP map $\mathcal{C}^S : \text{BL}(\mathcal{H}^S) \rightarrow \text{BL}(\mathcal{H}^S)$ in terms of a dilated picture of unitary dynamics in \mathcal{H}^{SE} , with a fiducial initial environment state $\tau^E \in \text{BL}(\mathcal{H}^E)$ and the environment being ignored following the evolution, depicted in Fig. 2.9, as

$$\sigma^S = \mathcal{C}^S(\rho^S) = \text{tr}_E \left[\mathcal{U}^{SE} \rho^S \otimes \tau^E \right]. \quad (2.36)$$

The dilation for a quantum channel is non-unique: many configurations of $\{\mathcal{U}^{SE}, \tau^E\}$ can give rise to the same dynamical map \mathcal{C}^S . Nonetheless, a dilation can always be found for an environment with dimension d_E^2 [86]. It is often quipped that this correspondence allows us to always *go to the Church of the Larger Hilbert Space*, considering any overall deterministic but irreversible (*i. e.*, non-unitary) transformation of the density operator of the system to be the manifestation of our subjective ignorance of part of a reversible (unitary) transformation in the larger joint Hilbert space.

Stinespring's theorem tells us that any CPTP map can always be dilated to a fixed (though non-unique) quantum circuit comprising a joint unitary transformation through which the system interacts with a fiducial environment state, the latter of which is finally discarded. On the other hand, a CP map can always be thought of as arising similarly through a joint unitary interaction with some initial environment state, although in this case, rather than tracing over the degrees of freedom of the environment following the evolution, one hypothetically performs a measurement and post-selects on an outcome [28]. Since such a procedure can only occur probabilistically, there can be no such *fixed* underlying dilation model attributed to the map, in contrast with the case

for CPTP maps; thus, the property of trace-preservation may be interpreted as one of an *overall deterministic* implementation of the transformation.

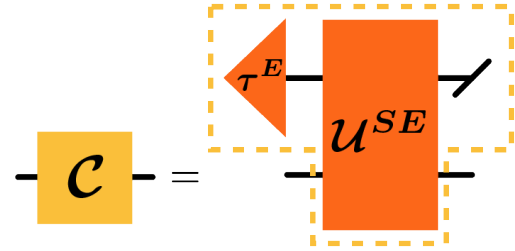


Figure 2.9: *Stinespring dilation of a quantum channel.* Any CPTP map \mathcal{C} can be dilated in terms of a unitary interaction \mathcal{U}^{SE} of the system with some fiducial environment state τ^E , which is finally disregarded (indicated by the diagonal slash). A dilation is non-unique: the channel contains all information enclosed in the yellow border on the right, but individual contributions of the environment state and the joint unitary (orange) cannot be delineated.

2.2.3 GKSL Equation

The above discussion culminates in the most general structure of a quantum evolution taking density operators to density operators. We can bring time back into the picture to see how a continuous-time evolution can naturally lead to [CPTP](#) dynamics on the system level. Although the exact evolution of the system is governed by its microscopic interactions with the environment, we can invoke the *approximation* that the evolution of the system only depends on its current state via some linear map, \mathcal{L}_t , and write down the following equation of motion for the system density operator [\[7\]](#)

$$\frac{\partial \rho_t^S}{\partial t} = \mathcal{L}_t(\rho_t^S). \quad (2.37)$$

When the generator is time independent, the formal solution to this equation constitutes a single-parameter semi-group, *i. e.*, a family of norm-continuous maps $\mathcal{C}_t^S = \exp(t\mathcal{L})$ satisfying $\mathcal{C}_{t+s}^S = \mathcal{C}_t^S \mathcal{C}_s^S$.⁶ This allows us to express the system state at arbitrary time t in terms of a map acting on an arbitrary initial state via

$$\rho_t^S = \mathcal{C}_{t:0}^S(\rho_0^S). \quad (2.38)$$

The family of dynamical maps $\mathcal{C}_{t:0}^S$ naturally inherit the properties of linearity, complete-positivity and trace-preservation. Furthermore, the semi-group property leads to the divisibility of the dynamical maps, *i. e.*, they satisfy $\mathcal{C}_{t:0}^S = \mathcal{C}_{t:s}^S \mathcal{C}_{s:0}^S \ \forall \ t \geq s \geq 0$. Intuitively, this means that we can calculate the system density operator at any time via a successive composition of previous dynamical maps. We will return to a deeper discussion of this concept in the coming section.

A breakthrough result for the field of open quantum systems was presented independently by *Gorini, Kossakowski and Sudarshan* [\[87\]](#) and *Lindblad* [\[88\]](#) ([GKSL](#))⁷, who provided the most general form of the generator \mathcal{L}_t in Eq. (2.37) such that the resulting solution forms a divisible [CPTP](#) semi-group, yielding the [GKSL](#) equation

$$\frac{\partial \rho_t}{\partial t} = -i[H, \rho_t] + \sum_i \gamma_i \left(L_i \rho_t L_i^\dagger - \frac{1}{2} \{ L_i L_i^\dagger, \rho_t \} \right). \quad (2.39)$$

Here, everything is understood to act on the space of the system only. Note that the anti-commutator is represented by $\{A, B\} := AB + BA$, the H represents a Hamiltonian (Hermitian) contribution to the system evolution, the $\{L_i\}$ are known as Lindblad operators and the $\{\gamma_i\} \geq 0$ are non-negative rates.

⁶ For a time dependent L_t , the solution would involve some time-ordered exponential and time integral, and would depend on two time parameters.

⁷ This equation has also been independently discovered by Franke in 1976 [\[89\]](#), although his name has not made it into the famous acronym yet.

To reiterate, the **GKSL** equation is the quantum generalisation of the classical **ME** (see Appendix B.1): the solution of the latter is a **P**-divisible family of stochastic maps, whereas that of the former is a divisible family of **CPTP** maps taking the density operator from one point in time to any other later time in a physically acceptable manner, as per Eq. (2.38) [87, 88]. Moreover, one can arrive directly at the form of the **GKSL** equation when starting from a system-environment model and invoking the Born-Markov and secular/rotating wave approximations [7, 82].

2.2.4 Tomographic Reconstruction of Quantum Channels

Recall that in the open systems framework, we assume that we cannot probe the state of the environment and do not know the joint unitary transformations. We now consider how one can experimentally reconstruct a description of the dynamical map for the evolution between pairs of timesteps by probing the system alone. By considering a fixed environment state in Eq. (2.36), the linear dynamical map is induced on the state space of the system. This property of linearity importantly permits a procedure to reconstruct quantum channels, known as *quantum process tomography* [8, 83, 90].

To understand process tomography, it is useful to first briefly consider quantum *state tomography*, where an unknown quantum state is inferred through measurement statistics. Given a state $\rho \in \text{BL}(\mathcal{H}^S)$, one begins by choosing a **POVM** that spans $\text{BL}(\mathcal{H}^S)$. Such a set of operators is called *informationally-complete* (**IC**) and necessarily contains at least d^2 **POVM** elements. By recording the probability for each measurement outcome, the quantum state can be uniquely reconstructed due to linearity [8, 83].

In direct analogy to state tomography, the protocol for quantum process tomography, depicted graphically in Fig. 2.10, is as follows [90]: **i)** take a basis of input states $\{\hat{\rho}^{(i)}\}_{i=1}^{d^2}$ that span the operator space of the system;⁸ **ii)** each of these input states are sent through the process; **iii)** the corresponding output states $\{\hat{\sigma}^{(i)}\}_{i=1}^{d^2}$ are determined via state tomography; and **iv)** the input-output relations deduced can be linearly inverted to uniquely determine the dynamical map.

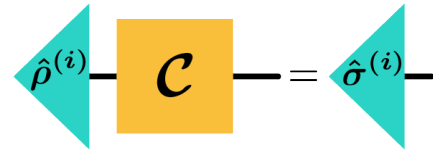


Figure 2.10: *Tomographic reconstruction of a quantum channel.* A quantum channel \mathcal{C} can be reconstructed by determining the output states $\{\hat{\sigma}^{(i)}\}$ corresponding to an **IC** set of inputs $\{\hat{\rho}^{(i)}\}$.

⁸ Here, we use the caret notation to indicate that the object belongs to a fixed basis, which is not necessarily normalised or orthogonal.

Mathematically, we make use of the natural inner product on $\text{BL}(\mathcal{H}^S)$, *i. e.*, for any elements $\mu, \nu \in \text{BL}(\mathcal{H}^S)$, we have the Hilbert-Schmidt inner product $(\mu, \nu) := \text{tr} [\mu^\dagger \nu]$. We can always construct a basis of $\text{BL}(\mathcal{H}^S)$ with a set of d^2 density operators. For any such basis, there exists a dual set of objects, $\{\hat{D}^{(i)}\}_{i=1}^{d^2}$, such that [51, 74]

$$\text{tr} [\hat{D}^{(i)\dagger} \hat{\rho}^{(j)}] = \delta_{ij} \quad \forall i, j. \quad (2.40)$$

We provide an explicit construction of the dual set to an arbitrary basis in Appendix B.2.

By linearity, determination of the output states for a basis of input states uniquely specifies the map. Thus the action of the CPTP map \mathcal{C}^S on an arbitrary input state $\rho \in \text{BL}(\mathcal{H}^S)$ can be linearly extended and expressed as

$$\mathcal{C}^S(\rho) = \sum_{i=1}^{d^2} \hat{\sigma}^{(i)} \text{tr} [\hat{D}^{(i)\dagger} \rho] \quad (2.41)$$

The elements of the dual basis to a basis of density operators are elements of $\text{BL}(\mathcal{H}^S)$, but not necessarily density operators, and the overall construction above is guaranteed to yield a positive semidefinite output.

Lastly, note that in this procedure it is assumed that the initial state of SE is uncorrelated, allowing us to treat the reduced initial state of the environment τ^E as a fixed constant of the problem. This property means that the dynamics is independent of the initial state of the system, allowing us to vary the inputs to the channel freely, as is a tenet of the process tomography protocol. For instance, GKSL dynamics allows for this clear delineation between the instantaneous state of the system and the subsequent dynamics at all times (see Eq. (2.37)), therefore leading to a CP-divisible description.

2.3 A PROBLEM OF FORMALISM

So far, we have explored the open quantum systems formalism and the dynamical map in describing the evolution of a quantum state to one at a later time. The usefulness of the dynamical map picture comes to the fore when considering its ability to be tomographically reconstructed in a finite number of experiments, providing an unambiguous description of the process in terms of accessible quantities. However, the dynamical map description does not accommodate for intermediary interventions on the level of the system. In order to understand memory effects in quantum stochastic processes, we would like to develop a similar operational framework that actively accounts for multi-time correlations. As we shall see throughout this section, a number of subtleties arise when attempting to do so.

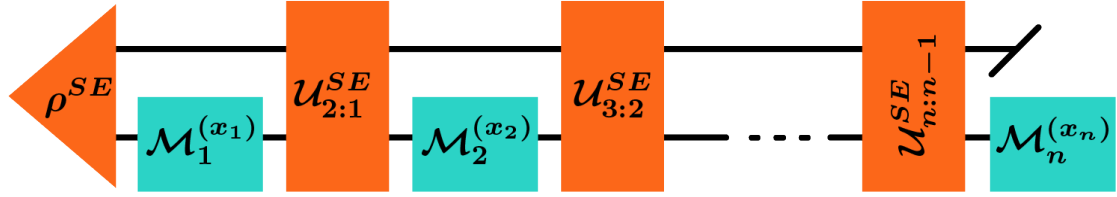


Figure 2.11: *Dilation of a quantum process interrogated in time.* A joint system-environment state ρ^{SE} is interrogated in time by measurements $\mathcal{M}_j^{(x_j)}$ interspersed throughout periods of joint unitary evolution $\mathcal{U}_{2:1}^{SE}$.

Consider an experimenter making measurements on some system over a set of timesteps, Λ_n , described by the measurement operators $\{M_j^{(x_j)}\}$ interspersed throughout periods of joint unitary dynamics, with $t_j \in \Lambda_n$ denoting the timestep of each measurement. The conditionally realised transformation of the system associated to observing the outcome x_j is described by the CP map $\mathcal{M}_j^{(x_j)}(\rho^{SE}) := M_j^{(x_j)} \rho_j^{SE} M_j^{(x_j)\dagger}$, where the measurement operators act on the system alone and identity operators on the environment are implied. The joint probability distribution of the statistics observed given a sequence of measurements applied over time is depicted in Fig. 2.11 and written

$$\mathbb{P}_{n:1}(x_n, \dots, x_1) = \text{tr} \left[\mathcal{M}_n^{(x_n)} \mathcal{U}_{n:n-1}^{SE} \dots \mathcal{M}_2^{(x_2)} \mathcal{U}_{2:1}^{SE} \mathcal{M}_1^{(x_1)} \rho^{SE} \right]. \quad (2.42)$$

To reiterate, here, the unitary maps $\mathcal{U}_{j:j-1}^{SE}$ are understood to act on the joint system-environment space, while the CP maps $\mathcal{M}_j^{(x_j)}$ act only on the system, and we have chosen not to include nested parentheses to avoid notational clutter, as we will often do throughout this thesis, with the understanding that all maps act on everything to the right of them.

As exemplified previously through the Stern-Gerlach example at the beginning of Section 2.2, a major problem of this description is that the resulting joint distributions do not satisfy the containment property of the KET. Logically, from a proper description of the dynamics over multiple timesteps, we expect to be able to deduce the correct description of the dynamics on any subset of timesteps. Its breakdown arises due to the necessarily invasive nature of measurements in quantum mechanics [14]: in contrast to classical theory, here, choosing not to interrogate the system is different from averaging over all possible measurement outcomes.

An additional point of concern is that for a system evolving in an open fashion, performing a measurement on the system can condition the state of the environment. This is not a purely quantum mechanical feature, but rather a symptom of stochastic processes with memory. For example, consider applying a rank-1 projective measurement

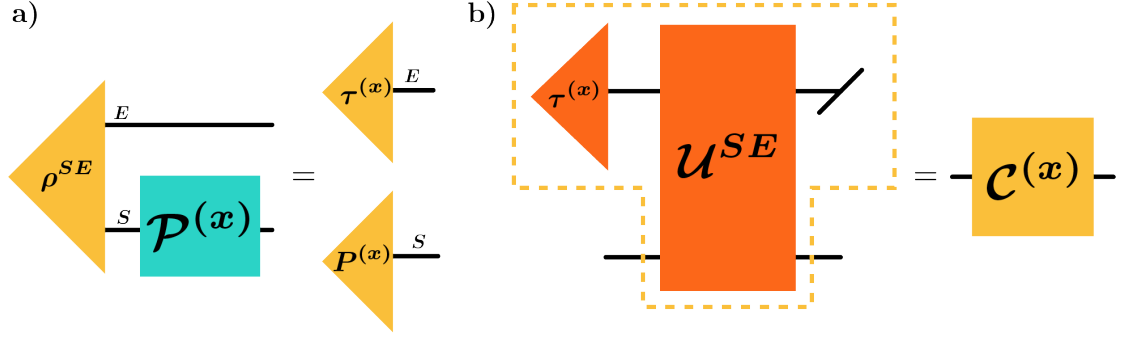


Figure 2.12: *Measuring the system conditions the environment and influences future dynamics.* In panel a), the S subsystem of an initially correlated system-environment state ρ^{SE} is subject to a projective measurement, which has the effect of conditioning the environment into some state $\tau^{(x)}$. In panel b), we consider the scenario where some system-environment dynamics \mathcal{U}^{SE} subsequently occurs. As the environment state that the system goes on to interact with is conditioned by the measurement outcome (rather than being some fixed, fiducial state), a *set of many different* dynamical maps $\{\mathcal{C}^{(x)}\}$ that describe the evolution is induced, one for each measurement outcome. Each of these maps has a non-deterministically occurring environment state in its dilation, represented by the orange objects enclosed by the yellow dashed line.

$\mathcal{P}^{(x)}(\bullet) := P^{(x)} \bullet P^{(x)}$ on the subsystem S of a correlated SE state, as shown in Fig. 2.12. Upon recording outcome x , the joint state maps to

$$\rho^{SE} \xrightarrow{x} P^{(x)} \otimes \tau^{(x)}, \quad (2.43)$$

where $\tau^{(x)} = \text{tr}_S [P^{(x)} \rho^{SE}] \in \text{BL}(\mathcal{H}^E)$. Although the post-measurement state is uncorrelated for each measurement outcome—it is a tensor product state—the post-measurement environment state is correlated with the outcome observed; as such, the future evolution of the system can depend upon knowledge of previous outcomes.

The points raised above have led many to the conclusion that an “intrinsic characterization and quantification of memory effects in the dynamics of open quantum systems... has to be based solely on the properties of the dynamics of the open system’s *density matrix*” [23]. In the coming chapter we defy this statement, highlighting that it is a problem of *formalism* that must be overcome to obtain a fully fledged description of quantum stochastic processes, rather than a fundamental issue. To put the cart before the horse, there exist operational frameworks that can account for multi-time correlations, superseding traditional open systems frameworks and subsequently allowing for a proper description of memory effects. We will first examine in detail how criteria based on two-point descriptions of quantum processes fail to accurately represent the underlying process when memory effects are present, motivating the fact that we are forced to go beyond such a paradigm to characterise generic quantum stochastic processes.

2.3.1 A Hierarchy of Notions of Non-Markovianity

Due to the breakdown of the [KET](#) on the level of observed probabilities in quantum mechanics—and therefore the lack of a well-defined notion of conditional statistics—the concept of Markovianity in the quantum realm has been mainly studied throughout the open systems community in terms of the time-evolution of the density operator, or, more operationally, the dynamical maps describing the evolution between pairs of timesteps. Of course memory effects display signatures that can be gleaned from such two-point considerations, giving rise to a myriad of non-Markovianity *witnesses*. The usefulness of such approaches is not to be underestimated: from a practical perspective, they often provide easy-to-check criteria that purport to verify the presence of memory effects. Prominent examples include those based on: the divisibility of the dynamics [66, 91]; the monotonically non-increasing nature of the distinguishability of quantum states subject to the evolution [92]; the detection of initial correlations [74, 93–98]; the positivity of the dynamical maps [99–101]; changes to quantum correlations or coherence [102, 103]; changes to the Fisher information [104]; and channel capacities and information backflow [105–108]; to name but a few. See Refs. [23, 76] for a thorough overview of these various concepts and their relations.

The main problems with such approaches are that they: **i)** lack a clear operational interpretation; **ii)** do not coincide with Markovianity in the classical limit; and **iii)** do not agree on the characterisation of whether a given process is Markovian or not. Consequently, different introduced “measures” of non-Markovianity disagree on both the degree of non-Markovianity and whether or not memory effects are present at all [49, 67, 106, 109, 110].

For instance, one of the most widely used criteria is based on the notion of [CP-divisibility](#) [66], which intuitively means that the open dynamics can be broken into a piecewise composition of dynamical maps satisfying⁹

$$\mathcal{C}_{k:i} = \mathcal{C}_{k:j} \mathcal{C}_{j:i} \quad \forall t_k > t_j > t_i. \quad (2.44)$$

The property of [CP-divisibility](#) implies some of the other aforementioned concepts, *e. g.*, the non-increasing distinguishability criteria is a direct consequence of the contractivity of the trace-distance under [CPTP](#) maps, although the converse does not hold [66]. Consequently, there are many processes that the former characterisation would deem to be non-Markovian, but the latter would disagree, *e. g.*, see Refs. [109, 110].

⁹ Note that [CP-divisibility](#) obviously implies [P-divisibility](#).

Moreover, considering how an experimenter might test for CP-divisibility in practice illuminates that there are (at least) two inequivalent definitions of divisibility [54]. This is because of the difficulties that arise when attempting to tomographically reconstruct the maps $\mathcal{C}_{k:j}$. By time t_j , the system will generally be correlated with its environment, hindering one's ability to vary the input state to the map independently of the environment and therefore breaking a crucial requirement of quantum process tomography. To circumvent this problem, under the assumption that no initial correlations are present at t_i so that the maps $\mathcal{C}_{k:i}$ and $\mathcal{C}_{k:j}$ can be reconstructed and that $\mathcal{C}_{j:i}$ is invertible, one may compute an artificial map $\mathcal{A}_{k:j} = \mathcal{C}_{k:i}\mathcal{C}_{j:i}^{-1}$ and test if it is CPTP.

Notwithstanding the fact that neither of these assumptions might be satisfied, it is unclear operationally what dynamics the inverse map $\mathcal{C}_{j:i}^{-1}$ and therefore the derived map $\mathcal{A}_{k:j}$ actually represents. An operational way around this is to simply discard any possible system-environment correlations present at t_j by, *e. g.*, making a measurement and preparing the output state in a fixed state that is independent of measurement outcomes, which are then averaged over. This allows for the experimental reconstruction of $\mathcal{C}_{k:j}$ which is guaranteed to be CPTP, and one must simply check whether Eq. (2.44) holds. Within the set of processes where the dynamical maps in the former construction are invertible, the latter operational criteria is stricter; in either case, just as in the classical setting, CP-divisibility does not imply Markovianity [27, 54].

Additional examples of a similar flavour can be found in the Supplemental Material of Ref. [49]: in particular, an example where the trace-distance distinguishability criteria would deem the process Markovian, and another in which no system-environment correlations are ever built up in the process, which is also often considered a hallmark of Markovianity [76]; however, both of these examples can display memory effects. In summary, each of the criteria proposed at the beginning of this section are based upon an inadequate description of processes with memory, giving rise to the aforementioned inconsistencies and leading some to the conclusion that there can be no unique condition for Markovianity for quantum stochastic processes. This is not true: the problem we must overcome is one of formalism, as we now discuss.

2.3.2 *Limitations of Traditional Approaches*

The reason that the traditional approaches considered above fail is because Markovianity is, at its core, a statement concerning multi-time conditional statistics: in its truest sense, determining whether a process is Markovian requires testing for conditional in-

dependence between the current statistics measured from those deduced at all earlier times. This demands an exponentially large set of conditions to be satisfied, even in the classical case, *i. e.*, Eq. (2.11) must hold for all x_1, \dots, x_n . Despite the added complication of the incompatibility of measurement statistics in quantum mechanics described above, the fact that any two-point description of dynamics cannot suffice to characterise Markovianity is more of a logical statement: it simply cannot be used to test all conditions, and therefore is inadequate at describing the complete story. In short, a proper treatment of non-Markovian processes *cannot* be a simple extension of the tools used to describe Markovian ones [25].

The fact that the dynamical map formalism fails to capture multi-time memory effects highlights its major shortcoming when it comes to describing quantum processes with memory. Moreover, the dynamical map description cannot even accurately address all two-point dynamics allowable in quantum theory; for instance, it is well-known that it fails to describe open dynamics in the presence of initial system-environment correlations [65, 73–75, 111–114]. As discussed earlier, the traditional treatment and reconstruction of open quantum system dynamics assumes an initially uncorrelated system-environment state. This hypothesis assigns a peculiar role to the initial time; even if it happens to be true at some time, which could be taken as the initial time leading to subsequent CPTP dynamics to any later time, the system will, at this later time, generally be correlated to its environment and we immediately face the problem of describing the dynamics beyond that later timestep. Thus, the proper description of quantum processes across multiple timesteps inherently concerns the initial correlation problem, which must be overcome as a starting point.

This problem came to the fore due to technological advances that allowed experimentalists to begin tomographically reconstructing quantum logic gates by means of process tomography [115–119]. Although the gates implemented were expected to be non-ideal, they were expected to be CPTP. However, to the surprise of many, this was not necessarily the case: the dynamical maps constructed were not CP. A notable theoretical explanation for the construction of non-CP quantum gates was immediately put forth by Pechukas, who realised that in the presence of initial system-environment correlations, the dynamical map formalism suggests that the subsequent dynamics of the system need not be CP [111]. This means that a density operator describing a quantum state can evolve, in the presence of some ancillary system, into a non-positive operator, whose physical interpretation is unclear. Fundamentally, this notion is at odds with the fact that we always observe positive probabilities in any experiment. Furthermore, relinquish-

ing the requirement of **CP** means sacrificing many physically important principles, such as the Holevo bound [120], data processing inequality [121], and the entropy production inequality [122]. An alternate approach is to give up the requirement of linearity [113], which also proves problematic: complete tomography is no longer possible in general when the dynamics is non-linear [51, 123]; the data processing inequality is also violated [8]; and numerous problematic implications have been shown to arise (see, *e. g.*, Refs. [124–126]). On the other hand, at first glance, it seems as if we are in a double-bind: forked between sacrificing either **CP** or linearity for a consistent description of the dynamics [114].

Despite their original operational motivations, dynamical map descriptions do not take active interventions into account, besides the limited scenario where the initial system state is uncorrelated from its environment and active preparation procedures can be enacted *without* influencing the subsequent dynamics. When initial system-environment correlations are present, on the other hand, one cannot probe the system without also affecting the environment (see Eq. (2.43)). Since the environment can be conditioned by an operation on the system and then can feed forward to play a role in the subsequent evolution (see Fig. 2.12 b)), any tomographically reconstructed description of the dynamical map will depend upon the interventions applied to the system, seemingly implying the lack of a process that exists independent of the experimenter. This is why, for instance, even non-Markovian generalisations of **MEs** with memory kernels, which are useful for simulating processes with memory [24], are insufficient to characterise them: unless they allow for active interventions on the system, their operational consequences are unclear. As we shall soon see, this blurriness between the the dynamics governed by the process and the transformations applied by a probing experimenter is directly related to the breakdown of the **KET** in quantum theory.

To summarise, on the one hand, the lack of consistency conditions on the level of probability distributions seemingly imply that there may not exist a unique, fixed process giving rise to the statistics observed. This renders any notion of conditioning, as is

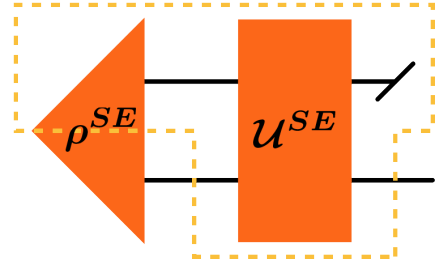


Figure 2.13: *Initial correlation problem.* The dynamical map formalism assumes the ability to consider the environment as some fixed state. This is not the case when initial correlations are present, since we cannot separate system states that are input to the subsequent dynamics from the dynamics themselves. This artefact is depicted by the yellow dashed line which (wrongly) attempts to ‘cut’ the initially correlated state.

required to characterise Markovianity and Markov order more generally, nonsensical. On the other hand, any framework that does not properly account for active interventions on the system, such as those based on dynamical maps, are necessarily inadequate to describe processes with memory. In order to generalise Markov order to the quantum realm, we require a reasonable picture of multi-time correlations. We stress, finally, that in any such case, these problems arise due to inadequacy of *formalism* rather than fundamental physical truths: a proper description of a stochastic process should ameliorate these aforementioned issues. To correctly describe the statistics observed for sequences of measurement outcomes, thereby capturing multi-time memory effects, we are forced to go beyond such traditional approaches and develop a framework that meaningfully accounts for active interventions, as we consider in the coming chapter.

QUANTUM STOCHASTIC PROCESSES

THROUGHOUT THE PREVIOUS CHAPTER, we explored in detail the nature of both classical and quantum stochastic processes, highlighting some key difficulties that arise when attempting to characterise memory in the quantum case. Here, we introduce a general formalism that accommodates an unambiguous study of processes with memory by making explicit the role of the *experimenter*.¹ By separating the underlying, uncontrollable system-environment dynamics of the process at hand from the controllable interventions an experimenter might choose to apply on the level of the system, we come to a robust operational framework for describing stochastic processes that actively takes interventions into account, thereby solving the aforementioned issues.

One such framework that accounts for possible interventions across multiple timesteps is that of the *process tensor* [49, 50], which is the fundamental mathematical object of the operational formalism we employ throughout the remainder of this thesis to describe quantum stochastic processes. It provides a multi-linear mapping from sequences of controllable operations applied by an experimenter to the final output density operator of the evolution. Additionally, by way of a generalised spatio-temporal Born rule, the process tensor yields the correct joint statistics for any temporal operation sequence an experimenter might implement [34]. It thereby encapsulates all (multi-time) memory effects in the process, which are crucial for a proper treatment of non-Markovian processes; indeed, within the context of open system dynamics, the process tensor was developed specifically to generalise traditional approaches of the open systems formalism [49, 50], which are limited in scope to two-point correlations.

Before properly introducing this versatile framework, it is worth mentioning that similar formalisms have been developed within various other contexts, initially by Lind-

¹ Although we speak explicitly of an experimenter, we wish to be clear that they need not actually be there; the notion of an experimenter simply provides an artifice that lends itself nicely to mental imagery and language.

blad [26] and Accardi, Frigerio and Lewis [27], with more modern incarnations going under the guises of: *quantum combs* in the context of generic quantum circuit architectures [28–30], *causal automata* or *non-anticipatory quantum channels* to address memory effects in quantum processes [47, 48], *process matrices* to study the nature of causality in quantum foundations [17, 32, 33], *operator tensors* [39, 40] and *superdensity operators* [41] pertaining to the development of quantum mechanics in spacetime, *quantum strategies* for quantum game theory [42, 43], and *causal boxes* regarding secure communication protocols [44]. While the motivations and subtle details behind these notions vary slightly, the common thread is that they all separate the controllable from the uncontrollable influence on the system; making perspicuous that it is everything that is *out of control* of an experimenter that constitutes the *process* itself.

In short, these frameworks describe a quantum stochastic process as a *collection* of joint probability distributions over the outcomes of *any possible sequence of measurements*. For example, the dynamics of a spin- $\frac{1}{2}$ particle can be uniquely described by recording the probability for the spin to be found in alignment with any sequence of independent directions an experimenter might choose to measure (at the timesteps of interest). Once this data has been recorded, generalised Kolmogorov conditions hold, and a generalisation of the KET can be recovered for quantum (and more general) stochastic processes [14, 27], thereby unifying previous approaches to open system dynamics [51]. Crucially, the process tensor framework provides both an unambiguous definition of quantum stochastic processes and a suitable notion of marginalisation in quantum theory. Perhaps most importantly for our present purposes, the formalism permits the development of a set of necessary and sufficient conditions for a quantum process to be classified as Markovian [49, 50]. In this chapter, we introduce the process tensor formalism, its properties, and other tools necessary for the remainder of this thesis.

3.1 PROCESS TENSOR FRAMEWORK

The operational perspective to quantum mechanics embraces the philosophy perhaps best stated by Peres [127]: “The simple and obvious truth is that quantum phenomena do not occur in a Hilbert space. They occur in a laboratory. If you visit a real laboratory, you will never find there Hermitian operators. All you can see are emitters (lasers, ion guns, synchrotrons and the like) and detectors. The experimenter controls the emission process and observes detection events.” An experimenter has access—in principle—to everything that can be measured on the level of the system, and hence a proper description of the

process is one that reproduces the correct measurement statistics. We now take up this perspective and show it resolves the problem of describing multi-time quantum dynamics by way of the process tensor.

3.1.1 Multi-time Quantum Experiments

We consider the scenario in which an experimenter probes a quantum system that is evolving according to some dynamics that they wish to characterise. Importantly, the experimenter is presumed to have complete instantaneous control over the choice of operations that they implement on the system over a number of timesteps, but no control over the intermediary dynamics; for this reason, we refer to the setting as a *multi-time quantum experiment*. A conceptual schema that describes any such physical experiment depicts it as the composition of the following steps.

Definition 3.1 (Multi-time quantum experiment). A multi-time quantum experiment proceeds according to the following protocol:

1. The initially unknown state of a system is prepared into a known state (which could be statistical in nature, *i. e.*, not pure).
2. The system is subsequently subject to some physical evolution.
3. An experimenter has access to probe the system.
4. Steps 2 and 3 repeat a number of times, with the system finally being measured.

We can concretely relate this multi-time experiment scheme to a dilated system-environment picture, as depicted in Fig. 3.1. In general, the initially unknown quantum state in step (1) of Def. 3.1 can display system-environment correlations and is described by a density operator $\rho_{1^i}^{SE} \in \text{BL}(\mathcal{H}_{1^i}^{SE})$. We label the state with the subscript 1^i to denote that it is the *input* state to the first interrogation procedure applied by the experimenter, namely the initial preparation procedure applied to the system. The initial preparation is only different in name to any of the subsequent probing operations that an experimenter will be allowed to implement in the multi-time setting; the sole reason for the distinction is to emphasise that in the special case where there are no initial system-environment correlations, any system state can be prepared by the experimenter independently of the process, and so the description of the process will begin on the space associated to the *output* of the preparation map, labelled 1^o . However, this is not generally the case: the possibility of initial system-environment correlations make the role of preparation of significant importance [65, 73, 74]. We will now examine

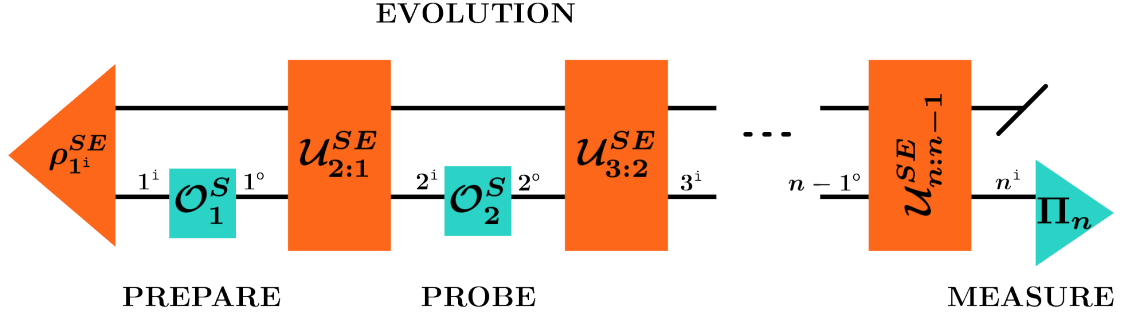


Figure 3.1: *Multi-time quantum experiment.* According to the steps of Def. 3.1, first an initially unknown state of a system, potentially correlated with its environment, is prepared according to the preparation procedure \mathcal{O}_1^S in Eq. (3.2). The system and environment subsequently evolve via the joint unitary evolution $\mathcal{U}_{2:1}^{SE}$. Then the experimenter probes the system, applying another CP map \mathcal{O}_2^S . The probing repeats a number of times, with the final state being measured with the POVM Π_n . We colour the operations that an experimenter can control, namely the preparation, the probing instruments and the final measurement, in green. In contrast, the underlying process over which an experimenter has no control, consisting of the initial system-environment state and the subsequent joint unitary evolutions, is coloured in orange. Note that each timestep is associated to an input and an output space, labelled from the perspective of the experimenter.

each of the individual elements in this setting in order to motivate the computation of joint probabilities. We begin with the two-time case, where an experimenter applies a preparation procedure (on a potentially initially correlated state) and a subsequent measurement, which is the first generalisation of quantum channels.

For instance, the experimenter might apply a projective measurement to yield a known system state with some probability. As we briefly touched on in Subsection 2.2.2, since such conditional transformations between states must occur with at most unit probability, any valid physical preparation procedure on a quantum system S must be a CP and trace-non-increasing map.² Specifically, these are transformations $\mathcal{O}_1^S : \text{BL}(\mathcal{H}_{1^i}^S) \rightarrow \text{BL}(\mathcal{H}_{1^o}^S)$, which takes input states $\rho_{1^i}^S \in \text{BL}(\mathcal{H}_{1^i}^S)$ to subnormalised output states $\rho_{1^o}^S \in \text{BL}(\mathcal{H}_{1^o}^S)$ via

$$\rho_{1^o}^S = \mathcal{O}_1^S(\rho_{1^i}^S) \quad \text{such that} \quad \text{tr}[\rho_{1^o}^S] \leq \text{tr}[\rho_{1^i}^S]. \quad (3.1)$$

The trace of the output state encodes the probability of the specific CP transformation being realised (for the given input state). Preparation procedures thus defined can correspond to any physically implementable transformation, including unitary evolutions, CPTP transformations and measurements. After the preparation procedure, the system state is a subnormalised density operator. More generally, in the multi-time scenario,

² Such maps are referred to as CP maps, with their trace-non-increasing nature implicit.

keeping track of success probabilities encoded in the trace of the associated output state will prove helpful for the computation of joint measurement statistics.

Any such preparation on the system acts non-trivially on correlations that the system shares with its environment. Given some initially correlated $\rho_{1^i}^{SE} \in \text{BL}(\mathcal{H}_{1^i}^{SE})$, a CP preparation procedure $\mathcal{O}_1^S : \text{BL}(\mathcal{H}_{1^i}^S) \rightarrow \text{BL}(\mathcal{H}_{1^o}^S)$ yields the joint output state

$$\rho_{1^o}^{SE} = \mathcal{O}_1^S \otimes \mathcal{I}_1^E (\rho_{1^i}^{SE}), \quad (3.2)$$

where $\rho_{1^o}^{SE} \in \text{BL}(\mathcal{H}_{1^o}^{SE})$ typically exhibits system-environment correlations.

Following this first interrogation by the experimenter, the joint system-environment state is subject to some uncontrollable unitary evolution, as per step (2) of Def. 3.1. The evolution of the joint state from time t_1 to t_2 is represented by the unitary map $\mathcal{U}_{2:1}^{SE} : \text{BL}(\mathcal{H}_{1^o}^{SE}) \rightarrow \text{BL}(\mathcal{H}_{2^i}^{SE})$. Note the inevitable awkwardness of the labelling convention that we run into here, which denotes the output of the unitary transformation with an input label: this is due to the fact that input and output labels are written from the perspective of the experimenter, and so, naturally, the output state of the uncontrollable unitary evolution becomes the input to the next operation applied by the experimenter. Following the first two stages of Def. 3.1, the system-environment state is

$$\rho_{2^i}^{SE} = \mathcal{U}_{2:1}^{SE} (\mathcal{O}_1^S \otimes \mathcal{I}_1^E (\rho_{1^i}^{SE})). \quad (3.3)$$

As a brief aside, note that it is clear that the act of initial preparation can strongly influence the dynamics of the state of the system, because the specific transformation realised can be correlated with the state of the environment, which in turn influences the subsequent evolution. The state of the system following the first two steps of a quantum experiment described above is related to the CP preparation procedure $\mathcal{O}_1^S : \text{BL}(\mathcal{H}_{1^i}^S) \rightarrow \text{BL}(\mathcal{H}_{1^o}^S)$ in the dilated picture via

$$\rho_{2^i}^S = \text{tr}_E [\mathcal{U}_{2:1}^{SE} (\mathcal{O}_1^S \otimes \mathcal{I}_1^E (\rho_{1^i}^{SE}))]. \quad (3.4)$$

To clearly see the influence of the initial preparation, first suppose that the experimenter applies a rank-1 projective measurement on the system, where observation of any particular outcome x corresponds to the conditional transformation $\mathcal{P}^{(x)}(\bullet) = P^{(x)} \bullet P^{(x)}$ being realised with probability $p_x = \text{tr} [P^{(x)} \rho]$. If initial correlations are present, the post-preparation state is³

$$\mathcal{P}^{(x)} \otimes \mathcal{I}^E (\rho_{1^i}^{SE}) = P_{1^o}^{(x)} \otimes \tilde{\tau}_{1^o}^{(x)}, \quad (3.5)$$

³ Here we drop the subsystem labels on conditional states and transformations to avoid clunky notation, with the understanding that the measurement acts on the system alone.

where $\tilde{\tau}_{1^\circ}^{(x)} = \text{tr}_S [\mathcal{P}^{(x)} \otimes \mathcal{I}^E(\rho_{1^\circ}^{SE})] \in \text{BL}(\mathcal{H}_{1^\circ}^E)$ and $P_{1^\circ}^{(x)} \in \text{BL}(\mathcal{H}_{1^\circ}^S)$ describe the post-measurement states. We immediately see that the state of the environment is conditioned by the measurement enacted by the experimenter: the dynamical map description of the subsequent unitary evolution would make it seem as if the process itself is dependent on the preparation, as a distinct quantum channel is induced on the level of the system for each conditional environment state. Consequently, this seems to imply that there is no proper process *per se* that is independent of the experimenter, which is indeed the reason why traditional approaches to open quantum dynamics break down when initial correlations are present.

If, on the other hand, the system is initially uncorrelated from its environment, then any preparation has no influence on the state of the latter. The initially product system-environment hypothesis of the tomographic scheme discussed in Subsection 2.2.4 corresponds to this special case, where states of the system can be prepared without affecting the environment, which retains its unique, fixed state $\tau_{1^\circ}^E \in \text{BL}(\mathcal{H}_{1^\circ}^E)$ throughout the preparation accordingly

$$\sigma_{1^\circ}^S \otimes \tau_{1^\circ}^E = \mathcal{O}_1^S \otimes \mathcal{I}_1^E (\rho_{1^\circ}^S \otimes \tau_{1^\circ}^E). \quad (3.6)$$

Satisfaction of Eq. (3.6) means that the dynamical map describing the subsequent evolution of the experiment is independent of the preparation procedure and uniquely defined, since the environment state can be treated as a constant of the problem. However, the assumption of an initial product state is not satisfied in many realistic circumstances, especially beyond the weak-coupling regime [45, 46, 65, 128].

Returning to the main point of this subsection, at step (3) of the procedure in Def. 3.1 the experimenter has access to the system part of $\rho_{2^\circ}^{SE}$ given in Eq. (3.3). Again, they can apply any CP map of their choosing, before the joint system-environment state is subject to another portion of unitary evolution, and so on. For a process where the experimenter has access to the system at n timesteps, the final state of the system after applying a sequence $\mathcal{O}_1^S, \dots, \mathcal{O}_{n-1}^S$ of CP maps is

$$\rho_{n^\circ}^S = \text{tr}_E [\mathcal{U}_{n:n-1}^{SE} \mathcal{O}_{n-1}^S \dots \mathcal{U}_{2:1}^{SE} \mathcal{O}_1^S (\rho_{1^\circ}^{SE})], \quad (3.7)$$

where all maps act upon everything to their right. Since the initial system-environment state, the subsequent joint unitary evolution and the discarding of the environmental degrees of freedom all occur deterministically, the trace of the output state $\rho_{n^\circ}^S$ is sub-normalised with respect to the probability of realising the sequence of transformations applied by the experimenter; by tracking the operations applied, we come to a meaning-

ful description of conditional quantum states. At the conclusion of the experiment, the final state is measured with a [POVM](#) Π_n .

3.1.2 Process Tensor

The important shift in perspective that allows for a consistent description of multi-time quantum dynamics is to separate what an experimenter has control over from what they do not. Although they cannot generally know, in advance, the initially correlated system-environment state, nor the subsequent portions of joint unitary evolution, what *can* be controlled is the choice of instrument interrogating the system at each timestep. An operationally meaningful framework for describing quantum evolution therefore necessarily consists of two parts: **i)** the *uncontrollable* underlying process which governs the joint unitary evolution of the system with some inaccessible environment, and **ii)** the interleaved *controllable* changes to the state of the system, effected by the probing operations implemented by an experimenter.

For instance, in describing open quantum dynamics in the presence of initial correlations (*i. e.*, the most general two-time quantum experiment as per Def. 3.1), what is desirable is *not* a dynamical map taking initial system density operators—which are uncontrollable—to final ones; indeed, such maps are not linear in general [111], or cannot be reconstructed [129]. Instead, we seek a *map on a map*, which takes any choice of the preparation map as input and outputs the final density operator of the system. To this end, one can define the *superchannel*, $\mathcal{T}_{2:1}$ [74], to represent everything that is *out of the control of an experimenter*, *i. e.*, everything on the r.h.s of Eq. (3.4) except for \mathcal{O}_1^S , as shown in Fig. 3.2

$$\mathcal{T}_{2:1}^S(\bullet) := \text{tr}_E \left[\mathcal{U}_{2:1}^{SE} \left(\bullet \otimes \mathcal{I}_1^E(\rho_{1i}^{SE}) \right) \right]. \quad (3.8)$$

The superchannel contains information regarding both the initially correlated state ρ_{1i}^{SE} and the subsequent joint unitary dynamics $\mathcal{U}_{2:1}^{SE}$; however, from the superchannel alone we cannot delineate the individual contributions of these elements (without access to the environment). On the other hand, the superchannel contains within it, by definition, everything required to determine the output state of a quantum experiment for *any* preparation procedure, making it the object of fundamental operational importance.

The superchannel acts on a [CP](#) preparation map and yields the correct output state in analogy to the way the dynamical map acts on an input density operator. It is linear in its argument by construction, which permits its tomographic reconstruction

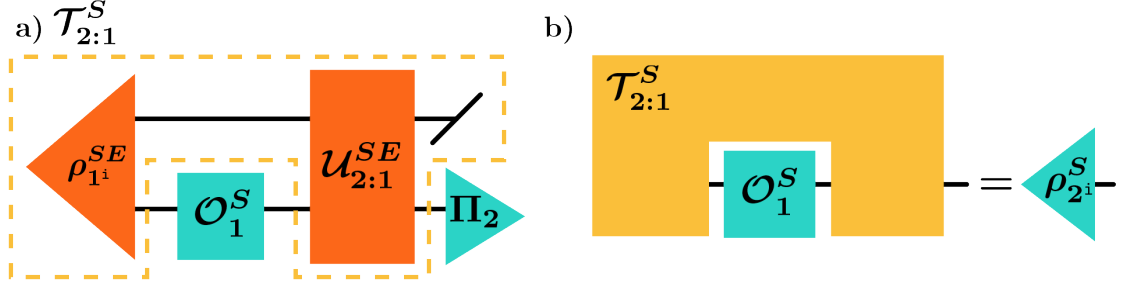


Figure 3.2: *Superchannel: resolution to the initial correlation problem.* In panel **a)** we depict a two-step quantum experiment. By separating what is controllable to the experimenter from what is uncontrollable, we can define the superchannel $\mathcal{T}_{2:1}^S$, which is everything enclosed in the dashed, yellow border. In panel **b)**, we depict how this higher-order map acts on any preparation \mathcal{O}_1^S to yield the correct output state ρ_{2i}^S of the system that is accessible to the experimenter at the second timestep.

in a similar way to that of a dynamical map (see Subsection 2.2.2), *i. e.*, by linearly extending input-output relations deduced for a basis set to uniquely determine its action on arbitrary inputs [74]. However, in contrast to the dynamical map, the inputs of the superchannel are CP preparation maps rather than density operators. While a set of d^2 linearly independent density operators can be chosen to span the space of quantum states, d^4 linearly independent CP maps are required to span the space of allowable preparations [51, 123]. Indeed, such an experimental reconstruction of a superchannel has recently been achieved in the laboratory to characterise the evolution of a photonic qubit that is initially correlated with a single-photon environment [130].

Thus, the superchannel is the natural logical extension of the operationally accessible input-output relations that motivated the dynamical maps description, taking controllable inputs to measurable outputs, with the generalisation allowing for the presence of initial correlations. Moreover, the superchannel satisfies natural extensions of the notions of complete-positivity and trace-preservation [74], as we will discuss for the more general process tensor shortly. Indeed, this approach operationally solves the problem of describing quantum dynamics in the presence of initial correlations. Most importantly for our purposes, the superchannel includes all two-point correlations of the process, allowing for a consistent calculation of multi-time statistics for any choice of interrogation by the experimenter.

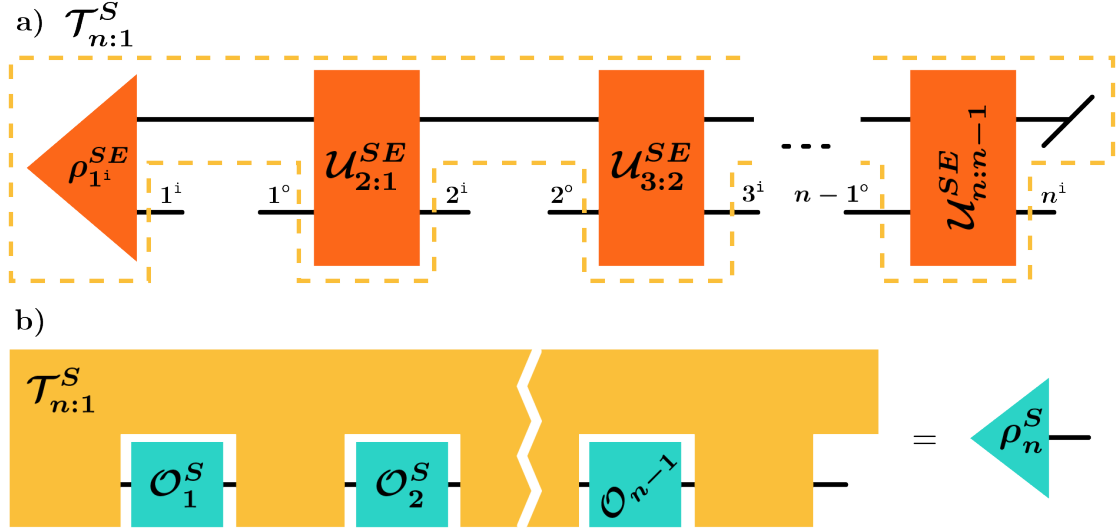


Figure 3.3: *Process tensor: an operational description of quantum stochastic processes.* In panel a), we show the abstraction of everything that is uncontrollable to an experimenter, which defines the process tensor, $\mathcal{T}_{n:1}^S$. In panel b), we depict how the process tensor acts on sequences of CP maps applied by an experimenter on the level of the system to the final density operator.

In analogy to the definition of the superchannel for the two-time case, one can abstract all that is uncontrollable in an open process across multiple timesteps as the *process tensor*, $\mathcal{T}_{n:1}$, depicted in Fig. 3.3, as follows⁴

$$\rho_n^S = \text{tr}_E \left[\mathcal{U}_{n:n-1}^{SE} \mathcal{O}_{n-1}^S \dots \mathcal{U}_{2:1}^{SE} \mathcal{O}_1^S \rho_{1^i}^{SE} \right] =: \mathcal{T}_{n:1}^S(\mathcal{O}_{n-1}^S, \dots, \mathcal{O}_1^S). \quad (3.9)$$

The process tensor is a multi-linear mapping from the sequences of CP control operations upon which it acts to the quantum states at the final output. Due to this linearity, it follows that the process tensor, like the superchannel and the quantum channels that it generalises, can be experimentally reconstructed in a finite number of experiments by way of an extended tomographic scheme [50, 123]. Since the process tensor acts only upon operations applied on the level of the system, from this point forth we will drop the superscript label S for any maps that act on/states that live on the system (unless potentially ambiguous).

One of the advantages of the process tensor formalism is that it directly relates to an operational picture which clarifies a number of concepts pertinent to open quantum dynamics. For example, recall the breakdown of the consistency conditions of the KET on the level of measured probability distributions in quantum theory. For classical stochastic

⁴ Here, on the rightmost side where everything acts on the system alone, we economise the timestep labelling with the understanding that each timestep is associated to an input and output Hilbert space by writing $j = \{j^i, j^\circ\}$ and further compressing ordered sequences of timesteps as $n:1 := \{1, \dots, n\}$.

processes, averaging over measurement outcomes amounts to doing nothing to the system on average, and so the descriptors of the process on any subset of times obtained via marginalisation of deduced statistics are compatible. In quantum theory, a distinction between marginalisation and doing nothing arises because measurements in different bases have different overall effects. As depicted in Fig. 3.4, doing nothing to the system is by no means equivalent to marginalising over the statistics observed for a fixed sequence of measurements.

It is obvious from the operational formalism developed that the process tensor fulfils the following natural consistency condition [14]: for any two sets of timesteps $\Lambda_k \subset \Lambda_n$, the descriptor of the process over the smaller set of times can be obtained from \mathcal{T}_{Λ_n} by letting it act on identity maps $\mathcal{I}_{\Lambda_n \setminus k}$ at the times $\Lambda_n \setminus k$, as highlighted in Fig. 3.5. Mathematically, the consistency condition reads

$$\mathcal{T}_{\Lambda_k}(\bullet) = \mathcal{T}_{\Lambda_n}(\mathcal{I}_{\Lambda_n \setminus k}, \bullet) =: \mathcal{T}_{\Lambda_n}^{|\Lambda_k}(\bullet) \quad (3.10)$$

where \bullet is a placeholder for operations one could implement at the remaining timesteps in Λ_k . By separating the controllable influence on the system from the underlying process, the process tensor formalism allows us to recover compatibility for the descriptors of quantum stochastic processes for different sets of times, and with this a generalised version of the KET can be derived [14, 27]. This result serves to define what we mean by a *quantum stochastic process*, paving the way for an unambiguous study of them.

Up until this point, we have stressed the intuitive picture of the process tensor as the object that allows an experimenter to compute all possible multi-time statistics they might deduce via actively probing the system of interest. We now show how these can be calculated directly for any sequence of interrogations. Recall that a measurement is represented by a POVM $\mathcal{J} = \{\Pi^{(x)\text{T}}\}$, where each of the elements corresponds to a possible outcome, and they satisfy the summation condition $\sum_x \Pi^{(x)} = \mathbb{1}$ ensuring that *some* outcome occurs with certainty. In the temporal setting where the post-measurement state of the system is of interest, the natural generalisation of a POVM is an *instrument*, which is a collection of CP maps $\mathcal{J} = \{\mathcal{O}^{(x)}\}$ that overall yields a CPTP map $\sum_x \mathcal{O}^{(x)} = \mathcal{O}^{\mathcal{J}}$ [6, 26]. Intuitively, the requirement that all the CP maps that make up the instrument sum up to a CPTP map means that *some* transformation to the system occurs with certainty. Thus, if an experimenter applies a sequence of CP maps $\mathcal{O}_1^{(x_1)}, \dots, \mathcal{O}_{n-1}^{(x_{n-1})}$, each of which are elements of an instrument $\mathcal{J}_1, \dots, \mathcal{J}_{n-1}$, with a measurement instrument

Figure 3.4: *Doing nothing vs. averaging over measurements.* In quantum theory there is a difference between doing nothing to the system, \mathcal{I} , and averaging over all measurement outcomes, $\{\mathcal{M}^{(x)}\}$.

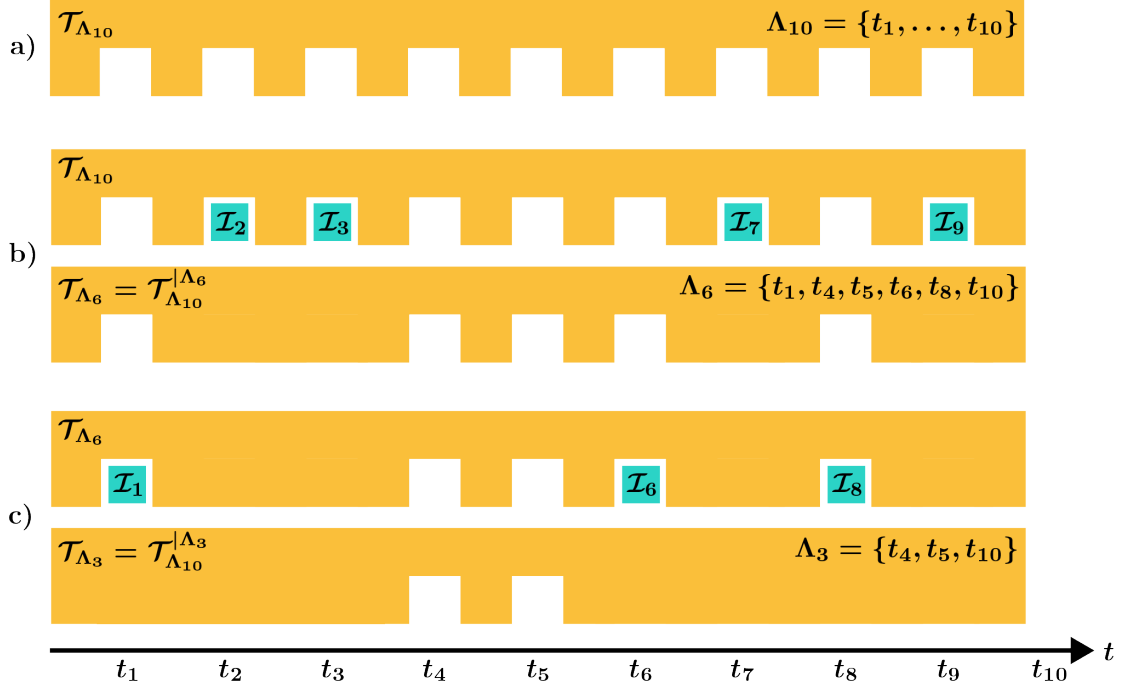


Figure 3.5: *Consistency condition for the process tensor.* The process tensor satisfies a natural consistency condition. For concreteness, in panel **a)** we depict a process tensor over ten timesteps. From this, the correct descriptor on any subset of timesteps can be derived by letting it act on identity maps at the appropriate times. For example, in panel **b)** we show how the correct description over times $\Lambda_6 = \{t_1, t_4, t_5, t_6, t_8, t_{10}\}$ can be obtained in this way from that defined on $\Lambda_{10} = \{t_1, \dots, t_{10}\}$. Moreover, in panel **c)**, we show the containment of \mathcal{T}_{Λ_3} in both descriptors \mathcal{T}_{Λ_6} and $\mathcal{T}_{\Lambda_{10}}$, where $\Lambda_3 = \{t_4, t_5, t_{10}\}$. The crucial point is that the unique maximal description contains within it the proper description of the process over any subset of timesteps.

$\mathcal{J}_n = \{\Pi_n^{(x_n)^T}\}$ applied to the final state, the joint probability distribution over the outcomes realised can be calculated from the process tensor directly via

$$\begin{aligned} \mathbb{P}(x_n, \dots, x_1 | \mathcal{J}_n, \dots, \mathcal{J}_1) &= \text{tr} \left[\Pi_n^{(x_n)^T} \rho_n \right] \\ &= \text{tr} \left[\Pi_n^{(x_n)^T} \mathcal{T}_{n:1} \left(\mathcal{O}_{n-1}^{(x_{n-1})}, \dots, \mathcal{O}_1^{(x_1)} \right) \right]. \end{aligned} \quad (3.11)$$

The process tensor contains all joint probability distributions for all possible measurement settings, and is thus the natural generalisation of classical stochastic processes, as well as quantum states (see below).

To summarise the developments so far in this subsection, recall that for classical stochastic processes, it is the hierarchy of compatible joint probability distributions over all timesteps that serves to characterise the underlying process. For quantum processes, each event must be associated to a **CP** map on the system, and it is the process tensor that characterises the process by mapping any possible multi-time sequence of **CP**

maps to the correct joint statistics via Eq. (3.11). By way of the generalised KET, one can straightforwardly deduce the correct descriptor of the process over any subset of timesteps, thereby alleviating conceptual difficulties regarding the proper characterisation of quantum stochastic processes. Indeed, by accounting succinctly for all possible sequences of interventions on a system of interest, the process tensor encodes all possible multi-time correlations between deducible statistics and therefore, on a sufficiently fine-grained set of timesteps, captures the most general evolutions possible in both quantum and classical physics.

We reiterate the important conceptual departure from traditional approaches to open quantum dynamics: there, descriptions of quantum processes typically involve tracking the state of the system as a function of time, which limits the ability to calculate the outcomes of measurements to at most two timesteps in any given trial of the experiment, inherently failing to capture multi-time memory effects that are critical to understanding processes with memory. Clearly, the system density operator at each timestep can be obtained from the process tensor by simply plugging in identity maps at all of the preceding timesteps, thereby unifying all such two-point descriptions. Besides providing a more intuitive operational picture, the process tensor description also subsumes the standard approaches to quantum processes with memory by way of non-Markovian master equations, which aim to account for the effects of some memory kernel on the evolution of the system. Admittedly, the process tensor is more of a characterisation than a dynamical description; nonetheless, on the timesteps upon which it is defined, it provides a more general description, accounting for all possible multi-time correlations deducible—rather than only those either derived microscopically or deduced phenomenologically—the process tensor goes beyond the realm of applicability of such approaches.

Crucially, we now have a way to meaningfully construct quantum generalisations of statements that are multi-time in nature, such as Markovianity [49, 50] and Markov order [1, 2], allowing for a consistent study of memory in quantum stochastic processes from an operationally sound perspective. Before we do so, we present a brief mathematical interlude to develop a useful representation of the process tensor as a multi-partite quantum state, which will prove fruitful for understanding its properties and proving statements throughout the remainder of this thesis.

3.2 REPRESENTING LINEAR MAPS

Up to this point, we have discussed dynamical maps and process tensors on rather abstract grounds as mappings. For explicit statements, it proves helpful to choose a representation that is well-adapted to the respective purpose. There are various different explicit representations of maps describing valid physical evolution in quantum mechanics, each suitable for certain purposes, such as the tomographic representation of Eq. (2.41), the so-called *A*- and *B*-forms of dynamical maps first introduced by Sudarshan *et al.* in Ref. [131], and the Kraus decomposition [80, 132], amongst others; for an overview of the inter-relations between such representations, see, *e. g.*, Ref. [51]. Here, we introduce and use exclusively the *Choi-Jamiołkowski Isomorphism* (CJI) [133, 134].

3.2.1 Choi-Jamiołkowski Isomorphism

This isomorphism allows us to consider any linear map taking elements of some input vector space to some other output vector space as a single element of the joint input-output vector space. Concretely, consider a linear map \mathcal{L} acting on the bounded linear operators on a Hilbert space $\mathcal{L} : \text{BL}(\mathcal{H}_i) \rightarrow \text{BL}(\mathcal{H}_o)$, where, for generality, we allow for the input and output Hilbert spaces to be distinct. This map can be represented as a bipartite operator $L_{oi} \in \text{BL}(\mathcal{H}_o \otimes \mathcal{H}_i)$ through its action on half of an unnormalised maximally-entangled state $\Psi := \sum_{ij}^d |ii\rangle\langle jj| \in \text{BL}(\mathcal{H}_i \otimes \mathcal{H}_i)$ as follows

$$L_{oi} := \mathcal{L} \otimes \mathcal{I}(\Psi). \quad (3.12)$$

See Fig. 3.6 for a graphical representation. Note that we consistently use upper-case Roman calligraphic letters to denote maps and their sans-serif variant to denote their corresponding representation in terms of the CJI. We refer to the matrix L_{oi} resulting from Eq. (3.12) as the *Choi operator* associated to \mathcal{L} .

The action of the map \mathcal{L} on an arbitrary element of its input space $\eta_i \in \text{BL}(\mathcal{H}_i)$ can be expressed in terms of its Choi operator L_{oi} via

$$\mathcal{L}(\eta_i) = \text{tr}_i \left[\left(\mathbb{1}_o \otimes \eta_i^T \right) L_{oi} \right], \quad (3.13)$$

where $\mathbb{1}_o := \sum_i^d |i\rangle\langle i| \in \text{BL}(\mathcal{H}_o)$ is the identity operator on the output Hilbert space and we slightly abuse notation by writing $\text{tr}_i[\bullet] := \text{tr}_{\mathcal{H}_i}[\bullet]$. The validity of Eq. (3.13) can be shown by direct insertion of Eq. (3.12) as follows

$$\begin{aligned}
 \text{tr}_i \left[(\mathbb{1}_o \otimes \eta_i)^T \mathbb{L}_{oi} \right] &= \text{tr}_i \left[\left(\mathbb{1}_o \otimes \eta_i^T \right) \mathcal{L} \otimes \mathcal{I}(\Psi) \right] \\
 &= \sum_{ij}^{d_i} \text{tr}_i \left[\left(\mathbb{1}_o \otimes \eta_i^T \right) \mathcal{L}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \right] \\
 &= \sum_{ijk}^{d_i} \mathcal{L}(|i\rangle\langle j|) \langle k | \eta_i^T | i \rangle \langle j | k \rangle = \sum_{ij}^{d_i} \mathcal{L}(|i\rangle\langle j|) \langle j | \eta_i^T | i \rangle \\
 &= \sum_{ij}^{d_i} \mathcal{L}(|i\rangle\langle j|) \eta_i^{ij} = \mathcal{L} \left(\sum_{ij}^{d_i} \eta_i^{ij} |i\rangle\langle j| \right) = \mathcal{L}(\eta_i),
 \end{aligned} \tag{3.14}$$

where the final line of equalities holds by the linearity of \mathcal{L} and the decomposition of an arbitrary element of $\text{BL}(\mathcal{H}_i)$ as $\eta_i = \sum_{ij}^{d_i} \eta_i^{ij} |i\rangle\langle j|$.

While the [CJI](#) holds for linear maps in general, for the types of evolution that are physically meaningful in quantum theory, Choi operators have particularly nice properties.⁵ Consider a [CPTP](#) map $\mathcal{C} : \text{BL}(\mathcal{H}_i) \rightarrow \text{BL}(\mathcal{H}_o)$, where again, for generality, we allow the input and output systems of the map to be distinct. Complete-positivity and trace-preservation for the quantum channel \mathcal{C} translate into the following properties of its Choi operator \mathbb{C}_{oi}

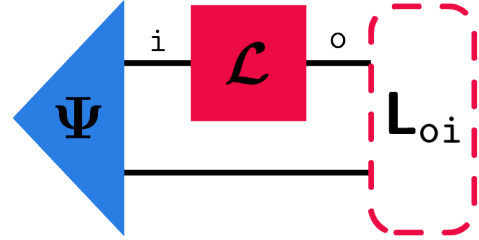


Figure 3.6: [CJI](#) of a linear map. Any linear map $\mathcal{L} : \mathcal{H}_i \rightarrow \mathcal{H}_o$ can be represented as a bipartite operator $\mathbb{L}_{oi} \in \text{BL}(\mathcal{H}_o \otimes \mathcal{H}_i)$ through its action on half of a maximally entangled state Ψ .

1. Complete-positivity : $\mathbb{C}_{oi} \geq 0$. (3.15)

2. Trace-preservation : $\text{tr}_o [\mathbb{C}_{oi}] = \mathbb{1}_i$.

Proof. 1. Clearly, a [CP](#) map \mathcal{C} corresponds to a positive semidefinite Choi operator $\mathbb{C}_{oi} \geq 0$ by definition. Any Hermitian,⁶ positive semidefinite operator admits a singular-value eigendecomposition $\mathbb{C}_{oi} = \sum_{\alpha=1}^{d_{oi}} \lambda_\alpha |\alpha\rangle\langle\alpha|$ with each $\lambda_\alpha \geq 0$ and $d_{oi} := d_o d_i$. The action of \mathcal{C} on an arbitrary state $\rho_i \in \text{BL}(\mathcal{H}_i^S)$ can then be written as

$$\begin{aligned}
 \mathcal{C}(\rho_i) &= \text{tr}_i \left[\left(\mathbb{1}_o \otimes \rho_i^T \right) \mathbb{C}_{oi} \right] \\
 &= \text{tr}_i \left[\left(\mathbb{1}_o \otimes \rho_i^T \right) \sum_{\alpha=1}^{d_{oi}} \lambda_\alpha |\alpha\rangle\langle\alpha| \right] = \sum_{\alpha=1}^{d_{oi}} \lambda_\alpha \sum_{ij}^{d_i} \langle i | \alpha \rangle \langle j | \rho_i^T | i \rangle \langle \alpha | j \rangle
 \end{aligned} \tag{3.16}$$

⁵ Indeed, it turns out that the Choi representation of a quantum dynamical map is equivalent to the *B*-form introduced in Ref. [131] for precisely this reason.

⁶ It is straightforward to show that the Hermiticity-preservation property of the dynamical map (which is implied by the [CP](#) condition) leads to an Hermitian Choi operator.

$$= \sum_{\alpha=1}^{d_{\text{o}}} \left(\sum_i^{d_{\text{i}}} \sqrt{\lambda_{\alpha}} |i\rangle\langle\alpha| \right) \rho_{\text{i}} \left(\sum_j^{d_{\text{i}}} \sqrt{\lambda_{\alpha}} |j\rangle\langle\alpha| \right) =: \sum_{\alpha=1}^{d_{\text{o}}} K_{\alpha} \rho_{\text{i}} K_{\alpha}^{\dagger},$$

where we have introduced the orthonormal basis vectors of the input space $\{|i\rangle\}, \{|j\rangle\}$ to perform the partial trace explicitly, made use of the non-negativity of the eigenvalues to write a unique square-root, and each the K_{α} are $d_{\text{o}} \times d_{\text{i}}$ matrices (since $|i\rangle\langle\alpha| \in \text{BL}(\mathcal{H}_{\text{o}})$). The expression of the dynamical map derived above is known as the Kraus or operator-sum representation, with each K_{α} known as the Kraus operator of the map. A fundamental result given in Ref. [132] states that a map is **CP** iff it can be written in the Kraus form above, concluding the proof. \square

Proof. 2. A **TP** map \mathcal{C} satisfies $\text{tr}[\mathcal{C}(\rho_{\text{i}})] = \text{tr}[\rho_{\text{i}}] \ \forall \ \rho_{\text{i}} \in \text{BL}(\mathcal{H}_{\text{i}})$. Writing this out explicitly in terms of the Choi operator \mathbf{C}_{oi} , we have

$$\text{tr}[\mathcal{C}(\rho_{\text{i}})] = \text{tr}_{\text{o}} \left[\text{tr}_{\text{i}} \left[\left(\mathbb{1}_{\text{o}} \otimes \rho_{\text{i}}^{\text{T}} \right) \mathbf{C}_{\text{oi}} \right] \right] = \text{tr}_{\text{i}} \left[\rho_{\text{i}}^{\text{T}} \text{tr}_{\text{o}} [\mathbf{C}_{\text{oi}}] \right] = \text{tr}[\rho_{\text{i}}], \quad (3.17)$$

which holds true for all ρ_{i} iff $\text{tr}_{\text{o}} [\mathbf{C}_{\text{oi}}] = \mathbb{1}_{\text{o}}$. \square

The set of physically allowable **CPTP** maps are therefore equivalent to bipartite operators $\mathbf{C}_{\text{oi}} \in \text{BL}(\mathcal{H}_{\text{o}} \otimes \mathcal{H}_{\text{i}})$ satisfying the conditions outlined in Eq. (3.15). A few remarks are in order.

i) It is clear from the trace-preservation property that $\text{tr}[\mathbf{C}_{\text{oi}}] = d_{\text{i}}$. Since the Choi operator of a **CPTP** channel must also be positive semidefinite, it can therefore be regarded as a supernormalised quantum state. That is, any such Choi operator lies in the convex cone of non-negative bounded linear operators of the joint input-output Hilbert space. More precisely, the set of Choi operators of **CPTP** maps corresponds to the intersection of the cone of positive semidefinite operators $\mathbf{A}_{\text{oi}} \geq 0$ with the hyperplane of those that satisfy $\text{tr}_{\text{o}} [\mathbf{A}_{\text{oi}}] = \mathbb{1}_{\text{i}}$.

ii) Whilst all **CPTP** maps can be uniquely identified with a (potentially supernormalised) density operator, not all states in $\text{BL}(\mathcal{H}_{\text{o}} \otimes \mathcal{H}_{\text{i}})$ represent a **CPTP** evolution, since, although they are positive semidefinite, a generic quantum state does not necessarily satisfy the additional trace-preservation constraint.

iii) A general **CP** map corresponds to a positive semidefinite Choi operator that does not necessarily satisfy the second property in Eq. (3.15), but instead must satisfy $\text{tr}_{\text{o}} [\mathbf{C}_{\text{oi}}] \leq \mathbb{1}_{\text{i}}$.

iv) Lastly, through the **CJI** we can think of states and effects as the Choi operators of **CP** maps with trivial input and output spaces respectively. States can be considered as the Choi operator of a **CPTP** map $\mathcal{R} : \mathbb{R} \rightarrow \text{BL}(\mathcal{H}_{\text{o}})$, and the normalisation of the

state $\text{tr}_o[\rho] = 1$ can be understood as a trace-preservation condition. In this sense, all normalised quantum states are **TP**, which is intuitive since they can be prepared deterministically. Effects $\mathcal{E}^{(x)} : \text{BL}(\mathcal{H}_o) \rightarrow \mathbb{R}$ map quantum states to real numbers (probabilities) and in this sense positivity of effects is equivalent to complete-positivity, since the output space is trivial.⁷ The action of an effect on a quantum state is written in terms of their Choi operators as $\mathcal{E}^{(x)}(\rho) = \text{tr}[\rho^T \Pi^{(x)}] = \text{tr}[\Pi^{(x)T} \rho]$, which unsurprisingly gives the Born rule (see Eq. (2.31)). In contrast to the case for states, the fact that the output space of effects is trivial implies that there is only one trace-preserving effect, namely the trace map $\mathcal{E}(\rho) = \text{tr}[\rho]$ whose Choi operator is the identity matrix $\mathbb{1}$ (see Fig. 3.7).

As a final brief example, note that we can write the tomographic representation of a quantum channel given in Eq. (2.41) in terms of the dual objects to the set of input spanning states, $\{\hat{D}_i^{(i)}\}$, and the corresponding output states, $\{\hat{\sigma}_o^{(i)}\}$, as follows

$$C_{oi} = \sum_i \hat{\sigma}_o^{(i)} \otimes \hat{D}_i^{(i)*}, \quad (3.18)$$

where $(\bullet)^*$ denotes complex conjugation. The validity of this construction can be seen directly by insertion

$$\begin{aligned} \mathcal{C}(\rho_i) &= \text{tr}_i \left[\left(\mathbb{1}_o \otimes \rho_i^T \right) C_{oi} \right] = \text{tr}_i \left[\left(\mathbb{1}_o \otimes \rho_i^T \right) \sum_i \hat{\sigma}_o^{(i)} \otimes \hat{D}_i^{(i)*} \right] \\ &= \sum_i \hat{\sigma}_o^{(i)} \text{tr} \left[\rho_i^T \hat{D}_i^{(i)*} \right] = \sum_i \hat{\sigma}_o^{(i)} \text{tr} \left[\hat{D}_i^{(i)\dagger} \rho_i \right]. \end{aligned} \quad (3.19)$$

3.2.2 Choi Representation of Process Tensor

We now employ the **CJI** to explicitly represent process tensors. Recall that its action is to take sequences of **CP** maps as its input and map them to some final output density operator. For generality, we allow the input and output systems of the **CP** maps to be distinct, as too the final output of the process tensor. With the **CJI**, we can represent any **CP** map $\mathcal{O}_j : \text{BL}(\mathcal{H}_{j^i}) \rightarrow \text{BL}(\mathcal{H}_{j^o})$ as a Choi operator $O_j \in \text{BL}(\mathcal{H}_{j^o} \otimes \mathcal{H}_{j^i})$. The process tensor acts on sequences of these as $\mathcal{T}_{n:1} : \text{BL}(\bigotimes_{j=1}^{n-1} \mathcal{H}_{j^o} \otimes \mathcal{H}_{j^i}) \rightarrow \text{BL}(\mathcal{H}_{n^i})$. As for the case of quantum channels, we can invoke the **CJI** to represent the process tensor as a positive matrix on the correct space, that satisfies certain trace conditions.

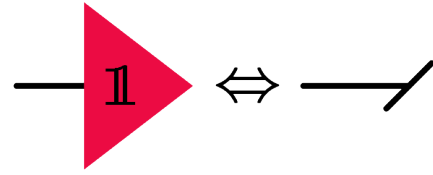


Figure 3.7: The unique trace-preserving effect. The only trace-preserving effect is the trace map, corresponding to an identity Choi operator $\mathbb{1}$. To abide by traditional notation, we depict this with a slash.

⁷ Note that the input space of the effect is labelled with o , in line with previous notation.

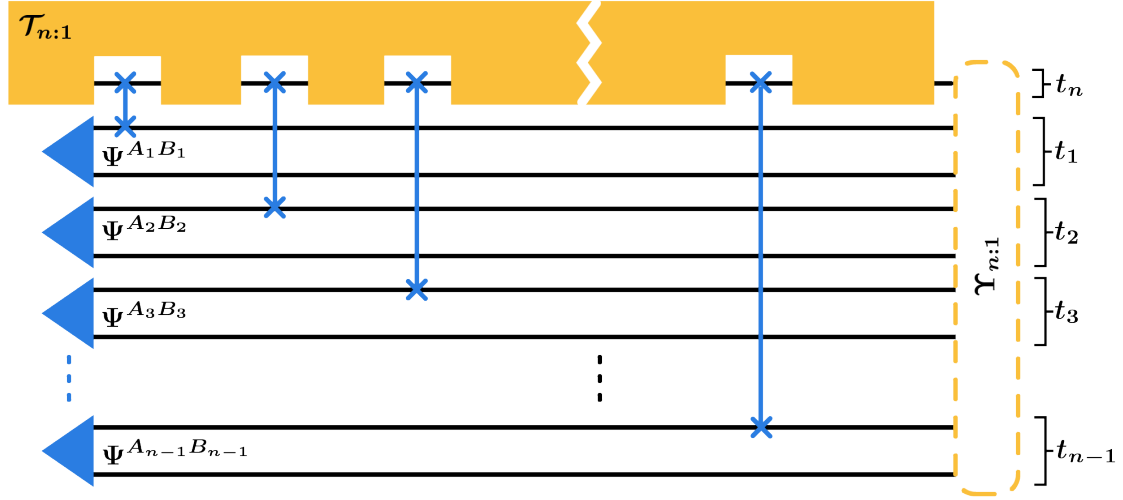


Figure 3.8: *CJI for the process tensor.* The multi-linear process tensor map, $\mathcal{T}_{n:1}$, can be represented as a many-body Choi operator, $\Upsilon_{n:1}$ through the *CJI*. At each timestep $t_j \in \{t_1, \dots, t_n\}$, half of an (unnormalised) maximally entangled state, $\Psi^{A_j B_j}$, is swapped into the process (blue crosses). The resulting $2n - 1$ body quantum state $\Upsilon_{n:1}$ contains equivalent information to the multi-linear map $\mathcal{T}_{n:1}$, with temporal correlations of the process being mapped to spatial ones between subsystems of its Choi operator. This is depicted by the brackets on the right, which signify the degrees of freedom of $\Upsilon_{n:1}$ that correspond to different times of the process t_j .

In detail, the many-body Choi operator corresponding to the process tensor is depicted in Fig. 3.8 and can be physically prepared (up to normalisation) as follows. Begin with $2(n - 1)$ ancillary systems $\{A_j, B_j\}$ of appropriate dimension $d_{j^i} := \dim(\mathcal{H}_{j^i})$ in unnormalised maximally entangled pairs, $\{\Psi^{A_j B_j}\}$, where each $\Psi^{A_j B_j} = \sum_{\alpha\beta}^{d_j} |\alpha\alpha\rangle\langle\beta\beta|$ with $d_j := d_{j^o} d_{j^i}$. At each timestep of the process, half of each pair is swapped with the system state through $\mathcal{G}_j^{SA_j}$. The resultant d^{2n-1} dimensional system-ancillary operator $\Upsilon_{n:1} \in \text{BL}(\mathcal{H}_{n^i} \otimes_{j=1}^{n-1} \mathcal{H}_{j^o} \otimes \mathcal{H}_{j^i})$ encodes equivalent information as the temporal map $\mathcal{T}_{n:1}$, and can be explicitly written as⁸

$$\Upsilon_{n:1} := \text{tr}_E \left[\mathcal{U}_{n:n-1}^{SE} \mathcal{G}_{n-1}^{SA_{n-1}} \dots \mathcal{U}_{2:1}^{SE} \mathcal{G}_1^{SA_1} \left(\Psi^{A_{n-1} B_{n-1}} \otimes \dots \otimes \Psi^{A_1 B_1} \otimes \rho_{1^i}^{SE} \right) \right]. \quad (3.20)$$

It is straightforward to show that the action of the process tensor—in clear analogy to the action of quantum maps in terms of their Choi operators—on an arbitrary sequences of *CP* maps can be expressed in terms of both the Choi operator of the process tensor itself and the maps on which it acts as follows

⁸ Due to the importance of the process tensor as the fundamental object in the operational view of open quantum dynamics, we denote its Choi operator by the special symbol Υ instead of a sans serif letter.

$$\mathcal{T}_{n:1}(\mathcal{O}_{n-1}, \dots, \mathcal{O}_1) = \text{tr}_{n-1:1} \left[\left(\mathbb{1}_{n^i} \bigotimes_{j=1}^{n-1} \mathcal{O}_{j^o j^i}^T \right) \Upsilon_{n:1} \right]. \quad (3.21)$$

More generally still, one could apply a sequence of interrogations correlated across timesteps, *e.g.*, by sending forward the ancilla that was used to implement an earlier operation. The corresponding correlated map $\mathcal{O}_{n-1:1}$ is a similar object to the process tensor (in a way that we will soon make explicit) and represents the most general kind of transformation one could implement over a sequence of timesteps. The action of the process tensor on such a correlated sequence of operations is, again, given by

$$\mathcal{T}_{n:1}(\mathcal{O}_{n-1:1}) = \text{tr}_{n-1:1} \left[\left(\mathbb{1}_{n^i} \otimes \mathcal{O}_{n-1:1}^T \right) \Upsilon_{n:1} \right]. \quad (3.22)$$

3.2.3 Properties of the Process Tensor

Since the process tensor acts on sequences of **CP** maps it makes no sense to speak of complete-positivity and trace-preservation in the original sense; however, from its definition, meaningful extensions of these concepts are satisfied [28, 50]. The following properties follow directly from the definition of the process tensor, but can also be motivated on more axiomatic grounds [28–30].

Complete-positivity means that if the process tensor acts on some sequence of **CP** operations $\mathcal{O}_1^{SA_1}, \dots, \mathcal{O}_{n-1}^{SA_{n-1}}$, where each operation $\mathcal{O}_j^{SA_j} : \text{BL}(\mathcal{H}_{j^i}^S \otimes \mathcal{H}_{j^i}^{A_j}) \rightarrow \text{BL}(\mathcal{H}_{j^o}^S \otimes \mathcal{H}_{j^o}^{A_j})$ acts on the system and some ancilla, the resulting transformation $\mathcal{T}_{n:1}^S \otimes \mathcal{I}_{n:1}^A(\mathcal{O}_{n-1}^{SA_{n-1}}, \dots, \mathcal{O}_1^{SA_1}) : \text{BL}(\mathcal{H}_{n-1^i}^{A_{n-1}} \otimes \dots \otimes \mathcal{H}_{1^i}^{A_1}) \rightarrow \text{BL}(\mathcal{H}_{n^i}^S \otimes \mathcal{H}_{n-1^o}^{A_{n-1}} \otimes \dots \otimes \mathcal{H}_{1^o}^{A_1})$ is a **CP** map, no matter the size of the ancillary spaces; this is represented graphically in Fig. 3.9 for the superchannel. As a special case, this implies that the output of the process tensor, for any physically allowed transformations the experimenter might apply, is always a valid quantum state (up to normalisation). The trace preservation property of quantum channels on the level of process tensors translates to the statement that for any overall deterministic sequences of operations applied, the output quantum state must have unit trace.

These properties are encoded naturally in the Choi operator of the process tensor as

$$\mathbf{1. \text{ Complete-positivity : } \Upsilon_{n:1} \geq 0.} \quad (3.23)$$

$$\mathbf{2. \text{ Trace-preservation : } \text{tr}_{j^i} [\Upsilon_{j:1}] = \mathbb{1}_{j-1^o} \otimes \Upsilon_{j-1:1} \quad \forall t_1 < t_j \leq t_n.}$$

Importantly, each $\Upsilon_{j-1:1}$ in the second condition is a proper process tensor describing the process on the timesteps t_1, \dots, t_{j-1} preceding each t_j . Thus the second property

encapsulates an entire hierarchy of trace conditions, which implies that the trace of a process tensor is equal to the product of the dimension of the system on all of its output Hilbert spaces: $d_{n:1}^o := d_{1^o} \times \dots \times d_{n^o}$. Clearly, the process tensor constructed via Eq. (3.20) satisfies both properties; conversely, every Choi operator satisfying them corresponds to a fixed open dynamics and thereby represents a valid process [50].

The CJI for the process tensor maps *temporal* correlations into *spatial* ones; thus, although almost all of the results to follow are presented in terms of the Choi operators of processes, these statements fundamentally address temporal properties of processes, such as correlations between observables measured over time on some evolving quantum system. As already encountered for the case of CPTP maps, all processes can be represented in

this way as supernormalised many-body quantum states, but not all such quantum states represent valid processes [37]. The set of possible temporal correlations are restricted, compared to their spatial counterparts, because the process tensor must satisfy the above hierarchy of trace-conditions in order that it can be dilated to a *fixed* system-environment model [28–30, 50].

The trace-preservation property on the process tensor can equivalently be viewed as a statement about causality. It is straightforward to show that if the hierarchy of trace conditions in Eq. (3.23) is satisfied, any later choice of instruments cannot influence earlier measurement statistics, and vice versa. This condition can be relaxed without leading to paradoxical situations [30, 32], but all process tensors naturally satisfy causality. We will only sometimes encounter non-ordered objects later on in this thesis; these are obtained from post-selection, which does not enforce causal order [30, 35, 39, 135].

From the vantage point provided by the Choi representation of processes and their properties, we now briefly reconsider from a more formal perspective a number of key concepts that have previously been loosely discussed. This will allow us to concretely calculate the joint probability distributions that are encoded in the process tensor with respect to any interrogating instrument sequence by way of a generalised spatio-temporal Born rule defined on the respective Choi operators.

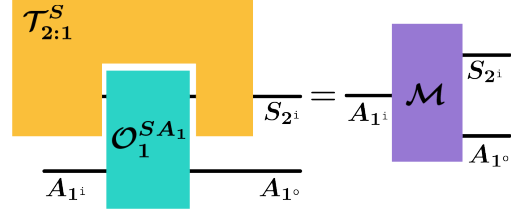


Figure 3.9: Complete-positivity for the superchannel. Complete-positivity for the superchannel $\mathcal{T}_{2:1}^S$ means that when acting on part of a CP map \mathcal{O}^{SA_1} , the resulting map $\mathcal{M} : \text{BL}(\mathcal{H}_{1^i}^{A_1}) \rightarrow \text{BL}(\mathcal{H}_{2^o}^S \otimes \mathcal{H}_{1^o}^{A_1})$ is a CP map, independent of the size of the ancilla.

3.3 SPATIO-TEMPORAL BORN RULE

Quantum theory is, at its core, about measurement statistics observed through experiment. Here we will take a closer look at **POVM** and instruments to make the connection between quantum stochastic processes and classical stochastic processes perspicuous.

It is well-known that quantum mechanics cannot be adequately described within standard probability theory due to inherently non-classical features, such as non-commutativity and contextuality. As discussed, a measurement instrument (*i. e.*, a **POVM**) is a set of such operators $\mathcal{J} = \{\Pi^{(x)\text{T}}\}$ that sum to the identity $\sum_x \Pi^{(x)\text{T}} = \mathbb{1}$. Here, each operator corresponds to an outcome, with the summation condition ensuring that *some* outcome occurs with certainty. With this, the probability for a measurement outcome to be realised for a given quantum state ρ are computed via the Born rule

$$\mathbb{P}(x|\mathcal{J}) = \text{tr} \left[\Pi^{(x)\text{T}} \rho \right]. \quad (3.24)$$

In the above equation, ρ is some *fixed*, *overall deterministic* object (*i. e.*, unit-trace quantum state) that contains all of the statistical information about the system of interest. This information can be deduced by means of the measurement instrument \mathcal{J} , where each of the possible outcomes corresponds to a (positive) **POVM** element, as depicted in Fig. 3.10. Here, the choice of measurement corresponds to what is controllable to an experimenter, and the state upon which it is implemented constitutes that which is uncontrollable. In the language of a generalised probability theory, intuitively, the

measurement instrument plays the role of providing a σ -algebra on the event space, with each constituent **POVM** element corresponding to an event.⁹

However, the Born rule, in its original form, does not properly assign joint probabilities to consecutive events [34]; in order to study temporal processes, one must track the transformations of the system over time upon observation of outcomes, which cannot be

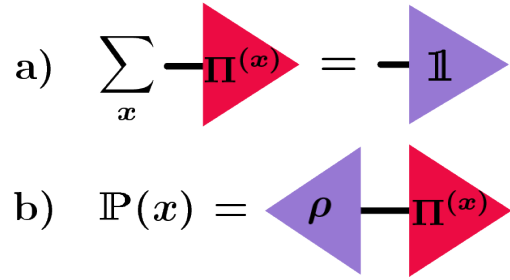


Figure 3.10: Graphical representation of a **POVM**. A **POVM** is a collection of positive operators $\mathcal{J} = \{\Pi^{(x)\text{T}}\}$ that sum to the identity operator $\sum_x \Pi^{(x)\text{T}} = \mathbb{1}$, shown in panel a). That the overall implementation is **TP** allows us to interpret the effect of each **POVM** element on a state, *i. e.*, the outputs of the Born rule, as probabilities, as per panel b).

⁹ Indeed, this is how the **POVM** got its name.

accounted for by **POVM**. As briefly discussed previously, in the temporal setting individual conditionally-realised events are elevated to **CP** transformations, and a collection of these that ensures *some* transformation occurs overall constitutes an instrument, which plays the analogous role to a **POVM**, as illustrated in Fig. 3.11 and formally defined as follows.

Definition 3.2 (Instrument [6, 26]). An *instrument*, \mathcal{J} , is a collection of physical transformations that overall correspond to a deterministic transformation. Concretely, an instrument is represented by a set of **CP** maps, $\{\mathcal{O}^{(x)}\}$, that sum to a **CPTP** map, *i. e.*, $\mathcal{O}^{\mathcal{J}} := \sum_x \mathcal{O}^{(x)}$ satisfies both conditions in Eq. (3.15).

Intuitively, the **CP** operation $\mathcal{O}^{(x)}$ describes how the state of the system is changed upon measuring outcome x , given that the instrument \mathcal{J} was used to interrogate the system. The concept of an instrument captures the most general overall deterministic transformation allowable within quantum theory that an experimenter could invoke at some point in time, including instruments with only a single deterministic ‘outcome’ corresponding to, *e. g.*, a unitary transformation. It is clear that the summation constraint imposed on **POVMs** is a special case of an instrument with a trivial output space, since the only trace-preserving effect is the identity operator. In contrast to the case of **POVMs**, however, the summation constraint on an instrument is non-unique; different instruments can correspond to different (unconditional) **CPTP** transformations.

We have already seen that the most general two-time quantum experiment is described by a superchannel $S_{2^i 1^o 1^i} \in \text{BL}(\mathcal{H}_{2^i} \otimes \mathcal{H}_{1^o} \otimes \mathcal{H}_{1^i})$. Consider then an experimenter applying an instrument $\mathcal{J}_1 := \{\mathcal{O}_{1^o 1^i}^{(x_1)}\}$ followed by a **POVM** on the final output $\mathcal{J}_2 := \{\Pi_{2^i}^{(x_2)}\}$ (since the process ends at t_2 , the post-measurement state is irrelevant and a **POVM** suffices calculate all statistics). For any specific realisation of the preparation procedure applied, the final output state is subnormalised with respect to the probability of the said preparation occurring and is given by

$$\rho_{2^i}^{(x_1)} = \text{tr}_1 \left[\left(\mathbb{1}_{2^i} \otimes \mathcal{O}_{1^o 1^i}^{(x_1)\text{T}} \right) S_{2^i 1^o 1^i} \right]. \quad (3.25)$$

The probability for implementation of a particular **CP** map at t_1 *and* the realisation of a measurement outcome at t_2 to occur is computed by applying the **POVM** on the output state via the standard Born rule, giving the joint statistics (see Eq. (3.11))

$$\begin{aligned} \mathbb{P}_{2:1}(x_2, x_1 | \mathcal{J}_2, \mathcal{J}_1) &= \text{tr} \left[\Pi_{2^i}^{(x_2)\text{T}} \rho_{2^i}^{(x_1)} \right] = \text{tr}_{2^i} \left[\Pi_{2^i}^{(x_2)\text{T}} \text{tr}_1 \left[\left(\mathbb{1}_{2^i} \otimes \mathcal{O}_{1^o 1^i}^{(x_1)\text{T}} \right) S_{2^i 1^o 1^i} \right] \right] \\ &= \text{tr} \left[\left(\Pi_{2^i}^{(x_2)\text{T}} \otimes \mathcal{O}_{1^o 1^i}^{(x_1)\text{T}} \right) S_{2^i 1^o 1^i} \right]. \end{aligned} \quad (3.26)$$

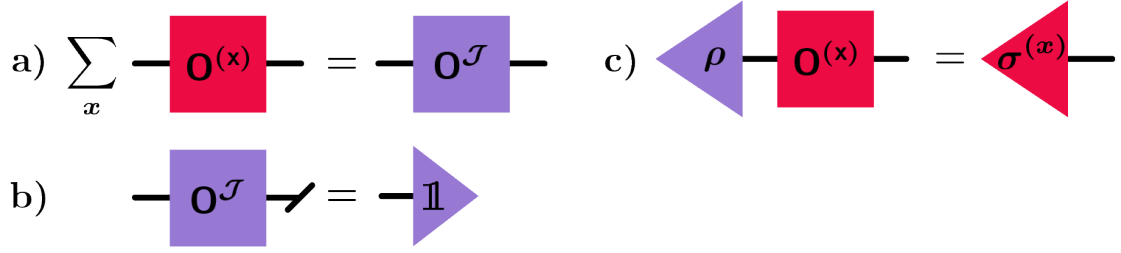


Figure 3.11: *Graphical representation of an instrument.* An instrument generalises the notion of a **POVM**. An instrument is a collection of **CP** maps $\mathcal{J} = \{O^{(x)}\}$ that sum to a **CPTP** map, shown in panel **a**). The trace-preservation constraint on **CPTP** maps entails that tracing over the output space of the map yields an identity operator effect, shown in panel **b**). Although in each run of the experiment a specific transformation is only conditionally realised $O^{(x)}(\rho) = \sigma^{(x)}$, as shown in panel **c**), the summation constraint ensures that overall the instrument transforms the state into a normalised one $\sigma = \sum_x \sigma^{(x)}$.

It is instructive to compare the structure of Eq. (3.26) with that of the standard Born rule in Eq. (3.24). In the two-time scenario, it is the superchannel that now plays the role of the density operator, containing the information about all the probabilities for all the ways in which a two-time process can be interrogated. In this sense, Eq. (3.25) constitutes a spatio-temporal generalisation of the Born rule, at least for uncorrelated probing operations.

In order to describe more general scenarios, we must first extend the definition of instruments to *testers*, after which we will provide a generalised Born rule that allows calculation of all multi-time probabilities. In the setting envisaged, an experimenter could apply a sequence of instruments that are correlated across timesteps, *e.g.*, by sending forward the ancilla that was used to implement an earlier operation. Since the experimenter does something with overall certainty, the (potentially correlated) **CP** operation sequence they implement corresponds to a realisation or trajectory of a process which, when summed over all possible outcomes, yields an overall deterministic transformation. In the previous subsection, we saw that any such multi-time process must have the structure of a process tensor, *i.e.*, satisfy both constraints of Eq. (3.23). We therefore introduce the following definition that generalises the notion of an instrument to the multi-time setting where operations are permitted to be temporally correlated, depicted in Fig. 3.12.

Definition 3.3 (Tester [30] / Instrument sequence). An n -tester or *instrument sequence*, $\mathcal{J}_{n:1}$, is a collection of physical transformations, which may be temporally correlated across n timesteps, that overall correspond to a deterministic transformation. Concretely, a tester is represented by a set of multi-time CP maps defined on n timesteps, *i. e.*, positive Choi operators $\{\mathcal{O}_{n:1}^{(x_{n:1})}\}$, that sum to a process tensor. That is, $\mathcal{O}_{n:1}^{\mathcal{J}_{n:1}} := \sum_{x_{n:1}} \mathcal{O}_{n:1}^{(x_{n:1})}$ satisfies both conditions in Eq. (3.23).

The individual tester elements need not satisfy the trace condition in Eq. (3.23) and represent the most general probing apparatus one could implement over a sequence of timesteps. To summarise, in decreasing levels of generality, a tester element is to a tester what a CP map is to an instrument what a POVM element is to a POVM. As an example, the (uncorrelated) instrument sequence considered in Eq. (3.26) is a 2-tester, $\mathcal{J}_{2:1} := \{\Pi_2^{(x_2)} \otimes \mathcal{O}_1^{(x_1)}\}$, since $\sum_{x_2 x_1} \Pi_2^{(x_2)} \otimes \mathcal{O}_1^{(x_1)} = \Pi_2^{\mathcal{J}_2} \otimes \mathcal{O}_1^{\mathcal{J}_1} = \mathbb{1}_2 \otimes \mathcal{O}_1^{\mathcal{J}_1}$, with $\mathcal{O}_1^{\mathcal{J}_1}$ CPTP by assumption; thus, the overall transformation implemented by this 2-tester is a two-step process comprising a CPTP transformation followed by a measurement. Here, the fact that the measurement instrument implemented does not depend on the realisation of the preparation—*i. e.*, that the respective probing instruments are independent—is reflected in the tensor product structure of each tester element.

Although the mathematical structure of the process itself and the tester used to probe it are identical, we emphasise the distinction between the underlying, *uncontrollable* process and the *controllable* tester. Lastly, to keep the language consistent with that which is most commonly employed throughout the community, we will often be lax and simply use the term ‘instrument’ to dub both instruments and testers.

In a straightforward generalisation of Eq. (3.26) to the case where a process, $\Upsilon_{n:1}$, is interrogated over n timesteps via a tester $\mathcal{J}_{n:1} = \{\mathcal{O}_{n:1}^{(x_{n:1})}\}$, one has the following spatio-temporal Born rule [34]

$$\mathbb{P}_{n:1}(x_{n:1}|\mathcal{J}_{n:1}) = \text{tr} \left[\mathcal{O}_{n:1}^{(x_{n:1})\text{T}} \Upsilon_{n:1} \right]. \quad (3.27)$$

The above equation is nothing more than a restatement of Eq. (3.11) in terms of solely the Choi operators of the process tensor and tester elements applied. This is perhaps the most pivotal expression in this thesis, concretely relating the description of a given process to the statistics observed with respect to any meaningful probing schema. Although we have arrived at it from an open systems perspective, Eq. (3.27) can be derived on axiomatic grounds via a generalisation of Gleason’s theorem applied to the process matrix [34].

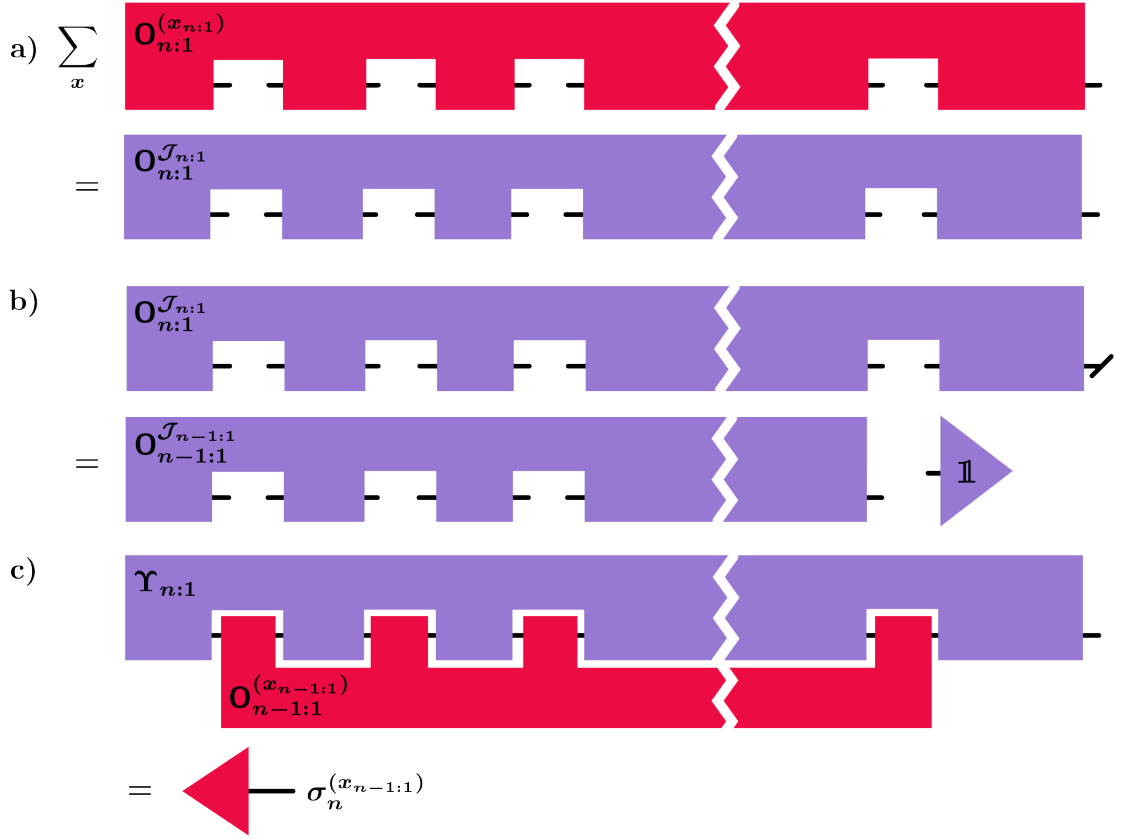


Figure 3.12: Graphical representation of a tester. A tester generalises the notion of an instrument to the case where the transformations may be correlated in time. An n -tester is a collection of positive Choi operators $\mathcal{J}_{n:1} = \{\mathcal{O}_{n:1}^{(x)}\}$ that sum to a multi-time CPTP transformation, *i. e.*, a process tensor, shown in panel **a**). The summation condition ensures some transformation happens with certainty and entails a hierarchy of constraints on the structure of the process tensor, illustrated in panel **b**). Given a process tensor $\Upsilon_{n:1}$, upon applying any sequence of operations $\mathcal{O}_{n-1:1}^{(x)}$ that constitute an $(n-1)$ -tester, some probabilistically-realised output state results $\sigma_n^{(x_{n-1:1})}$. Overall, a unit-trace state $\sigma_n = \sum_{x_{n-1:1}} \sigma_n^{(x_{n-1:1})}$ is the final output.

In clear analogy to Eq. (3.24), here tester elements correspond to a realisation of the process, *i. e.*, a sequence of potentially correlated events across the specified timesteps, and the tester itself plays an analogous role to a choice of σ -algebra over the space of possible trajectories. The spatio-temporal Born rule therefore provides the mapping to probabilities for interrogations applied to some fixed process tensor. Thus, the process tensor generalises the notion of the density operator to the temporal setting inasmuch as it contains sufficient information to deduce all possible joint probabilities corresponding to the realisation of any valid instrument sequence applied.

As mentioned previously, the system density operator at each timestep can be obtained from the process tensor, and hence the latter contains all existing descriptors

of open quantum processes as special cases. However, importantly, the process tensor additionally contains all of the multi-time information relevant to describing temporal correlations exhibited in any possible joint probability distribution that an experimenter might deduce. In this sense, quantum stochastic processes are described exactly like classical ones, as mappings from sequences of outcomes to probabilities, but with the crucial difference, that in quantum mechanics the concept of an instrument has to be inserted ‘in the middle’ to account for the possible invasiveness of measurements.

3.4 MARKOVIAN QUANTUM PROCESSES

With our ability to calculate joint statistics for quantum stochastic processes by way of the process tensor formalism and the generalised spatio-temporal Born rule, we now have an operational way to characterise Markovianity in quantum processes. Markovianity is, at its core, a multi-time statement regarding the conditional independence between the statistics observed at any given point in time and those of its entire history, given knowledge of its most recent state. We begin this section by considering how to meaningfully condition in quantum theory.

Since probabilities in quantum theory can only be calculated with respect to choices of probing instruments, a seemingly natural generalisation of Def. 2.4 is to demand that the following holds true for all times t_k on which a quantum stochastic process is defined

$$\mathbb{P}_k(x_k | \mathcal{J}_k; x_{k-1}, \mathcal{J}_{k-1}; \dots; x_1, \mathcal{J}_1) = \mathbb{P}_k(x_k | \mathcal{J}_k; x_{k-1}, \mathcal{J}_{k-1}). \quad (3.28)$$

Here $\mathbb{P}_k(x_k | \mathcal{J}_k; x_{k-1}, \mathcal{J}_{k-1}; \dots; x_1, \mathcal{J}_1)$ denotes the probability to measure x_k *given* that the outcomes x_{k-1}, \dots, x_1 were previously observed, with the instruments $\mathcal{J}_k, \dots, \mathcal{J}_1$ used to probe the dynamics. Note that on the right hand side, the probability conditioned on only the most recent outcome can only be meaningfully calculated with respect to the overall instruments applied throughout the prior history, $\mathcal{J}_{k-2}, \dots, \mathcal{J}_1$. Nonetheless, requiring Eq. (3.28) to hold for arbitrary sequences of historic instruments seems to provide a sensible notion of Markovianity, since it guarantees conditional independence of observed statistics. However, as we will now see, a subtlety arises when attempting to condition on the knowledge of previous measurement outcomes in quantum theory.

Given a quantum stochastic process, $\Upsilon_{n:1}$, we seek to calculate the conditional statistics above. Suppose the probing instruments chosen at each timestep are $\mathcal{J}_j = \{\mathcal{O}_j^{(x_j)}\}$. We must first calculate the joint statistics up to some specified time $t_k \in \{t_1, \dots, t_n\}$. By causality, the choice of future instruments \mathcal{J}_j for $t_j > t_k$ cannot influence the statistics

measured up until t_k , and it follows that the joint statistics can be computed from the reduced process tensor $\Upsilon_{k:1} = \text{tr}_{n:k^\circ} [\Upsilon_{n:1}]$ as follows

$$\mathbb{P}_{k:1}(x_k, \dots, x_1 | \mathcal{J}_k, \dots, \mathcal{J}_1) = \text{tr} \left[\left(\Pi_k^{(x_k)} \otimes \mathcal{O}_{k-1}^{(x_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right], \quad (3.29)$$

where we specify the **POVM** at t_k as $\mathcal{J}_k = \{\Pi_k^{(x_k)}\}$.

In order to condition on all of the previously observed statistics and calculate the l.h.s of Eq. (3.28), one simply wishes to divide the probabilities in Eq. (3.29) by those of the previous realisations

$$\mathbb{P}_{k-1:1}(x_{k-1}, \dots, x_1 | \mathcal{J}_{k-1}, \dots, \mathcal{J}_1) = \text{tr} \left[\left(\mathcal{O}_{k-1}^{(x_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right]. \quad (3.30)$$

However, here we immediately run into the problem that $\mathcal{O}_{k-1}^{(x_{k-1})}$ is an operator on $\text{BL}(\mathcal{H}_{k-1^\circ} \otimes \mathcal{H}_{k-1^i})$, whereas the reduced process tensor $\Upsilon_{k-1:1}$ only meaningfully provides information up to t_{k-1^i} , since $\text{tr}_{k^i} [\Upsilon_{k:1}] = \mathbb{1}_{k-1^\circ} \otimes \Upsilon_{k-1:1}$ (see Eq. (3.23)). Continuing from Eq. (3.30), we therefore have

$$\begin{aligned} & \mathbb{P}_{k-1:1}(x_{k-1}, \dots, x_1 | \mathcal{J}_{k-1}, \dots, \mathcal{J}_1) \\ &= \text{tr} \left[\text{tr}_{k-1^\circ} \left[\mathcal{O}_{k-1}^{(x_{k-1})T} \right] \otimes \left(\mathcal{O}_{k-2}^{(x_{k-2})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right]. \end{aligned} \quad (3.31)$$

The operator $\text{tr}_{k-1^\circ} \left[\mathcal{O}_{k-1}^{(x_{k-1})T} \right]$ necessarily does not capture any information about the output state of the interrogation at t_{k-1} , which could crucially impact the measurement outcomes observed at t_k . For instance, one can envisage situations where the experimenter performs a measurement $\{\Pi_{k-1^i}^{(r_{k-1})}\}$ and then prepares one of a set of (subnormalised) states with some outcome-dependant probability $\{\rho_{k-1^\circ}^{(s_{k-1}, r_{k-1})}\}$ to feed forward, implementing the instrument $\mathcal{J}_{k-1} = \{\Pi_{k-1^i}^{(r_{k-1})} \otimes \rho_{k-1^\circ}^{(s_{k-1}, r_{k-1})}\}$. Here, Eq. (3.29) would be sensitive to the choices of prepared states, whereas Eq. (3.31) would not. Thus, to divide the former by the latter would certainly not provide a meaningful notion of conditioning.

Indeed, conditioning necessarily breaks the information flow between past and future, while a generic operation need not. To ameliorate this issue we require the concept of a *causal break*, which is a particular type of instrument that clearly separates information about the inputs and outputs of its transformations. Intuitively, it corresponds to the scenario where an experimenter makes a measurement and then—in contrast to the example above—prepares a fresh, *independent* known state. A causal break instrument $\mathcal{J}_j = \{\mathcal{B}_j^{(x_j)}\}$ comprises elements (see Fig. 3.13)

$$\mathcal{B}_j^{(x_j)} := \rho_{j^\circ}^{(s_j)} \otimes \Pi_{j^i}^{(r_j)}, \quad (3.32)$$

where the labels x_j of the overall **CP** maps in the causal break instrument are split into those labelling the measurement outcome, r_j , and the subsequent independent reparation, s_j . Importantly, the Choi operators on the input and output spaces are completely uncorrelated; more generally, any operation whose output is independent of its input in the sense above constitutes a causal break.

Returning our consideration to the calculation of conditional statistics, we see that demanding the most recent instrument applied at time t_{k-1} to be a causal break immediately resolves the problem, allowing us to meaningfully condition on prior knowledge obtained through probing a quantum stochastic process. Now, Eq. (3.31) reads

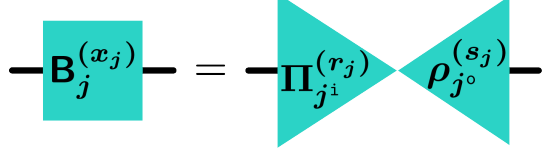


Figure 3.13: *Causal break.* A causal break at time t_j is a **POVM** $\{\Pi_{j^i}^{(r_j)}\}$ followed by an independent reparation into one of a known set of states $\{\rho_{j^o}^{(s_j)}\}$. This breaks any information/memory flow on the level of the system.

$$\begin{aligned} & \mathbb{P}_{k-1:1}(s_{k-1}, r_{k-1}, \dots, x_1 | \mathcal{J}_{k-1}, \dots, \mathcal{J}_1) \\ &= \mathbb{P}(s_{k-1} | \mathcal{J}_{k-1}) \operatorname{tr} \left[\left(\Pi_{k-1}^{(r_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right] \\ &= \operatorname{tr} \left[\left(\rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1}^{(r_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right], \end{aligned} \quad (3.33)$$

where $\mathbb{P}(s_{k-1} | \mathcal{J}_{k-1}) = \operatorname{tr} [\rho_{k-1^o}^{(s_{k-1})}]$ is simply the probability that the experimenter chooses to prepare the particular state $\rho_{k-1^o}^{(s_{k-1})}$, which is independent of the process. Similarly, Eq. (3.29) reads

$$\begin{aligned} & \mathbb{P}_{k:1}(x_k, s_{k-1}, r_{k-1}, \dots, x_1 | \mathcal{J}_k, \dots, \mathcal{J}_1) \\ &= \operatorname{tr} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1}^{(r_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right], \end{aligned} \quad (3.34)$$

By breaking the flow of information on the level of the system, a causal break allows us to compare like with like for each run of the experiment and therefore meaningfully condition in the quantum setting. Dividing Eq. (3.34) by Eq. (3.33) gives the probability to observe outcome x_k given knowledge of all previous outcomes for the instruments chosen, *i. e.*, the conditional probability $\mathbb{P}_k(x_k | \mathcal{J}_k; s_{k-1}, r_{k-1}, \mathcal{J}_{k-1}, \dots, x_1, \mathcal{J}_1)$.

Intuitively, conceptual introduction of the causal break is necessary to demarcate those memory effects that arise from the process, *i. e.*, are the manifestation of environmental influence, rather than as a result of the applied control operations feeding-forward information about the system's past. As the fresh preparation is independent of the previous measurement outcome, a causal break ensures that no temporal correlations

are transmitted through the system itself, breaking the causal link on the level of the system between the past $t_i \leq t_{k-1}^i$ and the future $t_j \geq t_{k-1}^o$. By the very nature of the causal break, the system itself cannot possibly transmit information beyond time t_{k-1} concerning the measurement outcome associated to $\Pi_{k-1}^{(r_{k-1})}$ or the operations implemented throughout its earlier history. In light of this, it is sensible to slightly revise Eq. (3.28) to demand that a Markovian quantum process should display statistics that are conditionally independent of all historic outcomes *and* their instruments, including the measurement outcome r_{k-1} of the causal break realised at t_{k-1}^i , given knowledge of the probabilities with which the output states of the causal break are prepared.

Definition 3.4 (Quantum Markov condition [49]). Consider a causal break \mathcal{J}_{k-1} at timestep t_{k-1} in which the measurement outcome is labelled by r_{k-1} and the prepared state by s_{k-1} . A quantum stochastic process is Markovian when the statistics observed with respect to an arbitrary measurement instrument \mathcal{J}_k at t_k are conditionally independent of all historic outcomes $r_{k-1}, x_{k-2}, \dots, x_1$ for any possible historic instruments $\mathcal{J}_{k-2}, \dots, \mathcal{J}_1$, given knowledge of only the state prepared during the causal break:¹⁰

$$\mathbb{P}_k(x_k | \mathcal{J}_k; s_{k-1}, r_{k-1}, \mathcal{J}_{k-1}; \dots; x_1, \mathcal{J}_1) = \mathbb{P}_k(x_k | \mathcal{J}_k; s_{k-1}, \mathcal{J}_{k-1}). \quad (3.35)$$

The definition above lends itself to an operational criterion for classifying Markovian quantum processes: a quantum process is non-Markovian iff there exists at least two different historic testers, $\mathcal{J}_{k-1^i:1} = \{\mathcal{O}_{k-1^i:1}^{(x_{k-1^i:1})}\}$ and $\mathcal{J}'_{k-1^i:1} = \{\mathcal{O}'_{k-1^i:1}^{(x'_{k-1^i:1})}\}$ such that, for some choice of preparation in the causal break, the density operator of the system at t_k is different, implying the violation of Eq. (3.35). This provides a valid and unambiguous method to witness memory effects. Conversely, fixing the preparation in the causal break and finding the subsequent density operator to be constant for all linearly independent historic tester elements implies that the process is Markovian.

Although we have made extensive use of the notion of causal breaks to provide an operational picture, the question of whether or not a process is Markovian does not depend on it. Indeed, Def. 3.4 directly leads to an unambiguous constraint on the structure of the process tensor itself, which is depicted in Fig. 3.14.

¹⁰ Interestingly, the same definition was introduced in Ref. [17], at roughly the same time as ours, using the process matrix formalism in the context of quantum causal modelling.

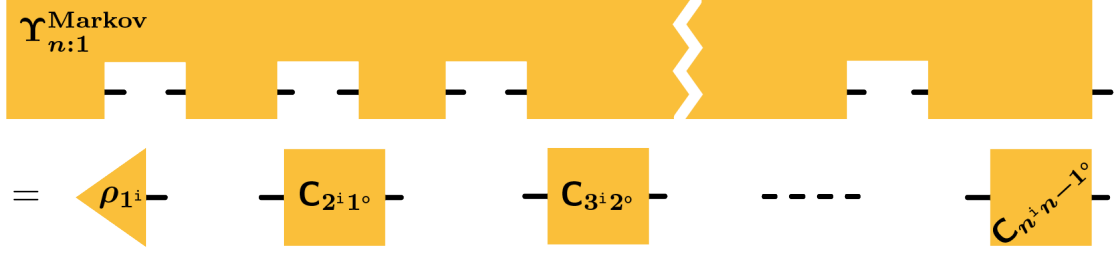


Figure 3.14: *Markovian process tensor.* The Choi operator of a Markovian process tensor is a tensor product of **CPTP** maps, representing uncorrelated evolution between timesteps.

Theorem 3.5 (Markovian quantum process). *A process tensor represents a Markovian process iff it has the following tensor product structure:*

$$\Upsilon_{n:1}^{\text{Markov}} = \mathbf{C}_{n^i n-1^o} \otimes \dots \otimes \mathbf{C}_{2^i 1^o} \otimes \rho_{1^i}, \quad (3.36)$$

where each $\mathbf{C}_{j^i j-1^o} \in \text{BL}(\mathcal{H}_{j^i} \otimes \mathcal{H}_{j-1^o})$ is the Choi operator of a **CPTP** map and $\rho_{1^i} \in \text{BL}(\mathcal{H}_{1^i})$ is the initial average system state.

Proof. The l.h.s of Eq. (3.35) is calculated as

$$\begin{aligned} & \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathbf{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right]}{\text{tr}_{k-1:1} \left[\left(\rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathbf{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right]} \\ &= \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathbf{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right]}{\mathbb{P}(s_{k-1} | \mathcal{J}_{k-1}) \text{tr}_{k-1:1} \left[\left(\Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathbf{O}_1^{(x_1)} \right)^T \Upsilon_{k-1:1} \right]} \\ &= \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathbf{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right]}{\mathbb{P}(s_{k-1} | \mathcal{J}_{k-1}) \mathbb{P}_{k-1:1}(r_{k-1}, \dots, x_1 | \mathcal{J}_{k-1}, \dots, \mathcal{J}_1)}. \end{aligned} \quad (3.37)$$

The r.h.s cannot be calculated readily by the spatio-temporal Born rule without first considering some historic instruments $\mathcal{J}_{k-2}, \dots, \mathcal{J}_1$. At all of these timesteps, we use their associated **CPTP** maps to calculate the probabilities; these are arbitrary and will eventually be varied freely to prove the result. Note also that summing over all of the most recent measurement outcomes, r_{k-1} , leads to the identity operator on the input space, $\mathbb{1}_{k-1^i}$. We thus have

$$\begin{aligned} & \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \mathbb{1}_{k-1^i} \otimes \dots \otimes \mathbf{O}_1^{\mathcal{J}_1} \right)^T \Upsilon_{k:1} \right]}{\text{tr}_{k-1:1} \left[\left(\rho_{k-1^o}^{(s_{k-1})} \otimes \mathbb{1}_{k-1^i} \otimes \dots \otimes \mathbf{O}_1^{\mathcal{J}_1} \right)^T \Upsilon_{k-1:1} \right]} \\ &= \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^o}^{(s_{k-1})} \otimes \mathbb{1}_{k-1^i} \otimes \dots \otimes \mathbf{O}_1^{\mathcal{J}_1} \right)^T \Upsilon_{k:1} \right]}{\mathbb{P}(s_{k-1} | \mathcal{J}_{k-1})}. \end{aligned} \quad (3.38)$$

Comparing Eqs. (3.37) and (3.38) shows that for a Markovian process we have

$$\begin{aligned} & \text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^\circ}^{(s_{k-1})} \otimes \mathbb{1}_{k-1^i} \otimes \dots \otimes \mathcal{O}_1^{\mathcal{J}_1} \right)^T \Upsilon_{k:1} \right] \\ &= \frac{\text{tr}_{k:1} \left[\left(\Pi_k^{(x_k)} \otimes \rho_{k-1^\circ}^{(s_{k-1})} \otimes \Pi_{k-1^i}^{(r_{k-1})} \otimes \dots \otimes \mathcal{O}_1^{(x_1)} \right)^T \Upsilon_{k:1} \right]}{\mathbb{P}_{k-1:1}(r_{k-1}, \dots, x_1 | \mathcal{J}_{k-1}, \dots, \mathcal{J}_1)}. \end{aligned} \quad (3.39)$$

This is clearly satisfied by the Markov process tensor structure given in Eq. (3.36), as can be seen by direct insertion. In the coming chapters, we will use similar arguments to derive structural properties of processes with finite Markov order.

In the converse direction, recall that this condition must hold for arbitrary choices of historical instruments. For any specific x_k and s_{k-1} , we can consider varying the outcomes for a fixed sequence of historical instruments, leading to the r.h.s changing but not the l.h.s; on the other hand, we can consider a fixed sequence of CP maps and vary the overall instruments, leading to the l.h.s changing but not the r.h.s. The only remaining way that Eq. (3.39) can be satisfied is if the process tensor itself splits as a tensor product $\mathbb{C}_{k^i:k-1^\circ} \otimes \Upsilon_{k-1^i:1}$. That $\mathbb{C}_{k^i:k-1^\circ}$ is a CPTP map follows directly from the properties of the process tensor. Lastly, repeating the argument for all times t_k on which the process is described yields Eq. (3.36). \square

Through Theorem 3.5 we finally have an intrinsic characterisation of Markovian quantum processes. Intuitively, the Choi operator of a Markovian process only displays correlations between adjacent times, meaning that the only temporal correlations in the process are those between preparations and the subsequent measurements at the next timestep. This is analogous to the fact that Markovian processes have non-trivial transition probabilities between adjacent timesteps only. Because the process tensor description includes the system density operator at all timesteps, the strict condition of Def. 3.4 presents both a unification and generalisation of previous definitions of Markovianity throughout the literature. For instance, this result confirms the well-known fact that Markovian processes are divisible; however, the converse direction fails to hold, as divisibility is not a strict enough criteria to force the process tensor into the required tensor product structure [54].

Moreover, the process tensor formalism can be used to explicitly calculate any of the measures of non-Markovianity introduced in the literature; several examples of memory effects that are not detected by conventional approaches but are within our operational framework are presented in the Supplemental Material of Ref. [49]. Indeed, from the structure of Markovian processes in Eq. (3.36), it is straightforward to define a measure that quantifies the amount of non-Markovianity, evaluated in terms of the distance

between the Choi operator of the process at hand and the nearest Markovian Choi operator with respect to any suitable distance metric, \mathbb{D} [49], as

$$\mathcal{N}(\Upsilon_{n:1}) := \min_{\Upsilon_{n:1}^{\text{Markov}}} \mathbb{D}(\Upsilon_{n:1} \| \Upsilon_{n:1}^{\text{Markov}}). \quad (3.40)$$

For instance, the measure chosen could be in terms of the relative entropy, as we will use in Subsection 6.1.2, or the Schatten 1-norm (trace distance), as used in Ref. [45] to study the distinguishability between Markovian and non-Markovian processes.

To summarise the story so far, we have here shown how the process tensor formalism naturally leads to a necessary and sufficient condition for a quantum process to be considered Markovian. This condition is stricter than those based on traditional approaches since the process tensor accounts for potential multi-time memory effects. By shifting to an operational description of stochastic processes, we have overcome the problem of formalism that led many to believe that there could be no unique criteria of Markovianity in open quantum dynamics. Indeed, the constraint on the structure of Markovian processes provides an intrinsic characterisation of a memoryless process.

We finally have a clear picture of what it means for a quantum process to have memory. As in the classical case, the description of generic processes grows rapidly in terms of complexity with the length of the memory; fortunately for us, many processes found in nature display an effectively finite-length memory. The remainder of this thesis presents novel work pertaining to characterising, quantifying and exploiting by way of simulation such finite memory processes.

Part III

UNSPOKEN WORDS

Not the weekend dance where you two-step to music you've heard before
and always know... but the Daily Dance with the wilder step, to a tune as
soundless as the eelgrass tune, to an echo of a song, or a song still unechoed.

— Ken Kesey, *Sometimes a Great Notion*.

MEMORY LENGTH

WE HAVE SEEN THAT FOR MARKOVIAN PROCESSES, once the state is known, the future is independent from the past. A central obstacle in predicting the future of a dynamical system is to understand how much of the system’s past acts as a relevant influence, which crucially determines the resources required for simulation. Although non-Markovianity is the rule rather than the exception when it comes to stochastic processes [25], their characterisation is resource intensive, both in the classical and quantum setting. Fortunately for us, in reality, even the most complex processes have a finite effective range—that is, a cause can only noticeably affect the future for a certain length of time—providing a natural notion of memory length, formally captured by the concept of Markov order.

A process with finite Markov order allows for an efficient description compared to the general non-Markovian case, as it only requires a portion of the history to optimally predict the future. Indeed, the substantial reduction in modelling complexity that ensues underpins the often-invoked higher-order Markov models to simulate non-Markovian stochastic processes [58–60]. The motivation for understanding processes with finite Markov order is thus two-fold: on the one hand, they exhibit genuine memory effects; on the other, these effects are constrained in time, rendering their description tractable [25].

Akin to the joint probability distribution describing a generic classical stochastic process, the description of a general quantum stochastic process has an exponentially increasing complexity with respect to the length of history that must be retained, with the added complication that all possible sequences of external interventions need to be accounted for. This naturally begs the question that will be the central focus of the present chapter: *are there quantum processes with finite-length memory, and hence a significantly reduced complexity?* We have already seen that the process tensor formalism naturally lends itself to a proper classification of Markovianity in the quantum regime; the first line of pursuit here is a similar generalisation of Markov order.

We will see that a reasonable and logically sound definition of Markov order for quantum processes can be formulated in terms of a constraint on the structure of the process tensor. However, in quantum mechanics, we have no choice here but to propose an *instrument-specific* notion of Markov order. Intuitively, the prospective definition implies that with respect to a given instrument sequence specified across ℓ timesteps of the memory, any deducible statistics across the history and future timesteps (*i. e.*, for arbitrary choices of history and future probing instruments) are guaranteed to be conditionally independent for each realisation of the memory tester in question. The logical consequence of this characterisation is that—perhaps unsurprisingly—quantum processes with memory exhibit different memory effects when probed in different ways. The aim of the present chapter is to precisely formulate and justify this notion of quantum Markov order which formally captures such behaviour.

Following the introduction of this instrument-specific notion of quantum Markov order, we provide two crucial pieces of supporting evidence of its plausibility. Firstly, that the definition reduces to the classical one in an appropriate fashion. Indeed, classical Markov order is implicitly defined in an instrument-specific way—with respect to the statistics deduced via sequences of *sharp* measurements. As soon as we allow for *fuzzy* measurements that coarse-grain information—or, more broadly, any active experimental interventions—the length of the memory depends on the instruments used to interrogate the process at hand, just like in the quantum case.

Within quantum theory, however, there exist a much richer array of instruments that an experimenter might apply to the system of interest than are available in the classical world, including generalised measurements, unitary transformations, and temporally correlated sequences implemented via a genuinely quantum control (*i. e.*, the most general testers). The second key result of this chapter shows that demanding a quantum process to have finite Markov order with respect to all possible instruments immediately trivialises the theory into only admitting Markovian processes. This result, in turn, leads to the realisation that any quantum process with memory has infinite Markov order with respect to generic instrument sequences and therefore requires the complete description. This fact notwithstanding, for many practical purposes, such a comprehensive understanding of the memory is often either unachievable—due to experimental or computational limitations—or otherwise unnecessary with regard to the specificity of the experimenter’s concerns.

Indeed, the landscape of memory effects in quantum processes is vast and ripe for exploration, as will become apparent in Chapter 5 where we analyse the structure of

quantum processes with finite Markov order with respect to familiar classes of history-blocking instruments. Finally, note that many of the concepts introduced, results derived and discussion that entails in the coming chapter are presented in Refs. [1, 2].

4.1 QUANTUM MARKOV ORDER

Before we begin the main discussion of this chapter, it is worthwhile to briefly consider some alternative suggestions that arise naturally when considering an extension of Markov order into quantum mechanics. This aside should both illuminate some crucial details that must be addressed in the temporal setting and distinguish our approach from a number of related concepts studied throughout the community.

Recall that, as discussed in Subsection 2.1.6, there are three equivalent characterisations for classical Markov order:

1. The joint probability distribution across the future and history, given events in the memory, factorises: $\mathbb{P}_{FH}(x_F, x_H|x_M) = \mathbb{P}_F(x_F|x_M)\mathbb{P}_H(x_H|x_M)$.
2. There exists a stochastic map $\mathcal{R}_{M \rightarrow FM}$ acting only on the memory block that can ‘recover’ the future statistics correctly: $\mathbb{P}_{FMH} = \mathcal{R}_{M \rightarrow FM}(\mathbb{P}_{MH})$.
3. The CMI between the history and the future given the memory vanishes: $I^{\text{cl}}(F : H|M) = 0$.

For some time, the recovery map has featured in the quantum information literature: here, quantum Markov chains have been defined as tripartite quantum *states* $\rho^{ABC} \in \text{BL}(\mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C)$ satisfying a quantum generalisation of recoverability [136–142]. Equivalently, quantum Markov chains have been defined as those states with vanishing quantum CMI, which is identical to those that saturate the strong subadditivity inequality [136, 143]. In the case of the former, the natural extension posits the existence of a CPTP recovery map $\mathcal{R}_{B \rightarrow BC} : \text{BL}(\mathcal{H}^B) \rightarrow \text{BL}(\mathcal{H}^B \otimes \mathcal{H}^C)$ such that $\rho^{ABC} = \mathcal{R}_{B \rightarrow BC}(\rho^{AB})$. Interestingly, unlike in the classical setting, the relation between quantum recoverability and the vanishing of the quantum CMI is not at all obvious; the proof of their equivalence is in fact a highly celebrated result [136, 137, 144]. Additionally, in Ref. [137] the authors introduced an important structural characterisation of the set of states with vanishing quantum CMI, which are therefore recoverable. The main theorem highlighting the existence of a projective measurement on the B subsystem such that, for each outcome of the measurement, the AC subsystem is steered into an uncorrelated state, providing a link by way of analogy to the quantum counterpart of the first concept listed above.

However, at first sight, it is unclear how such characterisations relate to *temporal processes*, where an experimenter has access to an evolving quantum system across multiple timesteps. As previously discussed, a number of concerns arise in the temporal setting because of the necessity to track the state of the system throughout the process in a meaningful way. Due to the breakdown of the [KET](#) for quantum processes on the level of joint probability distributions, there has heretofore been no sensible way to develop a generalisation of Eq. (2.19)—arguably the most fundamental definition of Markov order—to the temporal realm. The process tensor formalism, by way of encoding all possible joint distributions for all possible sequences of operations in time and permitting calculation of conditional statistics, provides a straightforward and unambiguous way to define Markov order for quantum processes. We will return to the relation between quantum Markov order and vanishing quantum [CMI](#) in the coming chapter.

4.1.1 Instrument-specific Quantum Markov Order

A natural approach to defining Markov order for quantum processes is to require that for every possible way of probing the process, the corresponding joint probability distribution satisfies the classical Markov condition. In other words, just as was the case for Markovian quantum processes, we fix the choice of instrument on the timesteps associated to the memory block and check for conditional independence between the history and the future. Thus, a sensible demand of a quantum process with finite-length memory is that *any* future statistics deducible—*i. e.*, no matter which future instruments are chosen—are conditionally independent of *any* historical instruments applied and their measurement outcomes, given knowledge of a length- ℓ instrument sequence applied to the memory block. Grouping together the timesteps as in Def. 2.5, we therefore define quantum Markov order as follows.

Definition 4.1 (Quantum Markov order [1]). A quantum stochastic process has Markov order $\ell = |M|$ with respect to an instrument \mathcal{J}_M when, for all possible instruments \mathcal{J}_H and \mathcal{J}_F on the history and future, the following is satisfied:

$$\mathbb{P}_F(x_F|\mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) = \mathbb{P}_F(x_F|\mathcal{J}_F; x_M, \mathcal{J}_M), \quad (4.1)$$

where $\mathbb{P}_F(x_F|\mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H)$ denotes the probability to measure x_F given that outcomes x_M and x_H were previously observed, with the instruments $\mathcal{J}_F, \mathcal{J}_M$ and \mathcal{J}_H used to probe the system dynamics.

The above definition is intuitive: we can imagine an experimenter who implements the instruments $\mathcal{J}_F, \mathcal{J}_M$ and \mathcal{J}_H . They would consider the process to display finite memory if, for any choice of \mathcal{J}_F and \mathcal{J}_H , the statistics observed on F and H are conditionally independent with respect to a *fixed* instrument \mathcal{J}_M . For any instruments the experimenter might use to probe the future evolution of the system, the full statistics is then completely determined by the instruments and outcomes of the most recent ℓ timesteps. Equivalently, given knowledge of the outcomes of the instrument across the past ℓ timesteps, the process governing the future is uncorrelated with that governing the history, guaranteeing that any possible statistics one might deduce on the history and future timesteps are independent in each run of the experiment. Again, it is crucial to note here that the notion of independence here is a conditional one.

Quantum Markov order thus defined is *instrument-specific* in the sense that it fixes the instrument sequence applied across M in order to sensibly calculate conditional statistics, stipulating that meaningful statements regarding memory length must be presented with the caveat regarding the instrument of choice. This instrument-specific definition boils down to the classical one if all instruments are fixed, sharp and classical, as we will discuss in Section 4.2. As we will soon see, demanding a quantum process to have finite Markov order for all possible instruments admissible in quantum theory is too restrictive, immediately trivialising the theory. Such a demand is equivalent to asking the process to have finite Markov order with respect to an arbitrary length- ℓ sequence of causal breaks, since they form a basis for the space of valid instrument sequences on the memory block. As this constraint only permits Markovian processes, we are forced to consider the more general instrument-specific scenario in order to meaningfully describe processes with memory, as we do here. Nonetheless, a process can have finite quantum Markov order with respect to an entire family of instruments.

In fact, we have already seen an example of this: Markovian quantum processes have Markov order 1 for all instruments consisting of only causal breaks. In contrast, when memory plays a non-negligible role in the evolution, operations performed on the system generally impact the environment significantly, which, in turn, inevitably influences the future dynamics. Thus, Markov order ℓ does not only refer to processes with conditional independence across a length ℓ sequence of causal breaks, since these may serve to ‘open a pathway’ for the history to influence the future via environmental conditioning. Indeed, this is a special case of the instrument-specific definition of quantum Markov order, as we will see in Chapter 5, where we will study a number of processes with finite-length

memory with respect to natural families of instruments, including those comprising unitary operations and extended sequences of causal breaks.

4.1.2 Finite Memory Constraint on the Process Tensor

We saw in Section 3.4 that Markovian quantum processes have a simple tensor product structure, and it is hence an interesting question to explore the implications of Def. 4.1 on the structure of the process tensor. To this end, consider a stochastic process described by the process tensor Υ_{FMH} and denote the probing instrument sequences by $\mathcal{J}_X = \{\mathcal{O}_X^{(x_X)}\}$. Then, the joint probability distribution over statistics observed is calculated via the generalised Born rule as

$$\begin{aligned} \mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H) \\ = \text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right]. \end{aligned} \quad (4.2)$$

From this expression, we can calculate well-defined conditional probabilities

$$\begin{aligned} \mathbb{P}_{FMH}(x_F | \mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) \\ = \frac{\text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right]}{\text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{MH} \right]}, \end{aligned} \quad (4.3)$$

where $\Upsilon_{MH} := \frac{1}{d_{F^\circ}} \text{tr}_F [\Upsilon_{FMH}]$ with $d_{F^\circ} := \prod_{j=k}^n d_{j^\circ}$ denoting the total dimension of the output spaces associated to F .

In Appendix C.1, we show that Def. 4.1 implies the following product structure on the process tensor, represented graphically in Fig. 4.1 [1]

$$\tilde{\Upsilon}_{FH}^{(x_M)} := \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right] = \Upsilon_F^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)} \quad \forall \mathcal{O}_M^{(x_M)} \in \mathcal{J}_M. \quad (4.4)$$

Here, the conditional future process, $\Upsilon_F^{(x_M)}$, is described by a proper process tensor whereas the unnormalised description of the historic process, $\tilde{\Upsilon}_H^{(x_M)}$, is garnished with a tilde to denote that it is an element of a tester, *i. e.*, when summed over all outcomes of the memory instrument, it yields a proper process tensor. We adhere to this notation throughout and will return to discuss this point shortly. Crucially, finite Markov order does not force the process tensor into an overall tensor product structure, but only conditionally; in analogy to Eq. (2.19), the conditional history and future processes are independent for each realisation of the instrument applied.

A few further comments are in order. Firstly, if Eq. (4.4) is satisfied, we say that the process has Markov order- ℓ *with respect to the history-blocking instrument sequence*,

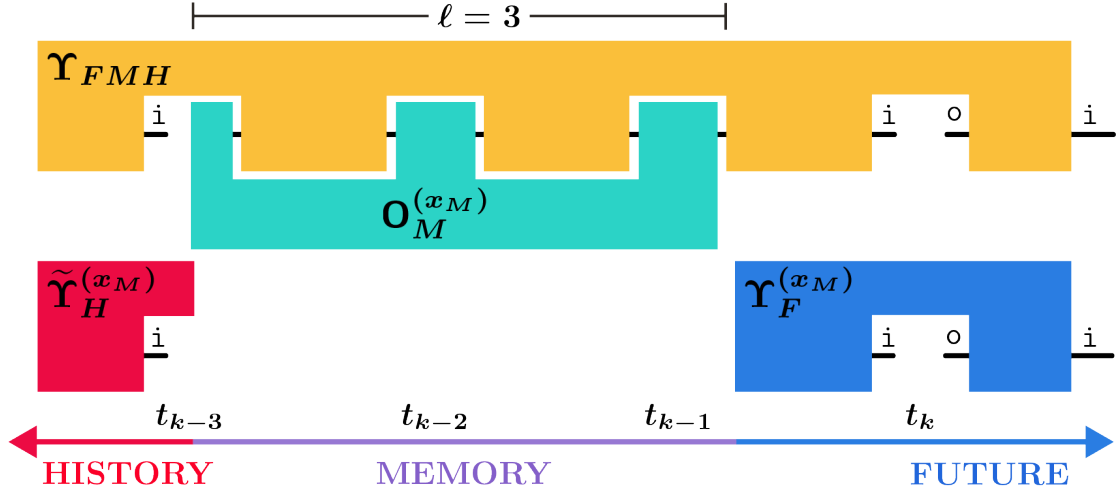


Figure 4.1: *Instrument-specific quantum Markov order.* An instrument sequence \mathcal{J}_M , comprising (temporally correlated) CP operations $\{\mathcal{O}_M^{(x_M)}\}$ (green) across a sequence of timesteps of length $\ell = 3$, is applied to a process Υ_{FMH} . The process is said to have Markov order ℓ with respect to this instrument sequence if, for each possible realisation of the instrument, labelled by x_M , the history (red, $\tilde{\Upsilon}_H^{(x_M)}$) and future (blue, $\Upsilon_F^{(x_M)}$) parts of the process are rendered conditionally independent. Note that here, the memory length of 3 timesteps corresponds to 5 Hilbert spaces on which the process tensor is defined (from t_{k-3}° to t_{k-1}°).

\mathcal{J}_M . The fact that the process is rendered conditionally independent for each realisation of the instrument—which is, overall, a *deterministic* implementation—means that an experimenter is *guaranteed* to block the effect of the history on the future upon its application (given that they know the outcome). More generally, there may exist individual operation sequences, *i. e.*, tester elements, that block the history. However, in contrast to instruments, since these could only be implemented with some probability, such operations act to *probabilistically* render the future and history conditionally independent. In this thesis, we focus only on overall *deterministic* history-blocking sequences, in which every constituent tester element in an instrument sequence acts to block the effect of history, guaranteeing conditional independence for each run of the experiment.

Secondly, satisfaction of Eq. (4.4) guarantees the conditional independence of *any* possible statistics one could obtain on the future and history given knowledge of the history-blocking instrument sequence. We can see clearly that the mutual information between the conditional future and history processes for any realisation of \mathcal{J}_M vanishes, since they are of product form

$$I(F : H)_{x_M} := S(\Upsilon_F^{(x_M)}) + S(\tilde{\Upsilon}_H^{(x_M)}) - S(\tilde{\Upsilon}_{FH}^{(x_M)}) = 0, \quad (4.5)$$

where $I(F : H)_{x_M}$ denotes the mutual information between the history and future processes given that the sequence corresponding to outcome x_M was realised, and $S(\bullet)$ is the von Neumann entropy.¹ The mutual information upper-bounds all possible correlations between arbitrary observables on F and H , and thus its vanishing implies the temporal regions of the future and history are uncorrelated given knowledge of each outcome x_M [145].

Thirdly, note that the conditional future process is a proper process tensor by construction, whilst the conditional history process represents an element of a tester. The former point arises from the fact that for each realisation of the instrument on the memory block, *some* fixed future evolution occurs. Mathematically, this can be seen by considering that the hierarchy of trace conditions in Eq. (3.23) holds for all x_M in Eq. (4.4), as each successive partial trace applied from the future backwards ends up acting only upon Υ_{FMH} . On the other hand, the fact that Υ_{FMH} is a proper process tensor no longer ensures the hierarchy of trace conditions hold for each conditional positive object $\tilde{\Upsilon}_H^{(x_M)}$ prior to the memory, as the multiplication with $\mathcal{O}_M^{(x_M)\text{T}}$ prevents us from continuing to trace back through the hierarchy; each realisation of the instrument in question amounts to a post-selection [1, 28, 35]. Nonetheless, since the overall implementation of the memory instrument is deterministic, when all possible outcomes are summed over, the history is described by a proper process tensor. In other words, on average, the history is described by a positive semidefinite Choi operator satisfying Eq. (3.23); however, the individual tester elements corresponding to each conditional outcome need not obey the causality conditions. In the special cases where they do, the probability of realising the associated sequence of outcomes of the history-blocking instrument can be extracted from the conditional history process, *i. e.*, we can write $\tilde{\Upsilon}_{FH}^{(x_M)} = \mathbb{P}(x_M | \mathcal{J}_M) \Upsilon_F^{(x_M)} \otimes \Upsilon_H^{(x_M)}$, with $\Upsilon_H^{(x_M)}$ a proper process tensor for each x_M , as we do at some points throughout this thesis.

Lastly, it is important to distinguish which input and output spaces constitute a memory block of length ℓ , as there is evidently some innocuous source of potential confusion in the way that we have discussed memory length so far. Any such block may begin and end on either the input or output Hilbert spaces associated with timesteps $t_{k-\ell}$ and t_{k-1} respectively (see Fig. 4.1: here, the memory block illustrated begins on the output space at $t_{k-\ell}$ and ends at the output space of t_{k-1}). To mitigate any possible confusion, one could describe the memory length in terms of the number of Hilbert

¹ Since entropies are only meaningfully-defined for normalised objects, any entropic quantity is calculated using the normalised process tensor, *i. e.*, $\Upsilon / \text{tr}[\Upsilon]$.

spaces comprising the memory block. However, intuitively, an experimenter is surely more concerned with the question of *how long ago something happened* rather than *how many Hilbert spaces back something happened*. As such, and to ease notation, we refrain from labelling each of these cases distinctly and simply consider the memory length to be the number of timesteps across which the history-blocking sequence spans. We will provide visual representations of each example considered throughout this thesis to elucidate how the memory block is defined in each case; for instance, in Fig. 4.1, the memory length is 3 timesteps rather than 5 Hilbert spaces.

As we have previously alluded to, a quantum process can have finite Markov order with respect to a family of instruments. Demanding that Eq. (4.1) holds with respect to all possible instrument sequences immediately trivialises the theory.

Theorem 4.2 (Only Markovian processes have finite Markov order for all instruments [1]). *The only quantum processes with finite Markov order with respect to all possible instruments are Markovian.*

The proof follows along the lines of that of Theorem 3.5 and is given in Appendix C.2. We make use of the fact that the set of tester elements forms a vector space to show that the only processes satisfying Eq. (4.4) for all possible instruments have trivial Markov order, *i. e.*, either $\ell = 1$ or 0 . Specifically, we show that if a process has finite Markov order with respect to a complete basis of CP maps on M , it cannot have the same Markov order with respect to any non-trivial linear combination of them.

Returning to the subtlety in defining memory length in quantum mechanics (*i. e.*, with respect to Hilbert spaces or with respect to timesteps), note that when a 1-step memory block is specified, there are two cases to be considered: the first being where the memory begins and ends on the most recent output Hilbert space (*i. e.*, $M = \{t_{k-1}^o\}$), and the second where it extends over both Hilbert spaces associated to the previous timestep (*i. e.*, $M = \{t_{k-1}^o, t_{k-1}^i\}$). Demanding the finite Markov order condition to hold for all instruments in the former case is equivalent to the Markovianity condition of Def. 3.4, and since all instruments on only a single output Hilbert space correspond to a unit-trace quantum state, it leads directly to the structure of Eq. (3.36).

On the other hand, requiring finite Markov order to hold for all possible instruments in the latter scenario is an even stricter demand than Markovianity, which only dictates that the condition hold for the family of causal break instruments (see Eq. (3.32)). If the future process is to be rendered independent from the history for all instruments spanning *both* Hilbert spaces of the previous timestep, the process is forced into the

following super-Markovian structure, where the future evolution is independent of even the most recent state preparation (like a fair coin toss)

$$\Upsilon_{n:1}^{\text{S-Markov}} = \rho_{n^i} \otimes \mathbb{1}_{n-1^o} \otimes \rho_{n-1^i} \otimes \dots \otimes \mathbb{1}_{1^o} \otimes \rho_{1^i}. \quad (4.6)$$

In either case, demanding finite Markov order with respect to all instruments on a memory block ending at some timestep t_{k-1} imposes a tensor product structure on the process tensor (either between t_{k-1}^i and t_{k-1}^o or between t_{k-1}^o and t_k^i) such that the future and history processes are independent of the measurement outcome; requiring this to hold at all timesteps enforces a Markovian or super-Markovian structure, thereby trivialising the theory by only admitting memoryless processes. This result immediately implies the following property for quantum stochastic processes.

Remark. Any non-Markovian quantum process has infinite Markov order with respect to a generic instrument sequence. Furthermore, applying random instruments can almost always witness memory effects.

This is because, if the process uses any memory at all, there always exists some instrument such that Eq. (4.4) does not hold (it only holds for all instruments if the process is Markovian). Applying random choices of operations on a finite-dimensional Hilbert space will eventually span the space and an experimenter can then find the suitable operation to witness memory effects.

It is clear that demanding Def. 4.1 to hold for all instruments is a very strong condition, as it requires the statistics observed by an experimenter to satisfy the Markov order- ℓ property *no matter* how they measure realisations of the process. Theorem 4.2 shows that this requirement is too restrictive in the quantum case. Thus, we have no choice but to characterise Markov order for quantum processes in an instrument-specific way. In light of this analysis, we see that classical Markov order is also defined in an inherently instrument-specific way, as it assumes the ability for an experimenter to sharply measure realisations of events. In many practical cases of interest, however, this assumption is either not satisfied due to experimental limitations that lead to noisy measurements [146–152], or insufficient to capture the scenario at hand, such as in causal modelling [16, 17]. In any framework that allows for active interventions, a vast array of memory effects can be captured by probing the system with different instruments, as we now consider.

4.2 RELATION TO CLASSICAL MARKOV ORDER

Importantly, quantum Markov order reduces to the classical Markov order statement in the correct limit: when the stochastic process at hand is entirely characterised by a joint probability distribution and the experimenter probes it appropriately via classical means.

Classical stochastic processes in the traditional sense, where interventions are not allowed, assume the existence of only one probing instrument, namely that of a measurement comprising only rank-1 projections onto one of a complete set of orthogonal (classical) states at each timestep: $\mathcal{J}^{\text{cl}} := \{\Pi_{\mathbf{i}}^{(x)} \otimes \Pi_{\mathbf{o}}^{(x)}\}$, where $\Pi_{\mathbf{i}}^{(x)} = \Pi_{\mathbf{o}}^{(x)} := |x\rangle\langle x|$ satisfy $\text{tr}[\Pi^{(x)}\Pi^{(y)}] = \delta_{xy} \forall x, y$. Clearly, as this is the only interrogating instrument allowed here, $\mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F^{\text{cl}}, \mathcal{J}_M^{\text{cl}}, \mathcal{J}_H^{\text{cl}}) = \mathbb{P}_{FMH}(x_F, x_M, x_H)$ by definition. Quantum Markov order therefore automatically reduces to the classical statement, since the latter stipulates conditional independence of statistics measured with the instrument above.

This can be readily seen by considering that the complete description of any stochastic process arising from classical physics, *i. e.*, its joint probability distribution, can be encoded in the diagonal of a process tensor written with respect to the local product basis that the classical measurements act in. Thus, the process tensor of a classical stochastic process has the structure $\Upsilon_{FMH}^{\text{cl}} = \sum_y \mathbb{P}_{FMH}(y_F, y_M, y_H) \Pi_{F^{\mathbf{i}}}^{(y_F)} \otimes \Pi_{M^{\mathbf{i}}}^{(y_M)} \otimes \Pi_{H^{\mathbf{i}}}^{(y_H)}$. Importantly, these block projectors are tensor products of local projectors. Calculating the probabilities for the classical measurement above gives the correct statistics.

4.2.1 *Classical Stochastic Processes with Interventions*

As noted by van Kampen, “a physical process... may or may not be Markovian, depending on the variables used to describe it” [25]; the same is true for the Markov order. The existence of perceived memory effects implicitly depends on our experimental abilities, both in quantum mechanics—where it is generally acknowledged—as well as in classical physics—where it is often forgotten. Indeed, the standard framework for studying classical stochastic processes assumes the ability to measure observations of the system sharply and it breaks down when one allows for fuzzy measurements or interventions more generally [14, 16, 17].

For instance, consider a scenario in which an experimenter is able to sharply measure realisations of a process, deducing the ‘true’ statistics of its description that satisfy the Markov order- ℓ condition, *i. e.*, for all timesteps $t_k \in \{t_{\ell+1}, \dots, t_n\}$, the following holds

$$\mathbb{P}_k(x_k | x_{k-1}, \dots, x_1) = \mathbb{P}_k(x_k | x_{k-1} \dots, x_{k-\ell}). \quad (4.7)$$

Suppose now that, instead of measuring the observations x , the experimenter is limited in resolution to only measuring some values y that coarse-grain over subsets of the possible x values. The conditional statistics of the observed outcomes y can be explicitly written in terms of the fine-grained variable x as

$$\begin{aligned} \mathbb{P}_k(y_k | y_{k-1}, \dots, y_1) &= \frac{\mathbb{P}_{k:1}(y_k, \dots, y_1)}{\mathbb{P}_{k-1:1}(y_{k-1}, \dots, y_1)} \\ &= \frac{\sum \mathbb{P}_{k:1}(x_k, \dots, x_1)}{\sum \mathbb{P}_{k-1:1}(x_{k-1}, \dots, x_1)} \neq \mathbb{P}_k(y_k | y_{k-1} \dots, y_{k-\ell}), \end{aligned} \quad (4.8)$$

where the summation runs over the x values that are lumped together for each y .

Even if the process displays finite Markov order with respect to the sharply observed events, it does not necessarily do so for their coarse-grained variants. The fact that coarse-graining can increase the memory length observed by an experimenter arises from the well-known property that the space of Markovian processes is not convex [4]. Interestingly, we can also have the opposite occur. A process can exhibit finite Markov order with respect to a fuzzy measurement sequence, but, given access to the system at a finer resolution, the experimenter would attribute a longer memory length to the process. Such a process has finite-length memory on average. To highlight this concretely, explicit examples of both situations are provided in Appendix C.3.

Indeed, when one allows for noisy interventions in the classical case, the product structure of Eq. (2.19) is no longer satisfied for each observation, even when the underlying process is Markovian, in the sense that additional demands must be satisfied to render the future independent from the past with noisy measurements [151].² This point raises significant concerns for the practical reconstruction of complex classical dynamics, since an experimenter cannot always be certain that their measurements are distinguishing outcomes to a sufficient level of granularity [154, 155]. Various ways of dealing with such measurement noise in realistic experimental scenarios have been proposed and analysed throughout the literature [146–152].

Formally, a measurement that can pick out a particular physical state with certainty is called *sharp*; one that is not so is called *fuzzy*. In what follows, we slightly

² Classical processes for which the property of Markovianity remains invariant with respect to aggregations of events have been studied formally under the guise of *lumpable* Markov chains and the conditions for satisfaction of lumpability clearly laid out in terms of the stochastic maps of the process [153].

generalise the definitions provided in Ref. [156] in order to account for the input and output Hilbert spaces associated to each timestep. In the classical setting, a sharp measurement corresponds to a set of rank-1 projectors that are pairwise orthogonal: $\mathcal{J}^{\text{Sharp}} = \{|y\rangle\langle y|_{\text{i}} \otimes |x\rangle\langle x|_{\text{o}}\}$ such that $\langle yx|zw\rangle_{\text{i o}} = \delta_{yz}\delta_{xw} \ \forall \ x, y, w, z$ and $\sum_{yx} |y\rangle\langle y|_{\text{i}} \otimes |x\rangle\langle x|_{\text{o}}$ is a stochastic map. Intuitively, this means that if the state is measured to correspond to the value y and the state of the system that is actually sent forward into the process corresponds to x , both of these are distinguishable with certainty. Fuzzy classical measurements $\mathcal{J}^{\text{Fuzzy}}$ correspond to instruments made up of higher-rank projectors, as is the case in the example above, where the aggregation of some of the granular x values lead to higher-rank projectors describing the measured y values, with non-zero overlap between the state measured and that sent forward. Similarly, sharp quantum measurements can be defined in terms of a set of pairwise orthogonal rank-1 projectors such that $\sum_{yx} |y\rangle\langle y|_{\text{i}} \otimes |x\rangle\langle x|_{\text{o}}$ is CPTP. However, fuzziness arises in quantum theory in two possible ways: firstly, through subjective ignorance that is made manifest in the same way as classical fuzziness, *i. e.*, through higher-rank projectors with non-zero overlap; and secondly through the non-commutativity of measurement operators, which means that even sets of rank-1 POVM elements can give rise to fuzziness, since they are not necessarily projective.

More generally than the case where some outcomes are coarse-grained over, memory effects must be understood with respect to probing instruments whenever experimental interventions that directly influence the state of the system are allowed. Such invasive operations lie at the core of the theory of classical causal modelling [16] (which contains standard classical stochastic processes as a special case [14]). Here, at each timestep an experimenter is permitted to implement transformations that map any probability distributions in the state space to any other—just like in the quantum case, the experimenter could perform non-TP maps. Due to the possibility of different choices of instruments, here too the standard Markov condition must be generalised to the causal Markov condition [16]. This, in turn, is a special case of quantum Markovianity [17].

In any operational scenario where active interventions are considered, unambiguous definitions of memory effects must inherently be made with respect to the instruments used to probe the process at hand. To this end, even in classical physics, we should say that if a classical process is considered to have Markov order- ℓ , it is *with respect to sharp observations of the process*. The generalisation of Markov order provided in Def. 4.1 reduces to unambiguously characterise memory length in any scenario where classical processes with fuzzy measurements and/or experimental interventions are allowed. In

an operational sense, it is more intuitive to think of finite memory length of a process as presupposing the ability for an experimenter to apply a sequence of instruments that serve to block the influence of history on the future statistics. Moreover, the framework of quantum stochastic processes contains within it all generalised classical stochastic processes, including those pertaining to scenarios where arbitrary interventions are allowed such as causal modelling [14, 17]. Thus, statements made in terms of the process tensor reduce naturally to their classical counterparts in the appropriate sense.

However, even in the most general classical setting of causal modelling, this instrument-dependence of Markov order is liftable, in the sense that it can be removed by changing perspective. By incorporating the experimenter *and* their choice of intervention into the description of the process, the standard definitions of Markov order apply on a higher level [15]. In other words, the above concern can be ‘explained away’ on grounds of principle. On the other hand, in the study of quantum stochastic processes, even sharp quantum measurements *look* fuzzy when they act on a state in general; the measurement-dependence issue is fundamentally unavoidable and must be acknowledged accordingly through an instrument-specific notion of Markov order as per Def. 4.1. We now further explore some of its consequences in terms of a dilated dynamics, following an example presented in Ref. [2], in order to build some intuition.

4.3 MEMORY LENGTH OF A GENERALISED COLLISION MODEL

Within the framework of open quantum dynamics, collision models have been introduced to great pedagogical effect to provide a concrete underlying mechanism describing memoryless processes [157–161]. In such models, a system interacts successively with an environment comprising independent ancillary subsystems through successive unitary ‘collisions’ with each ancilla. Because each ancilla is only interacted with once, there is no way for the environment to act as a mechanism for memory transport by influencing future dynamics. Since no physical model need be prescribed to the framework, it provides a flexible toy model that is applicable to studying a wide range of phenomena.

This setting can be generalised to allow for non-trivial memory effects: the most common approaches include beginning with an initially correlated environment [162, 163], allowing for ancilla-ancilla interactions [164–168], allowing for repeated system-ancilla collisions [169, 170], or some type of hybrid approach [78, 79, 171, 172]. Each of these scenarios can be motivated through realistic physical origins that demand some reasonable assumptions [173]; in any case, the environment acts as a memory by storing

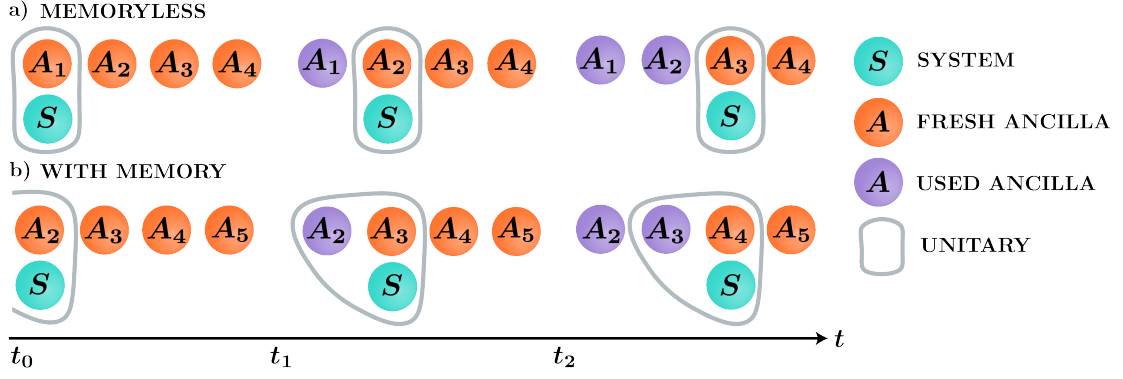


Figure 4.2: *Generalised collision model with memory.* The top panel **a)** depicts a standard memoryless collision model. The system S (green) interacts unitarily at each timestep once with each of a number of uncorrelated, fresh ancillary states A_j that constitute the environment (orange); the collision is represented by the grey boundary. Following the dynamics from t_0 to t_1 , the A_1 ancilla has been used and so stores information about the initial state of the system, indicated by the purple colour (see A_1 after t_1). However, each successive portion of evolution proceeds through an interaction with a fresh ancilla that has not yet interacted with the system. Thus, any memory of the system’s history cannot influence the future evolution, leading to Markovian dynamics. The bottom panel **b)** shows a generalised collision model, where the system is allowed to interact with multiple ancillas during each period of evolution. Here, between t_0 and t_1 , the system interacts with both A_1 (not shown) and A_2 , meaning that these ancillas can store information about the initial system state. The next portion of dynamics following t_1 involves A_2 again; thus, the future dynamics depend on the history. In this way, the ancillas serve to propagate memory effects through the process.

information about previous system states to govern future evolution (see Fig. 4.2 for illustration). Here, we focus on a special case of such dynamics with repeated system-ancilla interactions, which has application in studying phenomena with substantial time-delays between repeated interactions, *e. g.*, developing feedback-assisted control protocols [169, 170].

Consider specifically the following n -step process, depicted as a dilated quantum circuit in Fig. 4.3. A system, S , interacts with some inaccessible environment, E , which comprises $n + \ell - 1$ initially uncorrelated ancillary systems $\tau_0^E := \bigotimes_{x=1}^{n+\ell-1} \tau^{A_x}$. For simplicity, here we assume no initial system-environment correlations, and so we have a timestep t_0 at which an arbitrary system state can be prepared. The overall joint system-environment dynamics between timesteps t_{j-1} and t_j is represented by the map: $\rho_j^{SE} = \tilde{\mathcal{U}}_{j:j-1}(\rho_{j-1}^{SE})$. In this particular example, the joint evolution is broken up into an

ordered sequence of pairwise collisions between the system and ancillary states of the environment as follows

$$\tilde{\mathcal{U}}_{j:j-1} := \mathcal{U}_{j:j-1}^{SA_j} \dots \mathcal{U}_{j:j-1}^{SA_{j+\ell-1}}, \quad (4.9)$$

where the superscripts label the subsystems involved in the interaction. Following the dynamics between timesteps $t_{j-\ell}$ and t_j , the specific ancilla A_j will have interacted with the system ℓ times; it is then discarded and never again involved in the system's evolution. Note that in this model, we have not allowed for any initial system-environment correlations or ancilla-ancilla collisions; the type of evolution proposed here describes a time-translationally invariant microscopic model for processes with memory, which propagates through the ℓ ancillas that feed-forward to act like a linear memory tape. By design, we can see how memory effects arise: each ancillary system A_x can store information about the system, acquired during its first interaction through $\mathcal{U}_{x-\ell+1:x-\ell}^{SA_x}$, and use it to influence the future dynamics up until its final interaction with the system mediated through $\mathcal{U}_{x:x-1}^{SA_x}$.

Suppose then that an experimenter wishes to characterise the memory length of such a process. To do so, they must measure realisations of the state of the system at each timestep, immediately facing the problem that any such measurement both conditions the state of the environment and directly affects the state of the system which leads to different future dynamics dependent on both the measurement outcomes observed and the way in which they were measured. As we have discussed, the appropriate question relevant to understanding the memory length of the process is: *how can the experimenter block the effect of the history on the future dynamics over a finite number of timesteps?*

The representation of the process in Fig. 4.3 is particularly illuminating. We can see the possible ways in which information originating from the history, *i. e.*, about the initial system state, can perpetuate forward in time along connected paths (traced in red). For the particular collision model described above, an obvious history-blocking strategy involves discarding the system state emitted by the process and re-preparing one of a known set of states to feed into the process over a sequence of ℓ timesteps. It is clear that upon applying such a sequence of *trash-and-prepare* instruments, any possible path connecting the history to the future across ℓ timesteps is broken, thereby guaranteeing that the future evolution of the system is independent of anything that happened to it prior to the trash-and-prepare sequence.

In Appendix C.4, we prove that this trash-and-prepare protocol indeed blocks any possible influence that the history can have on the future evolution. Specifically, we show that for this particular process, at arbitrary time t_k , all future states of the sys-

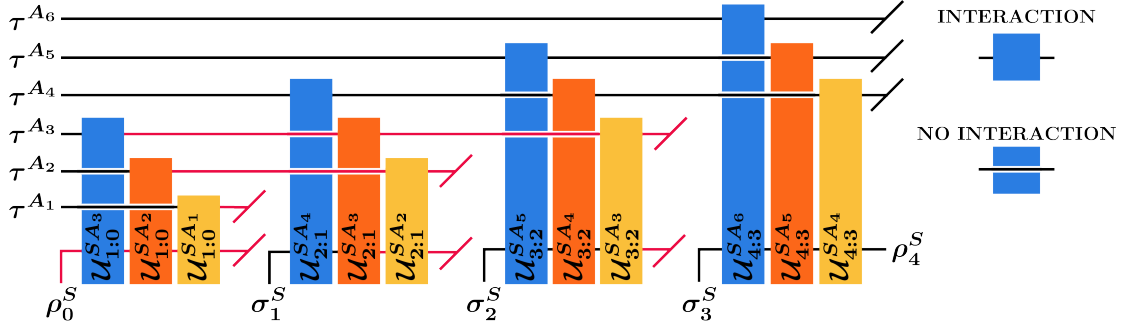


Figure 4.3: *Finite memory with respect to trash-and-prepare protocol.* The system-environment dynamics for the generalised collision model described, interspersed with the trash-and-prepare protocol applied to the system. Any possible influence stemming from the history persists to impact the future for at most $\ell = 3$ timesteps before being trashed. For instance, the red paths depicted signify the degrees of freedom that can be affected by the initial preparation, whereas the black ones cannot be. The final state is a function of only the most recent ℓ preparations, $\{\sigma_1^S, \sigma_2^S, \sigma_3^S\}$, and entirely independent of the initial system state, ρ_0^S . Any other instrument sequence on the system, *e.g.*, a measure-and-prepare rather than a trash-and-prepare instrument, would ‘open up’ a pathway for the initial state ρ_0^S to influence the future state ρ_4^S .

tem after application of any length- ℓ sequence of trash-and-prepare instruments can be uniquely described as a function of only the ℓ most recently prepared states, for any prior history. That is, the process has Markov order ℓ with respect to the entire family of trash-and-prepare sequences, implying that any possible statistics an experimenter might observe in the history and future are conditionally independent given this particular experimental control. Explicitly expressing the length- ℓ trash-and-prepare sequence in terms of operations on the system as

$$\mathcal{D}_{k-1:k-\ell}(\rho_{k-1}^S, \dots, \rho_{k-\ell}^S) := \sigma_{k-1}^S \text{tr} [\rho_{k-1}^S] \dots \sigma_{k-\ell}^S \text{tr} [\rho_{k-\ell}^S]. \quad (4.10)$$

In a slight abuse of notation, we can write

$$I(\{t_n, \dots, t_k\} : \{t_{k-\ell-1}, \dots, t_1\})_{\mathcal{D}_{k-1:k-\ell}} = 0. \quad (4.11)$$

By this we mean that the mutual information between any possible statistics recorded on the future and history timesteps, which quantifies any possible correlation between them, vanishes for all length- ℓ trash-and-prepare sequences $\mathcal{D}_{k-1:k-\ell}$.

Moreover, the experimenter could discard the states emitted by the process and feed in probabilistically-prepared states, *i.e.*, choose σ_j^S at random from the ensemble $\{\sigma_j^{(x_j)}\}$ with corresponding probabilities $\{p^{(x_j)}\}$ and the effect of history would still be blocked deterministically overall. In other words, the process also has finite

memory with respect to the trash-and-probabilistically-prepare family of instruments $\mathcal{J}_{k-1:k-\ell} = \{\mathcal{D}_{k-1:k-\ell}^{(x_{k-1}:x_{k-\ell})}\}$ with

$$\begin{aligned} \mathcal{D}_{k-1:k-\ell}^{(x_{k-1}:x_{k-\ell})}(\rho_{k-1}^S, \dots, \rho_{k-\ell}^S) &:= \\ p^{(x_{k-1})} \dots p^{(x_{k-\ell})} \sigma_{k-1}^{(x_{k-1})} \text{tr}[\rho_{k-1}^S] \dots \sigma_{k-\ell}^{(x_{k-\ell})} \text{tr}[\rho_{k-\ell}^S]. \end{aligned} \quad (4.12)$$

More generally still, the experimenter could even choose to feed in subsystems of an ℓ -partite entangled state at each timestep sampled from some ensemble $\{\sigma_{k-1:k-\ell}^{(x_{k-1}:k-\ell})}\}$ with probabilities $\{p^{(x_{k-1}:k-\ell})}\}$, overall implementing an multi-time (quantum) correlated instrument sequence of the form $\mathcal{J}_{k-1:k-\ell} = \{\mathcal{E}_{k-1:k-\ell}^{(x_{k-1}:k-\ell})}\}$ with

$$\mathcal{E}_{k-1:k-\ell}^{(x_{k-1}:k-\ell)} := p^{(x_{k-1}:k-\ell)} \sigma_{k-1:k-\ell}^{(x_{k-1}:k-\ell)} \text{tr}[\rho_{k-1}^S] \dots \text{tr}[\rho_{k-\ell}^S]. \quad (4.13)$$

In either case, for any realisation of the instrument sequences defined as the collection of CP operations in Eqs. (4.12) and (4.13), the future dynamics is conditionally independent of the history and we have satisfaction of Eq. (4.5) for all families of instruments in question

$$I(\{t_n, \dots, t_k\} : \{t_{k-\ell-1}, \dots, t_1\})_{x_{k-1:k-\ell}} = 0. \quad (4.14)$$

It is clear that the generalised collision model considered implies finite Markov order with respect to any of the aforementioned length- ℓ generalised trash-and-prepare sequences, but what can we say in the converse direction? That is, *does every process that is of finite Markov order with respect to trash-and-prepare instruments have a dilation as the one depicted in Fig. 4.3*? It turns out that having finite-length memory with respect to the trash-and-prepare protocol is a necessary but insufficient condition to deduce this generalised collision model dilation.

As a counterexample, consider two timesteps of dynamics in which two ancillary states of the environment are initially entangled, represented by the density operator $\tau^{A_1 A_2}$, and in product with the initial system state ρ_0^S . The system first interacts with A_1 via $\mathcal{U}_{1:0}^{S A_1}$, before A_1 is discarded, and then with A_2 via $\mathcal{U}_{2:1}^{S A_2}$, with a trash-and-prepare instrument $\sigma_1^S \text{tr}_S$ applied to the system in between. It is clear that the initial state ρ_0^S can have no influence on the future evolution, since the final system state can be written uniquely as a map acting only on the preparation fed into the process: $\rho_2^S = \text{tr}_{A_2} [\mathcal{U}_{2:1}^{S A_2} \sigma_1^S \otimes \tilde{\tau}^{A_2}]$, where $\tilde{\tau}^{A_2} := \text{tr}_{S A_1} [\mathcal{U}_{1:0}^{S A_1} \rho_0^S \otimes \tau^{A_1 A_2}] = \text{tr}_{A_1} [\tau^{A_1 A_2}]$ represents the reduced state of A_2 which, importantly, shows no memory of ρ_0^S . Therefore, the dynamics has finite Markov order $\ell = 1$ with respect to the trash-and-prepare protocol, but evidently does not have a dilation of the form depicted in Fig. 4.3; namely, because

the ancillas begin in an entangled state. In general, even if A_1 and A_2 interact with some unitary $\mathcal{V}_{1:0}^{A_2 A_1}$ *after* $\mathcal{U}_{1:0}^{S A_1}$, there is no dilation with initially correlated ancillas that can capture temporal correlations that might arise if some portion of the later dynamics is conditioned on the state of A_2 *before* that interaction.

To summarise, in this section we have introduced a specific type of generalised collision model which, by construction, perpetuates information about the history via a particularly simple mechanism. This allows us to study explicitly how memory effects arise from the perspective of the underlying dynamics and build an intuitive understanding of the necessity for instrument-specific quantum Markov order. The salient points to note are as follows.

i) The trash-and-prepare protocol does not block every type of memory. For arbitrary system-environment dynamics, following a length- ℓ trash-and-prepare sequence, ρ_k^S (and the future process more generally) will, generically, depend on both the known preparations $\{\sigma_{k-1}^S, \dots, \sigma_{k-\ell}^S\}$ *and* the previous historic states $\{\rho_{k-\ell-1}^S, \dots, \rho_0^S\}$. Thus, if an experimenter were to measure statistics on the future and history, they would be correlated, leading to a breakdown of Eq. (4.11) and, hence, an appreciable memory effect. In the coming chapter, we provide various examples of processes that exhibit finite Markov order with respect to other sequences of instruments, but not this one.

ii) Even for the special case of dynamics described above, application of a different sequence of instruments than the trash-and-prepare protocol would not lead to future dynamics that are independent of the history. For example, suppose that the experimenter were to perform a measurement at an intermediary timestep during a length- ℓ trash-and-prepare protocol. Here, the measurement would condition the state of the environment on its outcome, and hence the influence of the history could permeate through the memory block, leading to dependence of the final output on previous dynamics. Lastly, in Appendix C.5, we further explore some of the various other types of memory that can be built into collision models.

From the considerations outlined above, it is clear that knowing the history-blocking sequence for a given process gives us information about the process at hand, but not necessarily all of it. Although we have made no assumptions on the action of the unitaries, the dynamics examined here is a special case of generic quantum evolution and the trash-and-prepare protocol is just one of many possible sequences of instruments an experimenter might apply.

4.4 CHAPTER SUMMARY

In this chapter we have used the process tensor framework to formally provide an extension of the notion of Markov order to the quantum realm that reduces to the classical condition in the appropriate case. The intuition behind quantum Markov order remains unchanged—as in the classical case, the question boils down to: *can the future statistical evolution of the system be deduced completely, in principle, from the outcomes of the most recent ℓ instruments applied?* When any future evolution of the system is independent of any previous history following the application of some instrument sequence, the process exhibits conditional independence between the future and history. Importantly, whilst the instrument on the memory block must be specified to meaningfully define quantum Markov order, the historic and future instruments remain arbitrary: for each realisation of the instrument sequence in question applied to the memory, any possible statistics deducible on the history and future timesteps are guaranteed to be conditionally independent.

In quantum theory, we have no choice but to allow for such an active description of processes, dictating the necessity for an instrument-specific definition of memory length. We saw that demanding conditional independence between the history and the future for all possible instruments on the memory is too strong a restriction: no non-Markovian quantum process can display finite Markov order with respect to all possible interventions. Put differently, quantum stochastic processes exhibit different memory effects when probed with different instruments. Interestingly, such instrument-specific memory effects have been observed previously (see, *e. g.*, Ref. [174]); our characterisation formally captures and explains such behaviour.

This is also the case for classical stochastic processes where an experimenter can actively intervene with the system. Here, however, the issue is liftable in the sense that, in theory, we can always assume the ability to measure sharply. This is not the case in quantum mechanics: even sharp quantum measurements appear noisy as they do not generally reveal the full state of the system, and thus we have no choice but to account for the probing instruments employed. As luck would have it, by treading the path we are forced to take to understand memory effects in quantum processes, we develop a deeper comprehension of their classical counterparts.

To build intuition regarding quantum Markov order, we concluded the chapter by studying the memory length of a generalised collision model with repeated system-ancilla interactions and showed how the dynamics displays finite Markov order with respect to a

seemingly natural information-trashing instrument sequence, which serves to average out the effects that system-level operations have on the environment. The deeper exploration of memory effects in similar models in Appendix C.5 further motivates the necessity of instrument-specific Markov order for quantum processes and a better understanding of the microscopic mechanism for memory propagation.

The example analysed here provides but a special case of a quantum process with finite-length memory. There are a rich arsenal of instruments that an experimenter could, in principle, choose to apply in an attempt to block the effect of history on the future; indeed, quantum theory permits a rich landscape of memory effects, with many properties that distinguish it from the classical setting. Extending this line of investigation, we are now interested in what the satisfaction of Eq. (4.4) for a particular instrument sequence implies for the structure of the underlying process tensor. In the following chapter, we will explore the structure of quantum processes with finite Markov order with respect to certain choices of instruments, shedding light on distinguishing features and the relation between memory length, vanishing quantum CMI and the recoverability of the process. In particular, we ask which kinds of processes can have finite-length memory, and what can be inferred about the underlying process through knowledge of the history-blocking sequence.

PROCESSES WITH FINITE MEMORY LENGTH

IN THIS CHAPTER WE EXPLORE THE MATHEMATICAL STRUCTURE of processes with finite quantum Markov order. As was the case for the generalised collision model explored in Section 4.3, knowledge of the history-blocking sequence does not suffice to pin down the process at hand: in general, there can exist potential memory effects that are ‘hidden’ to the instruments in question but could be uncovered with a different probing scheme. Presently, we will examine the degree to which the structure of a process is constrained by having the property of finite Markov order.

We first detail the most general description of a process satisfying Eq. (4.4) for a particular instrument sequence. Although certain structural constraints are imposed, knowing that a particular instrument blocks the history does not necessarily tell us much about the effect of other instruments. For each key finding that arises, we present an associated example processes to build intuition regarding memory length in quantum processes. These examples are constructed in such a way as to highlight some key peculiarities of quantum Markov order, and their essence applies to processes more broadly.

With this structural understanding at hand, we will go on to analyse the connection between processes with finite Markov order and those with vanishing quantum CMI between the history and the future given the memory. From the structure of processes with vanishing quantum CMI, it follows that there exists an instrument sequence made up of orthogonal projectors that serve to render the history and future conditionally independent. Perhaps surprisingly, the converse does not hold, even in the case where the history-blocking sequence comprises only sharp, orthogonal projectors. This situation cannot happen for classical stochastic processes where, as we have discussed, finite Markov order and vanishing CMI are equivalent. We go on to lay out further restrictions imposed on the history-blocking instruments that ensures the quantum CMI vanishes.

The connection between the constraints placed on processes with finite Markov order and our desire to describe and reconstruct those with *approximately* finite memory

efficiently naturally brings us to the main concern of Chapter 6: *how can we meaningfully quantify the strength of memory and what are the subsequent implications for compression and recoverability?* Most of the results of this chapter stem from Ref. [2].

5.1 STRUCTURE OF QUANTUM PROCESSES WITH FINITE MARKOV ORDER

Our structural analysis is based on the fact that the process tensor is multi-linear in its arguments. Any ℓ -step operation sequence it acts upon can be considered as an element of a vector space $\mathbf{W}' := \text{BL}(\mathcal{H}_{k-1} \otimes \dots \otimes \mathcal{H}_{k-\ell})$ of dimension $d_{W'} := \dim(\mathbf{W}') = d^{4\ell}$, where d is the dimension of the quantum system of interest¹. As already mentioned, the only constraint on a set of operations that constitute a valid instrument sequence is that they sum to a proper comb, *i. e.*, they are positive operators whose sum yields an operator with the same causal ordering as the process tensor that acts on them, which is enough to guarantee their physicality [30]. In general, the CP elements constituting an instrument sequence need not span the entire space \mathbf{W}' , even though they are linearly independent. An instrument sequence that does span \mathbf{W} is called *informationally-complete* (IC), and any such instrument must contain a minimum number of $d_{W'}$ linearly independent elements. On the other hand, an instrument sequence that does not entirely span \mathbf{W}' is referred to as *informationally-incomplete*.

Informational-completeness and history-blocking are two distinct properties of an instrument sequence. In particular, an informationally-incomplete instrument sequence can block the history, *e. g.*, the trash-and-prepare sequence in Section 4.3. Informational-completeness pertains to whether or not an experimenter can completely characterise the process on the corresponding timesteps through knowledge of its action on each element. In the same way that an IC set of measurements must be performed to completely determine a quantum state, a process tensor can be uniquely tomographically reconstructed through knowledge of its action on an IC set of operations [50]. This property is of importance in this section, which identifies structure in the process tensor given that an experimenter knows that a certain instrument sequence blocks the history.

We focus first on the most general case, where one has satisfaction of Eq. (4.4) for an arbitrary instrument sequence. Suppose we have an informationally-*incomplete*

¹ Here, for simplicity, we consider a system of fixed dimension across all timesteps; the extension to the more general case where the dimension of the system varies at each timestep is straightforward. Note further that, in the case where the number of Hilbert spaces in the memory block is odd, $d_{W'} = d^{4\ell-2}$. Although we are not explicit regarding either of these points, their consideration only impacts the dimensionality of the underlying vector space and has no relevance to the results to be presented.

history-blocking sequence $\mathcal{J}_M = \{\mathcal{O}_M^{(x)}\}_{x=1}^c$ where $c < d_{W'}$. We can *complete* such an instrument sequence to span the entire space \mathbf{W}' by appending an additional collection of linearly independent operators, *i. e.*, construct the **IC** set $\mathcal{A}'_M = \mathcal{J}_M \cup \underline{\mathcal{A}}_M := \{\{\mathcal{O}_M^{(x)}\}_{x=1}^c, \{\underline{\mathcal{O}}_M^{(y)}\}_{y=c+1}^{d_{W'}}\} = \{\mathcal{O}'_M{}^{(z)}\}_{z=1}^{d_{W'}}$, where the underline signifies the objects that *are not* part of the original history-blocking instrument (we adhere to this notation throughout this chapter to be explicit). Note that the appended operators $\underline{\mathcal{A}}_M$ are not necessarily **CP**, nor do they necessarily form an instrument sequence; thus nor is the case for the overall construction \mathcal{A}'_M . All that is required is that $\underline{\mathcal{A}}_M$ is chosen as a linearly independent set spanning \mathbf{W}^\perp , so that \mathcal{A}'_M forms a basis for the entire space \mathbf{W}' on which the process tensor is defined. Since the entire collection \mathcal{A}'_M forms a linearly independent set (by construction), there exists an associated dual set of objects $\{\Delta_M'^{(w)}\}$ such that $\text{tr}[\mathcal{O}'_M{}^{(z)} \Delta_M'^{(w)\dagger}] = \delta_{zw} \forall z, w$ (see Appendix B.2). In terms of this (generally non-orthonormal) basis, we can (completely) represent any process tensor as

$$\Upsilon_{FMH} = \sum_z^{d_{W'}} \tilde{\Upsilon}_{FH}^{(z)} \otimes \Delta_M'^{(z)*}. \quad (5.1)$$

Since the instrument sequence \mathcal{J}_M acts to render the history and future independent for each outcome by hypothesis, we can further decompose the process tensor. We first partition the total dual set $\{\Delta_M'^{(z)}\}$ into the elements dual to those operations within the history-blocking sequence, $\{\Delta_M'^{(x)}\}_{x=1}^c$, and the rest, $\{\Delta_M'^{(y)}\}_{y=c+1}^{d_{W'}}$, such that $\text{tr}[\mathcal{O}_M^{(a)} \Delta_M'^{(b)\dagger}] = \text{tr}[\underline{\mathcal{O}}_M^{(a)} \underline{\Delta}_M'^{(b)\dagger}] = \delta_{ab} \forall a, b$ and $\text{tr}[\mathcal{O}_M^{(a)} \underline{\Delta}_M'^{(b)\dagger}] = \text{tr}[\underline{\mathcal{O}}_M^{(a)} \Delta_M'^{(b)\dagger}] = 0$. Now, the first c terms in the sum in Eq. (5.1) are $\sum_x \Upsilon_F^{(x)} \otimes \Delta_M'^{(x)*} \otimes \tilde{\Upsilon}_H^{(x)}$. By direct insertion, it is clear that this portion of the process tensor indeed satisfies Eq. (4.4). The remaining terms, which are inaccessible to the history-blocking instrument sequence \mathcal{J}_M , can be written as: $\sum_y \tilde{\Upsilon}_{FH}^{(y)} \otimes \underline{\Delta}_M'^{(y)*}$. These terms encapsulate future-history correlations that an experimenter might observe upon application of an alternative instrument. This leads to the following theorem, which outlines the most general structure a process with finite quantum Markov order must have.

Theorem 5.1. *Processes with finite quantum Markov order with respect to the instrument sequence $\mathcal{J}_M = \{\mathcal{O}_M^{(x)}\}$ are of the form:*

$$\Upsilon_{FMH} = \sum_{x=1}^c \Upsilon_F^{(x)} \otimes \Delta_M'^{(x)*} \otimes \tilde{\Upsilon}_H^{(x)} + \sum_{y=c+1}^{d_{W'}} \tilde{\Upsilon}_{FH}^{(y)} \otimes \underline{\Delta}_M'^{(y)*}, \quad (5.2)$$

where $c = |\mathcal{J}_M|$ is the number of constituent operations of the history-blocking instrument sequence, $\{\Delta_M'^{(x)}\}$ form the dual set to $\{\mathcal{O}_M^{(x)}\}$, satisfying $\text{tr}[\mathcal{O}_M^{(x)} \Delta_M'^{(y)\dagger}] = \delta_{xy} \forall x, y$, and $\{\underline{\Delta}_M'^{(y)}\}$ satisfy $\text{tr}[\mathcal{O}_M^{(x)} \underline{\Delta}_M'^{(y)\dagger}] = 0 \forall x, y$.

Importantly, each term in the first summation has F and H in tensor product, ensuring Eq. (4.4) is satisfied with certainty for each realisation of the instrument sequence in question. Such a decomposition must hold true for every timestep t_k at which a length- ℓ memory block ends (although the terms in it can change for different blocks). A quantum process with infinite Markov order with respect to every instrument cannot be written as per Eq. (5.2) with non-trivial terms in the first summation. The structure outlined makes it clear that, for an informationally-incomplete history-blocking sequence \mathcal{J}_M , the experimenter can only make meaningful statements about the memory length of the process with respect to the choice of instrument; the $\Upsilon_{FH}^{(y)}$ in the second term represents the portion of the process that can only be revealed through other probing schemes.

The generalised collision model explored in Section 4.3 is an example of such a process, since the trash-and-prepare protocol that blocks the effect of history constitutes an informationally-incomplete instrument sequence. This instrument sequence is, by its very nature, incoherent: an experimenter simply discards whatever states are output by the process and feeds in some of their own choosing. In contrast to this, one might expect that applying sequences of coherent, *i. e.*, unitary, operations to a process would always perpetuate memory effects from the history to the future by way of transmission through the level of the system alone. We now provide an explicit counterexample to this claim, *i. e.*, a process whose history is only blocked upon application of a sequence of coherent operations.

5.1.1 Unitary History-Blocking Instrument Sequences

Example 5.1 (Finite Markov order for a sequence of unitaries). Consider the process depicted in Fig. 5.1. It is constructed such that there is exactly one length- ℓ sequence of unitary operations that guarantees the history is blocked, such that the Markov order of the process is equal to ℓ . Between each timestep t_{j-1} and t_j , the process prepares an ancillary subsystem, A , in a maximally entangled state with S , $\psi^{AS} = \frac{1}{d} \sum_{xy} |xx\rangle\langle yy|$, which are together in tensor product with the rest of the environment E .

The joint EAS state undergoes dynamics according to some unitary map, \mathcal{U}_j , before an operation can be applied to the system S by the experimenter. Following this operation, the process applies the inverse \mathcal{V}_j^\dagger of some other unitary map \mathcal{V}_j on the system alone, where $\mathcal{V}_j^\dagger(\bullet) := V_j^\dagger \bullet V_j$. The joint EAS state then evolves according to the inverse unitary map \mathcal{U}_j^\dagger . Lastly, AS is subject to the following *memory cutting protocol*: a Bell basis measurement is implemented within the process, with another ancillary subsystem,

with the first sum containing a single term and the remainder of the process description encapsulated in the second term

$$\Upsilon_{FMH} = \frac{1}{d^\ell} \Upsilon'_F \otimes V'_{k-1} \otimes \dots \otimes V'_{k-\ell} \otimes \Upsilon_H + \sum_y \tilde{\Upsilon}_{FH}^{(y)} \otimes \underline{\Delta}_M^{(y)}, \quad (5.3)$$

where the V'_j/d are duals to the Choi states of the unitary maps \mathcal{V}_j^\dagger , and the conditional process tensor Υ'_F is the fresh future process initiated by successful implementation of the cutting protocol, *i. e.*, the process that ensues in line with the freshly prepared environment state.

The process tensor in Eq. (5.3) is evidently an expression of Theorem 5.1; however, some remarks are in order. Firstly, note that even in the special case $\ell = 1$, the process is non-Markovian, since it does not have the product structure outlined in Eq. (3.36) (and the coherent unitary operation at timestep t_{k-1} required to block the effect of history on the future operates on $\text{BL}(\mathcal{H}_{k-1^o} \otimes \mathcal{H}_{k-1^i})$). Loosely speaking, the history-blocking unitary operation does not serve to ‘cut’ the system line, but rather keep it ‘glued together’ so that temporal correlations can be transmitted through the system, in contrast to the effect of a causal break. It is these transmitted correlations that are then used to cut the information flow from the history to the future. Secondly, no sequence of unitary operations can be IC; by definition, an informationally-incomplete sequence cannot be used to extract full information about a process. Although we know that any future dynamics will be independent of the history with respect to this sequence, we cannot predict what the next state will be as a function of the history-blocking sequence.

5.1.2 Informationally-Complete History-Blocking Sequences

Interestingly, in Example 5.1, the influence of the history on the future is blocked only by a sequence of coherent operations. This is somewhat counter-intuitive, as one might expect unitary transformations to perpetuate memory effects. Clearly, the general structural constraint of Theorem 5.1 is rather flexible, since knowledge of such an incomplete history-blocking instrument sequence does not determine the structure of the process at hand. In many cases of interest, an experimenter makes use of an IC set of operations to probe the dynamics, *e. g.*, when attempting to tomographically reconstruct a generic process. In this case, since an IC instrument sequence does in fact span the entire space of operations (by definition), there can be nowhere for potential memory effects correlating the history and future to hide. The memory block can then be completely

decomposed onto an **IC** set of duals, uniquely specifying the entire process for each sequence of outcomes realised on the memory block. In this case, finding the future process to be conditionally independent of the history constrains the structure of the process tensor in a stricter manner than Eq. (5.2); we immediately have the following corollary.

Corollary 5.2. *A process with finite Markov order with respect to an informationally-complete instrument sequence must have the following structure:*

$$\Upsilon_{FMH} = \sum_x \Upsilon_F^{(x)} \otimes \Delta_M^{(x)*} \otimes \tilde{\Upsilon}_H^{(x)}. \quad (5.4)$$

*Note that this structure by no means implies that the process tensor is of a tensor product form, or that the history is necessarily blocked by any instrument sequence other than the **IC** one in question.*

With this corollary at hand, it is enlightening to re-examine Theorem 4.2, which states that the only processes with finite quantum Markov order with respect to all instrument sequences are memoryless. Its proof begins by demanding Eq. (4.4) to hold for all possible instruments. As such, we can consider an **IC** instrument sequence, in which case the process tensor must be of the form given by Eq. (5.4). Then, using the fact that one can construct arbitrary operation sequences spanning the space of operations on M , we can vary $\Delta_M^{(x)}$ freely. Demanding the structure of Eq. (5.4) to remain intact for arbitrary outcomes forces a tensor product between M and F or H (or both), meaning the process tensor is restricted to a single term in Eq. (5.4), *i. e.*, it is of product form. Requiring this to hold for any timestep leads to a memoryless process (either the Markovian or super-Markovian product structure of Eqs. (3.36) and (4.6) respectively).

An operationally motivated choice for an **IC** instrument sequence consists of applying a causal break at each timestep: recall that each operation here consists of an **IC** POVM followed by an independent preparation of one of an **IC** set of states to feed forward at each timestep (see Eq. (3.32)). The following example is constructed in such a way that the process is non-Markovian; however, it exhibits finite Markov order with respect to such an **IC** instrument sequence of causal breaks.

Example 5.2 (Finite Markov order with respect to an informationally-complete instrument sequence (causal breaks)). Consider the process depicted in Fig. 5.2, where, for simplicity, we present the case $\ell = 2$ for a 3-step process, with the extension to longer length memory immediate. Initially, the following tripartite state is prepared

$$\rho_{Y2^11^1} = \sum_y p_y \rho_Y^{(y)} \otimes \Delta_{2^1}^{(y)*} \otimes \rho_{1^1}^{(y)}, \quad (5.5)$$

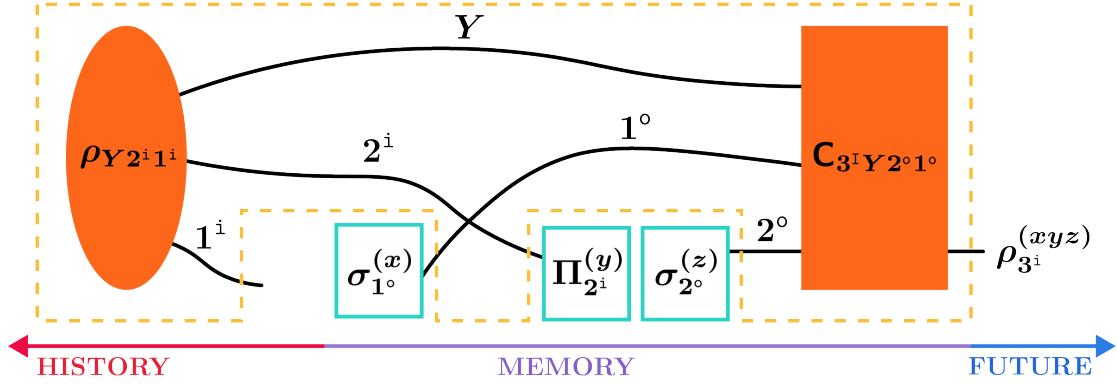


Figure 5.2: *Finite memory for an informationally-complete sequence.* Initially, a tripartite state $\rho_{Y2^i1^i}$ is constructed as per Eq. (5.5), with subsystems $1^i, 2^i$ of it fed out at consecutive timesteps as described in the text. The states fed back into the process on spaces $1^o, 2^o$ are fed forward as inputs to the CPTP map $\mathcal{C}_{3^iY2^o1^o}$ defined in Eq. (5.6). Upon applying any combination of the correct IC causal break sequence $\{\sigma_{1^o}^{(x)}, \Pi_{2^i}^{(y)}, \sigma_{2^o}^{(z)}\}$, one of d^6 final output states $\rho_{3^i}^{(xyz)}$ are output by the process in the future, each of which is conditionally independent of the historic $\rho_{1^i}^{(y)}$. If any other operations are applied, correlations can arise between the history and future.

with $\{\Delta_{2^i}^{(y)}\}$ forming the dual set to some IC POVM $\{\Pi_{2^i}^{(y)}\}$ and Y labelling an ancillary Hilbert space of the environment that is never accessible to the experimenter. The 1^i part of the initial state is fed out of the process at the first timestep, at which point the experimenter can implement any operation of their choice; similarly, the 2^i part is fed out at the second timestep. The output states of the experimenter's operations at timesteps 1^o and 2^o are mediated forward by the process, along with the Y part of $\rho_{Y2^i1^i}$, as inputs to a CPTP map $\mathcal{C} : \text{BL}(\mathcal{H}_Y \otimes \mathcal{H}_{2^o} \otimes \mathcal{H}_{1^o}) \rightarrow \text{BL}(\mathcal{H}_{3^i})$, whose Choi operator is

$$\mathcal{C}_{3^iY2^o1^o} := \sum_{xyz} \rho_{3^i}^{(xyz)} \otimes D_Y^{(y)*} \otimes D_{2^o}^{(z)*} \otimes D_{1^o}^{(x)*}, \quad (5.6)$$

where $\{D_Y^{(y)}\}$ are the dual set to $\{\rho_Y^{(y)}\}$, and $\{D_{2^o}^{(z)}\}, \{D_{1^o}^{(x)}\}$ respectively form the dual set to some IC set of preparations $\{\sigma_{2^o}^{(z)}\}, \{\sigma_{1^o}^{(x)}\}$. This map acts to take each one of the $\{\sigma_{1^o}^{(x)}, \rho_Y^{(y)}, \sigma_{2^o}^{(z)}\}$ combination of its inputs to one of d^6 unique states $\rho_{3^i}^{(xyz)}$, which are the final outputs of the process.

Stipulating the construction of $\rho_{Y2^i1^i}$ in Eq. (5.5) to be a positive semidefinite operator overall, and the map $\mathcal{C}_{3^iY2^o1^o}$ defined in Eq. (5.6) to represent a valid evolution, requires sufficient mixedness of each $\rho_{1^i}^{(y)}$ and $\rho_{3^i}^{(xyz)}$; additionally, choosing preparations $\rho_Y^{(y)}, \sigma_{2^o}^{(z)}$ and $\sigma_{1^o}^{(x)}$ such that $\sum_{xyz} D_Y^{(y)} \otimes D_{2^o}^{(z)} \otimes D_{1^o}^{(x)} = \mathbb{1}_{Y2^o1^o}$ ensures $\mathcal{C}_{3^iY2^o1^o}$ satisfies the necessary trace conditions of Eq. (3.15). Importantly, all of these conditions outlined above can be achieved simultaneously. It then follows that there exists an underlying

unitary dilation of the map $C_{3^i Y 2^o 1^o}$. The overall process tensor for the process described above is explicitly given by

$$\Upsilon_{3^i:1^i} = \sum_{xyz} p_y \rho_{3^i}^{(xyz)} \otimes D_{2^o}^{(z)*} \otimes \Delta_{2^i}^{(y)*} \otimes D_{1^o}^{(x)*} \otimes \rho_{1^i}^{(y)}. \quad (5.7)$$

Intuitively, the **IC** instrument sequence $\mathcal{J}_M = \{\sigma_{1^o}^{(x)}, \Pi_{2^i}^{(y)}, \sigma_{2^o}^{(z)}\}$ blocks any influence from the history to the future, as the measurement performed at 2^i leaves the initial state $\rho_{Y 2^i 1^i}$ in a product between Y and 1^i for each outcome, such that the final output state is then independent of any operation that could be performed at t_{1^i} . Indeed, for any realisation of the instrument sequence, the conditional future and history processes are in the product form of Eq. (4.4)

$$\text{tr}_{2^o 2^i 1^o} \left[\left(\sigma_{2^o}^{(z)} \otimes \Pi_{2^i}^{(y)} \otimes \sigma_{1^o}^{(x)} \right)^T \Upsilon_{3^i:1^i} \right] = p_y \rho_{3^i}^{(xyz)} \otimes \rho_{1^i}^{(y)}. \quad (5.8)$$

In this sense, the map $C_{3^i Y 2^o 1^o}$ has no bearing on whether the effect of history is blocked or not: an experimenter could coarse-grain over any of the preparations while applying the correct measurement, *e. g.*, feed in $p\sigma_{1^o}^{(x)} + (1-p)\sigma_{1^o}^{(x')}$, yielding a future state $p\rho_{3^i}^{(xyz)} + (1-p)\rho_{3^i}^{(x'yz)}$ that remains conditionally independent of the history $\rho_{1^i}^{(y)}$ given any measurement outcome y of $\{\Pi_{2^i}^{(y)}\}$ at t_{2^i} . Of course, simpler processes can lead to an independent history and future with respect to the outcomes of an **IC POVM** (see Example D.1 given in Appendix D.1). However, here we construct a more general process with $C_{3^i Y 2^o 1^o}$ defined as per Eq. (5.6) in order to yield d^6 *distinct* future states $\rho_{3^i}^{(xyz)}$ for each possible realisation of the causal break sequence, each one of which is conditionally independent of the history.

Just as in the generalised collision model of Section 4.3, in principle an experimenter can predict the next state of the system as a function of measurements and preparations in the causal break sequence. Furthermore, since the history-blocking sequence is **IC**, they could perform a process tomography to completely characterise the process as per Eq. (5.7). If, on the other hand, the experimenter were to apply a different instrument on the memory block, then correlations between the future and history would in general arise (but, as already mentioned, they could vary the preparations and not see any influence from the history).

The two examples provided in this section highlight significant properties of memory in quantum processes. Example 5.1 explicitly shows that there exist processes where specific sequences of unitary operations can break all possible temporal correlations between future and history, while Example 5.2 highlights that the operations of a history-blocking instrument sequence can comprise an **IC** (in general non-orthogonal) set of independent

measurements and preparations. So far, through Theorem 5.1 and Corollary 5.2, we have developed the structural constraints that a process tensor must satisfy in order to exhibit finite quantum Markov order for a given instrument sequence. However, this characterisation is difficult to check in practice, due to the non-uniqueness of possible decompositions for a process tensor. It is therefore natural to seek a function of such finite Markov order processes that vanishes iff there are no correlations between the history and future remaining once a memory block of length ℓ is specified. For classical stochastic processes (without interventions), it is straightforward to show that the CMI of the underlying joint probability distribution has the desired property. In contrast, in both of the above examples (and also in the generalised collision model of Section 4.3), the quantum CMI evaluated on the Choi operator of the process tensor between the history and future with respect to the memory is non-vanishing. For the sake of comprehensiveness, we explicitly construct the process tensor for a simple example and calculate the quantum CMI to be non-zero in Appendix D.1. This observation is insightful for a number of reasons which we address in the coming section, where we explore in detail the necessary conditions on the history-blocking instrument sequences for processes with vanishing quantum CMI, of which classical processes with finite Markov order are a special case.

5.2 QUANTUM MARKOV ORDER AND CONDITIONAL MUTUAL INFORMATION

In light of the observation above, we have the following theorem.

Theorem 5.3. *Vanishing quantum CMI guarantees the process has finite quantum Markov order; the converse is not true.*

The structure of processes with vanishing quantum CMI can be deduced from that of quantum states with vanishing quantum CMI, with the additional causality constraint imposed to ensure a valid process. The CMI of a quantum process is defined by $I(F : H|M) := S(\Upsilon_{FM}) + S(\Upsilon_{MH}) - S(\Upsilon_{FMH}) - S(\Upsilon_M)$, and it vanishes iff there exists a block orthogonal decomposition of the composite M Hilbert space as $\mathcal{H}_M = \bigoplus_m \mathcal{H}_{ML}^{(m)} \otimes \mathcal{H}_{MR}^{(m)}$, such that [137]

$$\Upsilon_{FMH}^{\text{CMI}=0} = \bigoplus_m p_m \tilde{\Upsilon}_{FML}^{(m)} \otimes \tilde{\Upsilon}_{MRH}^{(m)}. \quad (5.9)$$

Here, the decomposition of \mathcal{H}_M does not necessarily respect the temporal ordering of the underlying process; specifically, the Hilbert spaces $\{\mathcal{H}_{ML}^{(m)}\}$ do not need to describe events that occur strictly before or after those described in $\{\mathcal{H}_{MR}^{(m)}\}$.

The proof of Theorem 5.3 is given in Appendix D.2; the basic strategy is to explicitly construct a history-blocking instrument sequence for processes of the form in Eq. (5.9), and show that this structure is a special case of Eq. (5.2), meaning that vanishing quantum CMI implies finite Markov order. The history-blocking sequence we construct is, in fact, made up of the set of orthogonal projectors (which form a self-dual set) onto each of the m subspaces in the decomposition above.

An immediate reason for why the converse does not have to hold is that there is no reason why the memory blocking operations should be sharp, *i. e.*, that the projectors are rank-1 and pairwise orthogonal. In the case where the experimenter finds conditional independence with respect to a sequence of higher-rank fuzzy projectors, the future-history correlations hidden within each m subspace need not obey the constraint implied by Eq. (5.9), and hence the process can have non-vanishing quantum CMI, as shown explicitly in Example D.2 of Appendix D.3. As we have already seen in Section 4.2, similar behaviour arises in an operational interpretation of classical stochastic processes: if the experimenter cannot measure realisations of the process sharply then the statistics observed do not necessarily have vanishing classical CMI, even if the true underlying process is one of finite Markov order (see the examples in Appendix C.3).

This immediately begs the question: *do processes with finite Markov order with respect to an instrument sequence comprising only rank-1 orthogonal projectors necessarily have vanishing quantum CMI?* In addition to the proof of Theorem 5.3, we show in Appendix D.2 that—perhaps surprisingly—this is not the case. Intuitively, this is because such projectors that make up an instrument can live on an extended input-output Hilbert space at each timestep; in quantum theory, sharp measurements of composite systems can be fuzzy locally if they are made in an entangled basis. Thus, even if the instrument sequence comprises only rank-1, orthogonal projectors, the structural condition implied on the process tensor is still not strong enough to force a block-diagonal representation as necessary for the quantum CMI to vanish. Here we provide an explicit example that evidences this point.

Example 5.3 (Process with non-vanishing quantum CMI but finite Markov order for a sequence of rank-1, orthogonal projectors). Consider the process depicted in Fig. 5.3. The 4-dimensional ancilla qudit is initially in a coherent superposition $|\tau\rangle^A = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle + \delta|3\rangle$ with $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$ constituting the environment. Controlled on the state of this qudit, the process implements one of the four Pauli maps

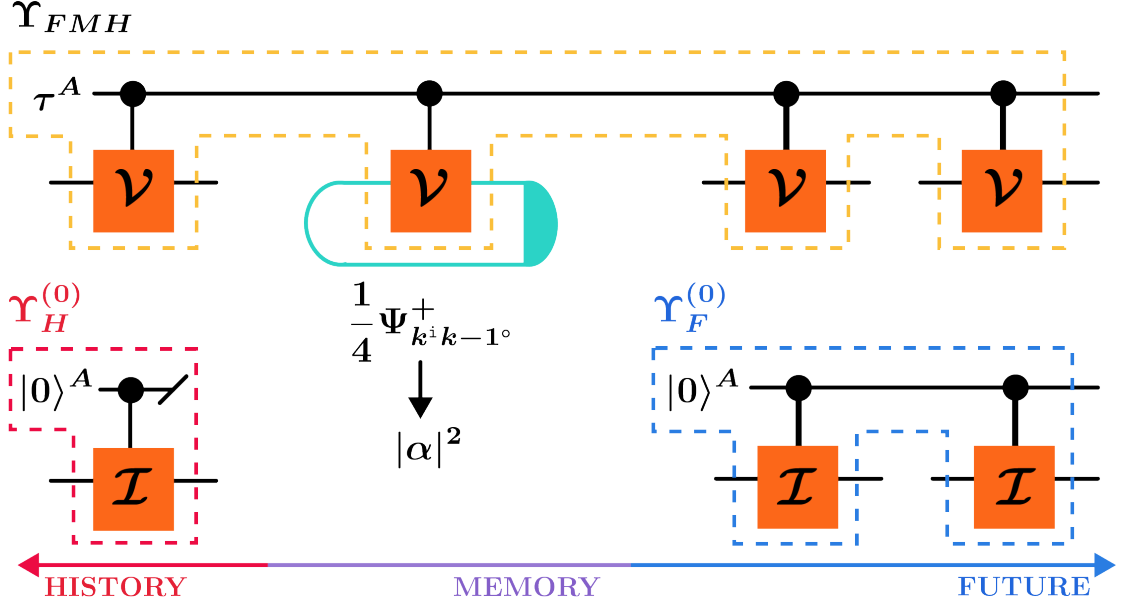


Figure 5.3: Process with non-vanishing quantum CMI. The environment is a 4-dimensional ancilla. Its initial state is a coherent superposition of the basis states $\{|0\rangle, \dots, |3\rangle\}$. The system-environment evolution is a control unitary, which implements one of the four Pauli rotations $\mathcal{V} := \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ on the system depending on the state of the ancilla (see top panel). The history-blocking instrument sequence consists of feeding in one half of a Bell pair and, at the next step, measuring the system and the other half in the Bell basis at the next timestep. For each outcome of this instrument, one can infer which of the four Pauli rotations was applied, and the history and future processes are conditionally independent. For illustrative purposes, the bottom panel depicts the conditional processes that arise from successful implementation of the operation $\frac{1}{4}\Psi^+$, which occurs with probability $|\alpha|^2$.

(including the identity map), $\mathcal{V} := \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$, on a single qubit system. The Choi operators of these maps are the projectors of the four (unnormalised) Bell pairs

$$|\Psi^\pm\rangle := |00\rangle \pm |11\rangle \quad \text{and} \quad |\Phi^\pm\rangle := |01\rangle \pm |10\rangle. \quad (5.10)$$

Suppose that the process continues for n timesteps and, at the end of the process, the ancilla is fed out with the system in order to retain the quantum features of the process. For simplicity, we also assume that there are no initial system-environment correlations, allowing us to define the process as one beginning on an output wire at t_1 . The corresponding process tensor is $\Upsilon_{n:1} = |\Upsilon\rangle\langle\Upsilon| \in \text{BL}(\mathcal{H}_{n^i}^A \otimes \mathcal{H}_{n^i}^S \otimes \mathcal{H}_{n-1^\circ}^S \otimes \dots \otimes \mathcal{H}_{1^\circ}^S)$, where

$$\begin{aligned} |\Upsilon\rangle := & \alpha|0\rangle_{n^i}^A \otimes |\Psi_{n^i n-1^\circ}^+ \dots \Psi_{2^i 1^\circ}^+\rangle + \beta|1\rangle_{n^i}^A \otimes |\Phi_{n^i n-1^\circ}^+ \dots \Phi_{2^i 1^\circ}^+\rangle \\ & + \gamma|2\rangle_{n^i}^A \otimes |\Phi_{n^i n-1^\circ}^- \dots \Phi_{2^i 1^\circ}^-\rangle + \delta|3\rangle_{n^i}^A \otimes |\Psi_{n^i n-1^\circ}^- \dots \Psi_{2^i 1^\circ}^-\rangle. \end{aligned} \quad (5.11)$$

Note that this is not a Markovian process (it is not of the product form of Eq. (3.36)), nor is it a classical probabilistic mixture of such processes; rather, the process tensor is a pure state representing a coherent superposition of implementing sequences of the four Pauli maps, corresponding to a genuinely quantum memory.

Consider the instrument sequence where, at some timestep t_{k-1} , an experimenter inputs half of one of the Bell pairs, feeds the other half forward to the next timestep t_k , and then makes a Bell basis measurement on the fed-forward ancilla and the system state output by the process (see Fig. 5.3). This instrument is made up of the Choi states $\mathcal{J}_{k^i k-1^\circ} = \{\mathcal{O}_{k^i k-1^\circ}^{(x)}\} := \frac{1}{4}\{\Psi_{k^i k-1^\circ}^+, \Phi_{k^i k-1^\circ}^+, \Phi_{k^i k-1^\circ}^-, \Psi_{k^i k-1^\circ}^-\}$. Since all cross terms in $\Upsilon_{n:1}$ are orthogonal to any of these, for each outcome observed upon their application, the experimenter observes one of the following four conditional processes

$$\begin{aligned}\tilde{\Upsilon}_{FH}^{(0)} &= |\alpha|^2 \Psi_F^+ \otimes \Psi_H^+, & \tilde{\Upsilon}_{FH}^{(1)} &= |\beta|^2 \Phi_F^+ \otimes \Phi_H^+ \\ \tilde{\Upsilon}_{FH}^{(2)} &= |\gamma|^2 \Phi_F^- \otimes \Phi_H^-, & \tilde{\Upsilon}_{FH}^{(3)} &= |\delta|^2 \Psi_F^- \otimes \Psi_H^-\end{aligned}\tag{5.12}$$

where $\Psi_F^+ := |0\rangle_{n^i}^A \otimes \Psi_{n^i n-1^\circ}^+ \otimes \dots \otimes \Psi_{k+1^i k^\circ}^+$, $\Psi_H^+ := \Psi_{k-1^i k-2^\circ}^+ \otimes \dots \otimes \Psi_{2^i 1^\circ}^+$, and the superscript label corresponds to each possible realisation *e.g.*, the label (0) corresponds to the experimenter feeding in half of the state $\Psi^+/2$ and successfully measuring it, which occurs with probability $|\alpha|^2$, and similarly for the other quantities defined. The conditional tester elements can thus each be normalised to a proper process tensor by simply dividing by the appropriate probability *e.g.*, $\Upsilon_{FH}^{(0)} = \tilde{\Upsilon}_{FH}^{(0)}/|\alpha|^2$.

Intuitively, once an outcome of the instrument described is observed, the experimenter can deduce which of the four independent control operations were applied to the system and hence the state of the ancilla, which collapses onto one of its computational basis states and does not change further throughout the process. This means that the history and future processes are known with certainty and are thus conditionally independent with respect to knowledge of the instrument outcome. In contrast, suppose that the experimenter were to perform an incoherent operation, such as feeding in the maximally mixed state before averaging over all measurement outcomes at the subsequent timestep. In this case, the conditional future-history process is now a probabilistic mixture of the four control operations being applied *i.e.*, $\Upsilon_{FH} = \sum_x \tilde{\Upsilon}_{FH}^{(x)}$, with $\{\tilde{\Upsilon}_{FH}^{(x)}\}$ defined in Eq. (5.12). Such a mixture of Markovian processes is non-Markovian due to the correlations between the future and history: indeed, in this scenario an experimenter could condition the future dynamics by performing certain operations in the history.

A simple calculation shows that the quantum CMI between the history and future given the memory for the process tensor in Eq. (5.11) does not vanish; rather, it is equal

to the Shannon entropy of the distribution $\mathbb{P}(x) = \{|\alpha|^2, |\beta|^2, |\gamma|^2, |\delta|^2\}$. Lastly, note that, had we chosen to discard the ancilla, rather than feed it out at the final timestep, the corresponding process tensor is a probabilistic mixture of sequences of the four Pauli maps applied, *i. e.*, the projector of Eq. (5.11) *without any cross terms*. In this case, the process tensor is of the form in Eq. (5.9) and the quantum CMI vanishes.

In summary, here we have an example of a process which has finite Markov order with respect to an instrument sequence comprising only rank-1, orthogonal projectors, but nonetheless has non-vanishing quantum CMI. The example presented here represents a genuinely quantum mechanical memory effect with no classical analog. The intuition behind the distinction is that in the classical setting, d orthogonal projectors are IC for a d -level system, whereas this fails to hold true in the quantum realm. An additional restriction on the “off-diagonal” terms of the process tensor must be adhered to in order to ensure the quantum CMI vanishes, which is formulated in Appendix D.2.

5.3 CHAPTER SUMMARY

In this chapter, we have outlined some of the key features of memory length in stochastic processes, many of which are peculiar to quantum mechanics. We began by tackling the general problem: *given a sequence of operations that acts to erase the effect of history on the future of a process, what can we say about its overall structure?* In Section 5.1 we detailed the generic constraint on process tensors with finite quantum Markov order, providing the most general structure deducible to an experimenter who knows the history-blocking instrument sequence in question. The first special case of this structure was exhibited in Example 5.1, where we studied a process whose history is blocked by a sequence of unitary operations. Although such unitary sequences can serve to block the effect of history, they provide minimal information to the experimenter about the process at hand and thereby represent the extremal case of informationally-incomplete history-blocking instrument sequences. On the other extreme are IC instruments which stipulate the complete description of the process. We then considered processes with finite-memory with respect to such IC instrument sequences, in particular an IC POVM followed by an independent reparation of a state from an IC set in Example 5.2.

Following this, we analysed the connection between quantum Markov order and the vanishing of the quantum CMI. As mentioned previously, until the recent introduction of the process tensor, there was no meaningful way to develop a sensible notion of Markov order in the temporal setting, since the statistics observed in time depend upon how an

experimenter probes the process and are thus inherently instrument-dependent. Despite this concern, many efforts throughout the literature are concentrated on Markov chains, defined as tripartite quantum states that are recoverable or (equivalently) have vanishing quantum CMI. On the other hand, the general theory of quantum Markov order for processes introduced here is captured by the conditional independence statement of Eq. (4.4); this instrument-dependent statement is in stark contrast with the aforementioned definitions on quantum states, which make no mention of the instrument sequence of choice. Therefore, it is not immediately clear how such characterisations concretely relate to temporal processes with finite quantum Markov order.

Nonetheless, the CJI lets us consider temporal processes in terms of their corresponding Choi operator, allowing us to concretely examine the link between the two inequivalent notions. In Section 5.2, we proved that in the quantum realm, processes with finite Markov order with respect to a sequence of instruments need not necessarily have vanishing quantum CMI. Although a similar departure can arise in the study of classical stochastic processes where fuzzy measurements are permitted, in quantum mechanics there can exist processes with finite Markov order with respect to a sequence of sharp, orthogonal projectors that has non-vanishing quantum CMI—in direct contradistinction to the classical setting. An explicit construction is provided in Example 5.3 and the additional constraints on the process tensor required to guarantee vanishing quantum CMI are detailed in Appendix D.2.

The results uncovered in this chapter raise some interesting avenues for future exploration. For instance, many realistic physical scenarios are often modelled by specific forms of interactions, *e. g.*, nearest-neighbour interaction spin chains evolving in a time-translationally invariant manner. In such a scenario, whilst a generic sequence of instruments such as the trash-and-prepare protocol will typically not act to block the historic influence, in practice it may be the case that such a sequence almost always *approximately* blocks the influence of history. A natural extension to this work would involve a deeper exploration of memory effects in specific physical models with respect to the instrument-specific quantum Markov order formalism.

A first step to understanding processes with approximately finite-length memory is to quantify memory strength, that is, the amount of temporal correlations remaining between the history and the future processes after an instrument sequence of choice is implemented on the memory block. In the coming chapter, we will introduce an instrument-specific notion of memory strength and explore its flexibility in characterising memory effects in an exactly solvable non-Markovian model.

MEMORY STRENGTH

SO FAR, WE HAVE LOOKED AT THE STRUCTURAL PROPERTIES of processes with finite Markov order. Now, we will quantify the *strength* of said memory effects, if they exist; that is, the degree to which they influence the observed statistics. This can be quantified by the temporal correlations remaining between the history and the future processes for a specified instrument on the memory block. In other words, we wish to understand the deviation from future-history independence with respect to the instrument sequence in use.

The ultimate goal concerning the quantification of memory strength is to understand the circumstances under which one can effectively describe processes in an efficient manner. Indeed, we are unlikely to find processes with strictly finite Markov order in nature, since, with respect to generic instruments, non-Markovian processes typically exhibit infinite memory length. Understanding how strong the memory effects across a given duration in time are has significant implication for the simulation of processes with *approximately* finite memory length.

For instance, numerical techniques for open dynamics often invoke finite memory approximations, where rapidly vanishing temporal correlations are truncated [24, 175–177]. This is tantamount to treating the process as having finite Markov order with respect to the identity instrument (*i. e.*, doing nothing to the process), although memory approximations involving other choices of instruments can also be made [178]. Another related result concerns quantum *states* with small quantum CMI. Recent bounds on the fidelity of recovery show that these tripartite quantum states allow for the existence of a CPTP map that approximately recovers the total state from only partial information, *i. e.*, by acting upon the conditioning subsystem alone [139]. Although this result does not directly pertain to processes, the CJI permits similar considerations in the temporal setting. However, as we have seen in the previous chapter, the quantum CMI is a poor

quantifier for the memory strength, as it does not necessarily vanish for processes with finite Markov order.

As we have laid out so far in this thesis, genuine memory effects in quantum processes must be specified with respect to the instrument sequence used to probe them. It is thus natural that a proper quantification of strength must also be instrument-specific. To this end, in this chapter, we will first develop a number of viable candidates for quantifying memory strength in quantum processes. We will then examine these quantifiers and the resulting measures for memory strength for an exactly solvable system-environment model, introduced in Ref. [179]. This model is amenable to the analysis of memory effects for a large host of physical situations by tuning parameters in the corresponding system-environment Hamiltonian, and therefore allows us to explicitly study many of the main concepts introduced throughout this thesis. We first highlight that multi-time memory effects, which remain uncaptured by two-time witnesses of non-Markovianity, are present throughout the parameter space, before exploring its memory structure with respect to natural families of instrument sequences. Here we see that over a fixed length of the memory block, the process indeed displays different memory strength for different probing instruments.

An important special case is where the instrument sequence is chosen to be the identity transformation on the system, as this pertains to the ‘natural’ memory of the process at hand, in the sense that it ties in most nicely with existing discussions of memory. For instance, standard numerical open systems techniques can be appropriately implemented when the memory effects in such cases die off rapidly [24]. We go on to study this situation, exploring the length of time over which the memory in the process naturally decays.

6.1 QUANTIFYING MEMORY STRENGTH

We begin by introducing an instrument-specific notion of memory strength. We build this up by first introducing quantifiers of the memory effects for each specific outcome of the instrument sequence in question, before motivating a number of suitable aggregations that compress these numbers into a single quantifier for the overall memory strength for the instrument.

6.1.1 Instrument-specific Memory Strength

We have already encountered a quantity that serves as a suitable starting point to the quantification of memory: the mutual information between the conditional history and future processes for each realisation of an instrument on the memory block (see Eq. (4.5)). We take these values as the foundation of the definition of memory strength.

Definition 6.1 (Outcome-specific memory strength). The memory effects of a process across $\ell = |M|$ timesteps are characterised by the mutual information between the conditional history and future processes for *each outcome* of the instrument:

$$\mathcal{S}_{\mathcal{J}_M}(x_M) := I(F : H)_{x_M} = S(\Upsilon_F^{(x_M)}) + S(\tilde{\Upsilon}_H^{(x_M)}) - S(\tilde{\Upsilon}_{FH}^{(x_M)}), \quad (6.1)$$

where $\mathcal{J}_M = \{\mathcal{O}_M^{(x_M)}\}$ and $\tilde{\Upsilon}_{FH}^{(x_M)} = \text{tr}_M [\mathcal{O}_M^{(x_M)} \Upsilon_{FMH}]$.

Note that this is not yet a complete definition of memory strength for the full instrument, but a quantifier for each specific outcome. Nonetheless, the quantity vanishes for all x_M iff the process has finite memory with respect to the instrument sequence \mathcal{J}_M .

One possible way to aggregate these quantifiers would be to take the average with respect to the probabilities of each realisation x_M ; however, this is not necessarily valid, as the corresponding probabilities in general depend on the historic probing instruments. Nonetheless, an average with respect to the *uniform distribution* serves to quantify the overall memory effect one might expect for the given instrument, when taking the uniform distribution as the unbiased prior for the occurrence probability of the respective outcomes. In this way, we define the *average* memory strength for \mathcal{J}_M as follows.

Definition 6.2 (Average instrument-specific memory strength). The average memory strength for the instrument \mathcal{J}_M is computed by taking the average of the outcome-specific mutual informations between the history and future (as per Def. 6.1) with respect to the uniform distribution:

$$\mathcal{S}_{\mathcal{J}_M}^{\text{avg}} := \frac{1}{|\mathcal{J}_M|} \sum_{x_M} \mathcal{S}_{\mathcal{J}_M}(x_M). \quad (6.2)$$

Alternatively, taking the *maximum* over all outcomes of the instrument sequence provides a quantifier for the extremal scenario that an experimenter might see in an individual run of the experiment.

Definition 6.3 (Maximum instrument-specific memory strength). The maximum memory strength for the instrument \mathcal{J}_M is the maximum of the outcome-specific mutual informations between the history and future (as per Def. 6.1) over all possible outcomes:

$$\mathcal{S}_{\mathcal{J}_M}^{\text{max}} := \max_{x_M} \mathcal{S}_{\mathcal{J}_M}(x_M). \quad (6.3)$$

Both Defs. 6.2 and 6.3 are no longer *outcome-specific*, but rather *instrument-specific* in the sense that they pertain only to quantifying the overall memory effect across a specified instrument. They both vanish iff the process has quantum Markov order ℓ with respect to \mathcal{J}_M . In contrast, taking, say, the minimum value as a memory strength quantifier and finding it to be zero does not necessarily imply the process to have finite Markov order, since the individual tester element corresponding to the minimum only provides a probabilistic history-blocking, whereas the remaining CP maps belonging to the instrument need not.

The instrument-specific memory strength definitions allow us to deduce the temporal correlations of the process over any time interval with respect to any set of instruments. For a fixed ℓ , minimising any suitable function of $\mathcal{S}_{\mathcal{J}_M}(x_M)$ —such as the average or maximum over outcomes—over all instruments provides a quantification of the *intrinsic* memory strength. That is, the amount of temporal correlations across ℓ timesteps that *cannot* be erased by any interrogation sequence, which vanishes iff the process is of Markov order ℓ . The instrument corresponding to the minimum value, $\text{argmin}_{\mathcal{J}_M}(\mathcal{S}_{\mathcal{J}_M})$, can be interpreted as providing the optimal history-blocking sequence across the given length. On the other hand, $\text{argmax}_{\mathcal{J}_M}(\mathcal{S}_{\mathcal{J}_M})$ relates to the optimal strategy for transmitting information across the memory block; the maximum value provides a novel quantification of the *process capacity*, which is an alternate way of quantifying the channel capacity in the presence of memory [47, 48].

Performing such optimisation is a difficult task in general.¹ Moreover, implementing the optimal sequences in either case may not be possible in practical laboratory setups, as it will generally require the ability to control temporally-correlated quantum operations. However, there are at least two specific instrument sequences that are of immediate relevance, namely the identity instrument and the completely-noisy instrument. The former leaves the system state unchanged and thus relates to the “natural” transmission of information on the system level across the memory block, whereas the latter considers the case where the experimenter tries to actively erase this information by discarding the state emitted by the process and repreparing white noise.

Consequently, we now focus on specifying two particularly important instruments and define the memory strength with respect to them. Note that both the identity map and the completely-noisy instrument have only a single outcome, and thus no aggregation

¹ Since $\mathcal{S}_{\mathcal{J}_M}^{\max}$ simply maximises over all CP maps, and $\mathcal{S}_{\mathcal{J}_M}^{\text{avg}}$ optimises over all CPTP maps—both of which are convex sets—the optimisation should be phraseable as a semidefinite program. However, it is unclear if this is also the case for more general functions of $\mathcal{S}_{\mathcal{J}_M}(x_M)$.

over outcomes is required and we simply have an instrument-specific memory strength denoted $\mathcal{S}_{\mathcal{I}_M}$. We define the *natural* memory strength across a length ℓ as follows.²

Definition 6.4 (Natural memory strength). The natural memory strength across a duration of length $\ell = |M|$ is the mutual information between the history and the future upon application of a sequence of identity maps on the memory block:

$$\mathcal{N}_\ell := \mathcal{S}_{\mathcal{I}_M} = I(F : H)_{\Upsilon_{FH}^{\mathcal{I}_M}}, \quad (6.4)$$

where $\Upsilon_{FH}^{\mathcal{I}_M} := \text{tr}_M [\Psi_M^+ \Upsilon_{FMH}]$, with $\Psi_M^+ = \bigotimes_{j=k-\ell}^{k-1} \Psi_{j \circ j^i}^+$ denoting the Choi operator of the sequence of identity maps.

We also define the *noise-resistant* length- ℓ memory.

Definition 6.5 (Noise-resistant memory strength). The noise-resistant memory strength across a duration of length ℓ is the mutual information between the history and the future upon application of a sequence of maps that discard the state emitted by the process and feed in the maximally-mixed state:

$$\mathcal{R}_\ell := \mathcal{S}_{\mathbb{1}_M} = I(F : H)_{\Upsilon_{FH}}, \quad (6.5)$$

where the mutual information is simply calculated on the future-history marginal of the process tensor, *i. e.*, $\Upsilon_{FH}^{\mathbb{1}_M} = \Upsilon_{FH} = \frac{1}{d_M^o} \text{tr}_M [\Upsilon_{FMH}]$, with the normalisation accounting for the preparation of maximally-mixed states.

This instrument sequence corresponds to applying any **POVM** without recording the outcomes followed by preparing the maximally-mixed state. The Choi operators corresponding to any such erasure sequence amount to the identity matrix on all of the input and output Hilbert spaces associated to each timestep in the memory block. In turn, this implies that \mathcal{R}_ℓ quantifies the amount of memory that survives the erasure on average.³

The notions introduced in this section are depicted in Fig. 6.1. We can either consider a fixed length ℓ and examine how the memory strength behaves over that duration for a variety of instruments, or, on the other hand, consider a fixed instrument and examine the memory as the number of (uncorrelated) sequential applications of said instrument grows with respect to ℓ . We now explicitly construct a process tensor describing an exactly solvable non-Markovian system-environment model in order to study the behaviour

² Since there is only a single outcome for the instruments considered in the following definitions, we depart from the notation introduced in Def. 6.1. Instead, we resort to the notation prevalent in the literature.

³ For processes with vanishing quantum **CMI**, the marginal future-history marginal is separable [137], implying that any faithful entanglement measure $\mathcal{E}(F : H)_{\Upsilon_{FH}}$ (see Ref. [85]) vanishes and any correlation between the history and future upon implementing a length- ℓ erasure sequence is classical.

a) NATURAL



b) NOISE-RESISTANT



c) CAUSAL BREAK (AVG / MAX)

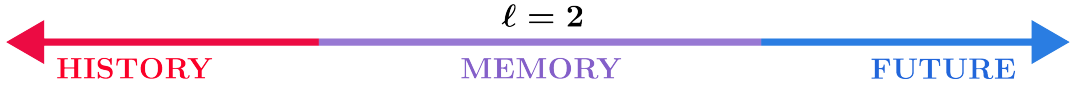


Figure 6.1: *Instrument-specific memory strength.* We depict the instrument-specific memory strength across a length $\ell = 2$. In panel **a)** is the *natural* memory strength, which is calculated with respect to the identity instrument. In panel **b)** is the *noise-resistant* memory-strength, in which an experimenter discards the system states and feeds in white noise. In both cases, the instrument has only one outcome, and the memory strength is simply the mutual information between the history (red) and future (blue) conditional processes. Panel **c)** shows the memory strength with respect to a causal break, where one must choose a suitable aggregate over the mutual information for each outcome, such as averaging or maximising over outcomes.

of both the average and maximum memory strength in Eqs. (6.2) and (6.3) respectively for a causal break sequence, and both the natural and noise-resistant memory strength defined in Eqs. (6.4) and (6.5) respectively.

6.1.2 Memory Length for an Exactly Solvable Model

Here, we focus on a model involving a qubit coupled to another qubit that is interacting with an additional bath (which is regarded as external to the system-environment), as introduced in Ref. [179]. Its low dimensionality makes it easy to solve analytically and therefore permits the investigation of memory effects for different model parameters.

The system-environment evolution proceeds according to the GKSL master equation

$$\frac{\partial \rho_t^{SE}}{\partial t} = -i\xi[\sigma_x^S \otimes \sigma_x^E, \rho_t^{SE}] + \kappa \mathcal{L}[\sigma_-^E](\rho_t^{SE}), \quad (6.6)$$

where the dissipator acts on the environment alone and is defined as: $\mathcal{L}[\sigma_-^E](\rho_t^{SE}) := \sigma_-^E \rho_t^{SE} \sigma_+^E - \frac{1}{2}\{\sigma_+^E \sigma_-^E, \rho_t^{SE}\}$, with $\sigma_\pm^E := \sigma_x^E \pm i\sigma_y^E$. The dynamics describes a qubit system interacting with a qubit environment with X - X coupling strength ξ and a cooling process on the environment due to its interactions with the external bath at rate κ .

The authors of Ref. [179] examined the non-Markovianity of the system dynamics using the breakdown of CP-divisibility, and the increase of the trace-distance distinguishability between arbitrary input states (introduced in Refs. [66] and [92] respectively) as measures of the existence of temporal correlations. Due to the simplicity of the model, the analytic form of the equation of motion on the level of the system alone can be derived and is written [179]

$$\frac{\partial \rho_t^S}{\partial t} = -\frac{\dot{c}_t}{2c_t} \mathcal{L}[\sigma_x^S](\rho_t^S), \quad (6.7)$$

where

$$c_t = \begin{cases} \exp\left(-\frac{\kappa t}{4}\right) \left(\frac{\kappa \sinh \frac{t}{4} \sqrt{\kappa^2 - 64\xi^2}}{\sqrt{\kappa^2 - 64\xi^2}} + \cosh \frac{t}{4} \sqrt{\kappa^2 - 64\xi^2} \right) & \text{for } \kappa^2 > 64\xi^2 \\ \exp\left(-\frac{\kappa t}{4}\right) \left(\frac{\kappa \sin \frac{t}{4} \sqrt{64\xi^2 - \kappa^2}}{\sqrt{64\xi^2 - \kappa^2}} + \cos \frac{t}{4} \sqrt{64\xi^2 - \kappa^2} \right) & \text{for } \kappa^2 < 64\xi^2 \\ \exp\left(-\frac{\kappa t}{4}\right) \left(1 + \frac{1}{4}\kappa t \right) & \text{for } \kappa^2 = 64\xi^2. \end{cases} \quad (6.8)$$

A necessary and sufficient criteria for the dynamics to be CP-divisible is that the coefficients of the dissipation terms in the above master equation for the system, *i. e.*, $-\frac{\dot{c}_t}{2c_t}$, are non-negative for all times [7]. Explicit calculation shows that for $\kappa^2 \geq 64\xi^2$, $-\frac{\dot{c}_t}{2c_t}$ is always non-negative, whereas for $\kappa^2 < 64\xi^2$, $-\frac{\dot{c}_t}{2c_t}$ is negative whenever $\cot \frac{1}{4}t \sqrt{64\xi^2 - \kappa^2} < -\frac{\kappa}{\sqrt{64\xi^2 - \kappa^2}}$. We therefore see an abrupt transition between CP-divisible and non-CP-divisible dynamics across the line $\kappa^2 = 64\xi^2$, as shown in Fig. 6.2.

In the CP-divisible regime, the trace-distance between any two states subject to the evolution is always non-increasing [92]. This fact allows for the quantification of (two-time) non-Markovianity by integrating any increases in the trace distance over all time, which is shown in Ref. [179] to yield the analytic result

$$\mathcal{N}^{\text{Two-time}} = \frac{1}{\exp\left(\frac{\kappa\pi}{\sqrt{64\xi^2 - \kappa^2}}\right) - 1}, \quad (6.9)$$

for $\kappa^2 < 64\xi^2$ and zero otherwise.

However, as discussed, CP-divisibility does not imply Markovianity; as such, Fig. 6.2 does not provide a comprehensive picture of the many prevalent memory effects for

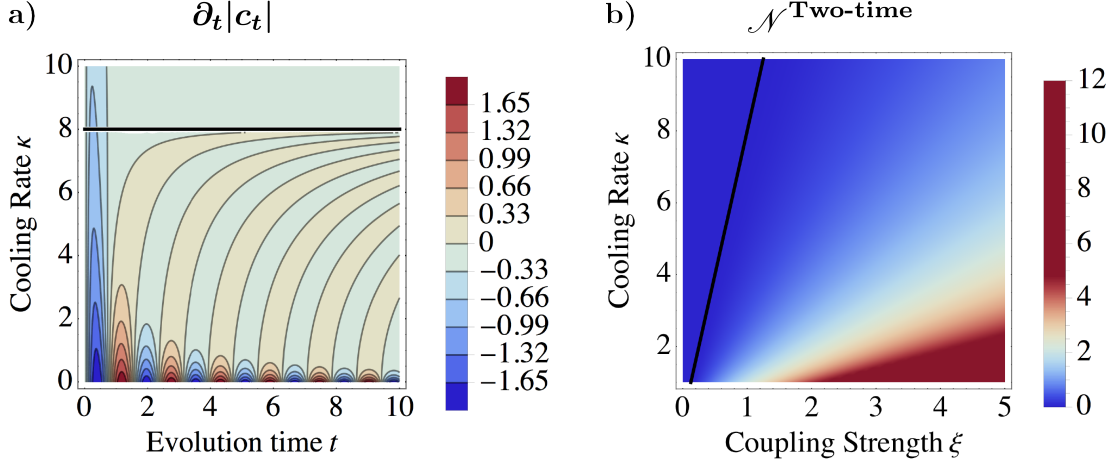


Figure 6.2: Abrupt transition between *CP*-divisible and non-divisible dynamics. In panel a), we plot $\partial_t |c_t|$ with $\xi = 1$. As $\text{sgn}(\frac{\dot{c}_t}{c_t}) = \text{sgn}(\partial_t |c_t|)$, this implies the dynamics is *CP*-divisible for $\kappa \geq 8$, but not for $\kappa < 8$. In particular, there is an abrupt transition along the line $\kappa = 8$. In panel b), we plot the two-time non-Markovianity $\mathcal{N}^{\text{Two-time}}$ as per Eq. (6.9). This is plotted in the parameter space $\xi \in [0, 5]$ and $\kappa \in [1, 10]$ to allow for comparison with Fig. 6.3. Note that this measure of non-Markovianity blows up exponentially for small κ , which is why we have excluded $\kappa \in [0, 1)$ from the plot, and that it vanishes for everything above the black line $\kappa = 8\xi$.

different choices of parameters κ and ξ . Here, we explicitly calculate the process tensor for the dynamics and show that it is non-Markovian for the entire parameter regime, before exploring the behaviour of the instrument-specific memory strength quantifiers introduced in the previous subsection.

We consider a parameter grid $\xi \in [0, 5]$ and $\kappa \in [0, 10]$ with increments of 0.1 in each direction and construct the $n = 6$ step process tensor, $\Upsilon_{6:1}(\xi, \kappa)$. Here, for simplicity, we assume an initially uncorrelated system-environment state, such that the process tensor begins on an output space. We also choose uniform spacing between timesteps of $dt = 0.3$, which corresponds to the natural timescale over which the trace distance between arbitrary initial system states increases for most values in the parameter space [179]. This means that the final time of the process tensor is $T = 1.5$, which corresponds to where the *CP*-divisibility criteria would witness non-Markovianity for a range of parameters.

At each point, we can calculate the non-Markovianity in the process by considering the distance to the nearest Markovian process, as per Eq. (3.40). Here, we choose the (pseudo-)distance to be the relative entropy, $\mathcal{D}(\Upsilon_{6:1}(\xi, \kappa) \parallel \Upsilon_{6:1}^{\text{Markov}})$, in which case the minimum occurs for the Markovian process that is built up from the marginals of the original process tensor [180], *i. e.*, using the relative entropy circumvents the normally

necessary minimisation. The corresponding results are depicted in Fig. 6.3, which indicates that the process is non-Markovian for all parameters in the chosen range. In particular, there is no abrupt transition between regimes. Although the non-Markovianity is small above the line $\kappa = 8\xi$ —at which the dynamics transitions from CP-divisible to non-divisible—it is non-zero, indicating a weak but detectable memory. The two-time witness of non-Markovianity in Eq. (6.9) is insensitive to such effects, which leads to the abrupt transition between regimes; by capturing all multi-time correlations, the non-Markovianity calculated via the process tensor shows this transition to be artificial. This result begs the question: *how long does the memory persist?* We now move to study the behaviour of some of the memory quantifiers proposed in the previous subsection.

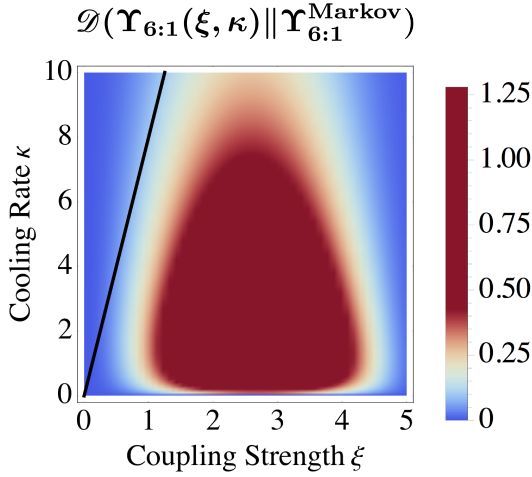


Figure 6.3: Heatmap of non-Markovianity. Non-Markovianity of $\Upsilon_{6:1}(\xi, \kappa)$. Although the non-Markovianity is small above the line $\kappa \geq 8\xi$, it is non-zero. Moreover, there is no abrupt transition between regimes, as all memory effects are accounted for.

To this end, we consider three fixed process tensors: one in the almost Markovian and CP-divisible regime, one in the intermediary regime, and one in the strongly non-Markovian regime, respectively defined as follows

$$\Upsilon^{\text{CP}} := \Upsilon_{6:1}(0.5, 6) \quad (6.10)$$

$$\Upsilon^{\text{Int}} := \Upsilon_{6:1}(3, 8)$$

$$\Upsilon^{\text{SNM}} := \Upsilon_{6:1}(3, 2).$$

We first consider, for fixed $\ell \in \{1, \dots, 4\}$ ranging from t_2 to t_5 , the natural and noise-resistant memory strength defined in Eqs. (6.4) and (6.5) respectively. We also consider the memory strength for length- ℓ sequences of causal breaks, by

first constructing the collection of outcome-specific values in Eq. (6.1) and subsequently calculating the average and maximum memory strength in Eqs. (6.2) and (6.3) for the instrument. The causal break is chosen to be a symmetric single-qubit IC POVM (defined in Example D.1) followed by the independent reparation into one of an IC set of states $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle^x\langle +|^x, |+\rangle^y\langle +|^y\}$, where $|+\rangle^{x/y}$ is the +1 eigenstate of σ_x/σ_y .

The results are summarised in Fig. 6.4. Interestingly, all three of these processes have vanishing memory strength with respect to the completely noisy instrument, meaning that an experimenter can erase the temporal correlations in the process by acting at a single timestep. The other three quantities do not vanish and display stark differences.

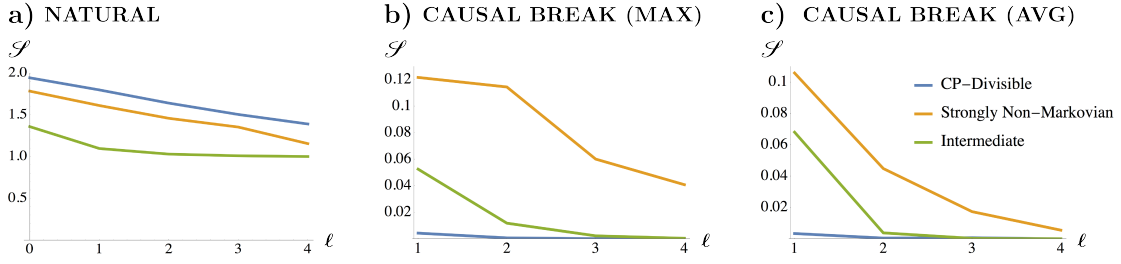


Figure 6.4: *Instrument-specific memory strength.* Here we plot the natural memory strength (panel **a**), and both the maximum and average memory strength with respect to a causal break sequence (panels **b**) and **c**) respectively) as a function of ℓ . Note the changing vertical scales and the legend on the right. Interestingly, the process tensor in the CP-divisible regime (blue) has the strongest memory strength with respect to sequences of the identity instrument, followed by that in the strongly non-Markovian regime (yellow) and lastly that in the intermediate regime (green). By comparing panel **a**) with panels **b**) and **c**), it is clear that for the CP-divisible process, most of the memory strength arises by way of the identity instrument transmitting information on the level of the system; when active probing interventions such as those of a causal break stop this flow of information, its memory strength becomes negligible. On the other hand, for the intermediate and strongly non-Markovian processes, significant multi-time memory effects are exhibited for all instruments shown.

All three processes display the strongest memory strength with respect to the identity, suggesting that a non-negligible amount of memory is transmitted on the level of the system alone by way of the identity operator. Moreover, it is the process tensor in the CP-divisible regime that exhibits the strongest natural memory strength, with that in the intermediate regime displaying the weakest. However, the effects of active interventions come to light when we consider the causal break instrument. Here, the CP-divisible process tensor displays almost vanishing memory strength with respect to either aggregation, with the intermediate process showing some memory effects and the strongly non-Markovian one exhibiting stronger temporal correlations still. This is not surprising, as a causal break acts to block the flow of information on the level of the system, which should effectively wipe-out any temporal correlations in a CP-divisible process. In all cases, the memory strength decays across longer memory blocks, as is intuitively expected.

6.2 CHAPTER SUMMARY

In this chapter we have proposed a number of instrument-specific definitions of memory strength. By accounting for all multi-time memory effects that are potentially present in a process, these definitions serve to unambiguously characterise how strong the temporal correlations between the history and the future are for any chosen sequence of instruments applied by an experimenter. Indeed, such instrument-specific notions of memory strength are directly relevant for experimental applications and computational simulation techniques: they provide an operational approach that permits making memory cutoff approximations of choice, rather than relying on the natural timescales of decay imposed by the system-environment dynamics, as are prevalent throughout numerical techniques [24]. Indeed, techniques of a similar flavour are being developed through extensions of the transfer tensor formalism [178, 181–184], bridging the connection between characterisation and efficient simulation of quantum processes with memory.

In studying the exactly solvable model, we showed a prime example of how this might be applied in practice, by examining how the behaviour of memory strength varies for different instruments and across different timescales. By tuning the parameters of the model appropriately, one could simulate dynamics that is amenable to short-time memory approximations with respect to sequences of causal breaks, for instance, by *e.g.*, constructing the CP-divisible process of Eq. (6.10). Indeed, the decay of memory effects over longer sequences of active interventions is related to various operational protocols such as dynamical decoupling [11], erasure of information or transmission of information [47, 48]. The results developed here for the simple two-qubit model are already interesting in their own right, and are suggestive of the possible insights that may be uncovered by considering the memory strength and related approximation techniques for appropriate choices of instruments in a variety of situations.

Part IV

ENVOI

You can trip off to places so wild and so wiggy that you don't know where you are until you get back. And sometimes not even know you tripped off at all because you never get back to know that you've left...

— Ken Kesey, *Sometimes a Great Notion*.

SUMMARY

NON-NEGLIGIBLE MEMORY EFFECTS PLAY A CRUCIAL ROLE in phenomena studied throughout a vast range of sciences—physical, biological, chemical, neurological, economical, ...; the list goes on. Many such effects are highly complex and their treatment lies beyond the cutting edge of current knowledge and experimental reach; on the other hand, when adequately understood, memory can be manipulated in order to develop state-of-the-art technologies. Thus, a proper theoretical description of processes with memory is of immense practical importance, marking *a kind of glory* that is deserving of pursuit.

To develop such a comprehension requires, quite naturally, a thorough understanding of the limitations of the prevailing descriptions. This was the focal point of Chapter 2: we first studied classical stochastic processes, which are well-understood and can easily be defined in an unambiguous way. This served as a point of reference against which open quantum evolution stands in juxtaposition when multi-time correlations are of importance. Up until recently, the murkiness surrounding quantum stochastic processes has led to ambiguity in defining many key concepts, culminating in—amongst other confusions—a myriad of incompatible definitions of memoryless quantum processes.

This comparison also made clear that the critical problem is one of formalism, rather than anything fundamental: two-time descriptions of stochastic processes, as they are used ubiquitously in the literature, are inadequate to describe general processes with memory. This is a statement of logic that transcends the physical theory to which it applies, holding equally true in both the classical and quantum realms. In Chapter 3 we saw that in order to capture multi-time correlations in a quantum process, due to the invasive nature of measurement at the nanoscale, we must clearly delineate between the role of the experimenter and that of the underlying physical evolution. This line of thinking naturally led us to a clear definition of quantum stochastic processes by way of the process tensor formalism.

The process tensor ameliorates the aforementioned obstacles, at once providing a linear, completely-positive and trace-preserving (in the correct sense) description that completely captures any causally-ordered evolution on a discrete set of timesteps admissible by quantum theory. Its definition stems from the dilated joint unitary evolution, which offers an intuitive connection to the physical mechanism that drives the process. By abstracting the uncontrollable process at hand from the sequence of controllable interrogations an experimenter might choose to apply, this operational picture captures all possible multi-time statistics that can be deduced with respect to any valid probing schema conceivable. Thereby, it both unifies and generalises previous definitions of memorylessness for quantum processes.

Thus, Part II summarises the story so far, leading us to a *narrative on the edge*: with the correct tools for describing general stochastic processes with memory, we stand at the precipice of a univocal study of memory effects in quantum processes, which has hitherto proved elusive.

The previously *unspoken words* of this thesis follow in Part III. Since the general description of stochastic processes grows, with respect to the length of time considered, to rapidly defy reasonably available computational resources—both in classical and quantum theory—our step towards understanding processes with memory concentrates on examining those for which the memory length is finite in duration. Fortunately, such processes can be efficiently modelled by only taking into account information regarding the most recent ℓ timesteps, rather than the entirety of history, when making predictions. Classically, this scenario is captured by the concept of Markov order, which provides a characteristic timescale for the memory length of any stochastic process.

In Chapter 4 we extended Markov order to quantum mechanics. While the idea remains unchanged from the standard intuition, the corresponding phenomenology of quantum processes with finite memory is significantly richer and more complex than its classical counterpart. As in the classical case, we asked if the future evolution of the system can be deduced, in principle, entirely from the most recent sequence of measurement outcomes. However, for quantum stochastic processes, the Markov order—and therefore the memory length of the process—is fundamentally dependent on the instruments used to probe the process.

We formulated the conditions that capture this concise and comprehensive definition of memory length for quantum processes, which is fully reducible to its classical counterpart in the appropriate limit, in terms of a constraint on the process tensor. From this, we

saw that no quantum process with memory can have finite Markov order with respect to all possible instruments.

In addition, our analysis highlighted that this dependence of Markov order on the interrogating instruments persists even in classical physics, as soon as active interventions on the system are permitted; in quantum theory, this issue is fundamental and must be acknowledged. In short, quantum processes with memory exhibit distinct memory effects when probed with different instruments. Our work provides the first formal classification of such behaviour.

In Chapter 5, we analysed the structure of quantum stochastic processes that display finite memory effects. We detailed the structural constraints that must be satisfied for the underlying process to have finite Markov order with respect to a given instrument sequence, including some specific natural and experimentally-relevant classes of probing instruments, such as unitary operations and an informationally-complete set of measurements and re-preparations. The connection between finite memory-length and the underlying system-environment dynamics was elucidated through a series of pedagogical examples, which served to outline a broad taxonomy of the memory effects possible in non-Markovian quantum processes. Lastly, we explored the relationship between quantum Markov order and vanishing quantum conditional mutual information, showing explicitly that although the latter implies the former, processes with finite quantum Markov order need not have vanishing quantum CMI.

In Chapter 6, we proposed instrument-specific measures of the memory strength for a quantum process. By explicitly analysing the behaviour of the memory strength for a simple model dynamics with respect to a range of instruments over varying timescales, we uncovered numerous interesting insights. For instance, the approaches developed here can be seen as an operational way to make finite-memory approximations to numerically simulate complex dynamics; or, from the perspective of information-processing, many of the notions discussed are related to finding the optimal sequence of operations to perform a given task. The intriguing results for the toy-model studied already allude to tantalising possibilities that might be harnessed by using such techniques to understand the strength and complexity of memory in quantum processes.

In conclusion, the results of this thesis solve the long-standing problem of extending the definition and quantification of memory effects from the classical to the quantum realm. We expect our approach to fundamentally shift the way memory in quantum processes is described by the community for a number of reasons, including—but not limited to—the following. For one, instrument-dependence of effects measured in quan-

tum processes is close to experimental reality, as the boundary of what is experimentally accessible shifts continuously as technology evolves. Moreover, while not all real-world processes will display finite Markov order, many processes are likely to do so approximately, and the developed insights will be of fundamental importance for their experimental reconstruction and simulation.

7.1 OUTLOOK

We now move to discussing some of the broader implications of our work. On the foundational side, it is clear that, upon the set of timesteps on which it is defined, the process tensor provides the most generic description of causally-ordered processes allowable within quantum theory. Examining properties of its structure, as we have in this thesis, provides fundamental insight into understanding the space of quantum processes and temporal correlations. Indeed, similar frameworks that do away with the axiom of causality, such as those based on the process matrix, are actively being developed to study the most general spatio-temporal correlations allowable [32, 33, 37], shedding light on the distinguishing features of classical, quantum and post-quantum theories.

On the practical side, the process tensor contains all the information one could ever hope to learn about a process. This, unfortunately, can make it computationally daunting to approach. In light of this, its usefulness lies in our ability to develop compression and extraction methods to approximate complex physical evolutions with overlapping process tensors of finite length for efficient simulation of long-term dynamics. Indeed, this is the flavour of many methods proposed throughout the literature, such as the transfer tensor approach [178, 181–184]. A deeper understanding of optimal compression and recovery schemes for processes with approximately finite memory length will have significant consequences for efficient quantum simulation.

Moreover, understanding memory effects has immediate relevance to developing near-term quantum technologies, particularly concerning the construction of error-correcting codes to combat correlated noise [64, 185–189] and the design of feedback protocols for coherent control [169, 170]. The far-reaching implications of these possible avenues of exploration highlight the substantial relevance of our novel approach for researchers interested in open quantum systems, quantum control, stochastic modelling, complexity science and quantum foundations, amongst other fields.

Part V

APPENDIX

The reverberation often exceeds through silence the sound that sets it off; the reaction occasionally outdoes by way of repose the event that stimulated it; and the past not uncommonly takes a while to happen, and some long time to figure out.

— Ken Kesey, *Sometimes a Great Notion*.



NOTATION SUMMARY

The mathematical underpinnings of quantum theory involves linear operators on a Hilbert space. We restrict ourselves to finite-dimensional Hilbert spaces; as such, linear operators can be represented as matrices. The notational conventions employed throughout this thesis are summarised in Table A.1 below. Whenever it is unambiguous, to avoid notational clutter, we drop timestep or subsystem labels. For the same reason, we avoid brackets wherever possible, with maps acting on everything to the right of them by convention. An Hermitian operator X is called positive (written $X \geq 0$) whenever its spectrum contains only non-negative values. We use $\log := \log_2$ and set $\hbar = 1$.

Lastly, note the colour-coding schema we employ throughout the figures in this thesis. We consistently use green to denote preparations, transformations and measurements that are controllable by an experimenter; yellow to denote those that are uncontrollable; and orange to represent components of the non-unique dilation of an uncontrollable process. When we wish to describe maps from an abstract mathematical perspective, we colour the components with red, blue or purple to minimise any possible confusion.

GENERAL

$\mathbb{C}, \mathbb{R}, \mathbb{N}$	complex, real and natural numbers
t_j, t_k, \dots	lowercase letters used as subscripts represent timesteps
$k : j$	an ordered sequence of timesteps $\{t_j, \dots, t_k\}$ for $k > j \in \mathbb{N}$
Λ_n	a set of timesteps of cardinality $n \in \mathbb{N}$ (not necessarily an ordered sequence, <i>e. g.</i> , Λ_3 could denote $\{t_1, t_4, t_5\}$)
$\mathbb{P}_{k:j}, \mathbb{P}_{\Lambda_n}$	a joint probability distribution defined on the timesteps indicated by the subscript
\mathbf{i}, \mathbf{o}	input and output spaces associated to each timestep

NOTATION SUMMARY

A, B, C, \dots	systems are labelled with capital letters
\mathcal{H}^A	Hilbert space associated to the system A
$\dim(\mathcal{H}^A) =: d_A$	dimension of the Hilbert space \mathcal{H}^A
$\langle \bullet , \bullet \rangle$	bra and ket
tr, tr_A	trace and partial trace over \mathcal{H}^A
A^*, A^T, A^\dagger	conjugate, transpose and conjugate transpose of A
$\text{Spec}(A)$	spectrum of A
$[A, B]$	commutator between A and B , <i>i. e.</i> , $[A, B] := AB - BA$
$\{A, B\}$	anti-commutator between A and B , <i>i. e.</i> , $\{A, B\} := AB + BA$
$A \otimes B$	tensor product of A and B
$A \oplus B$	direct sum of A and B
<hr/>	
OPERATORS	
<hr/>	
$\text{BL}(\mathcal{H}^A)$	set of bounded linear operators on \mathcal{H}^A
$\mathcal{A}, \mathcal{B}, \mathcal{C} \dots$	maps are denoted by calligraphic capital letters and act on everything to the right of them
$A, B, C \dots$	the Choi operators of maps are denoted by their sans-serif counterparts
$\mathcal{I}^A, \mathbb{1}^A$	identity map and identity operator on \mathcal{H}^A respectively
<hr/>	
ENTROPIES	
<hr/>	
$H^{\text{cl}}(\mathbb{P})$	Shannon entropy of the probability distribution \mathbb{P} , <i>i. e.</i> , $H^{\text{cl}}(\mathbb{P}) := -\sum_{x \in \mathcal{X}} \mathbb{P}(x) \log \mathbb{P}(x)$
$I^{\text{cl}}(A : B)$	classical mutual information between A and B , <i>i. e.</i> , $I^{\text{cl}}(A : B) := H(\mathbb{P}_A) + H(\mathbb{P}_B) - H(\mathbb{P}_{AB})$
$I^{\text{cl}}(A : C B)$	classical conditional mutual information between A and C given B , <i>i. e.</i> , $I^{\text{cl}}(A : C B) := H^{\text{cl}}(\mathbb{P}_{AB}) + H^{\text{cl}}(\mathbb{P}_{BC}) - H^{\text{cl}}(\mathbb{P}_{ABC}) - H^{\text{cl}}(\mathbb{P}_B)$

$\mathcal{D}^{\text{cl}}(\mathbb{P} \mathbb{Q})$	relative entropy (Kullback-Liebler divergence) between two distributions \mathbb{P} and \mathbb{Q} , <i>i. e.</i> , $\mathcal{D}^{\text{cl}}(\mathbb{P} \mathbb{Q}) := -\sum_{x \in \mathcal{X}} \mathbb{P}(x) \log \frac{\mathbb{P}(x)}{\mathbb{Q}(x)}$
$S(\rho)$	von Neumann entropy of the density operator ρ , <i>i. e.</i> , $S(\rho) := -\text{tr} [\rho \log \rho] = -\sum_{\lambda \in \text{Spec}(\rho)} \lambda \log \lambda$
$I(A : B)$	quantum mutual information between A and B , <i>i. e.</i> , $I(A : B) := S(\rho^A) + S(\rho^B) - S(\rho^{AB})$
$I(A : C B)$	quantum conditional mutual information between A and C given B , <i>i. e.</i> , $I(A : C B) := S(\rho^{AB}) + S(\rho^{BC}) - S(\rho^{ABC}) - S(\rho^B)$
$\mathcal{D}(\rho \sigma)$	quantum relative entropy between ρ and σ , <i>i. e.</i> , $\mathcal{D}(\rho \sigma) := \text{tr} [\rho \log \rho - \rho \log \sigma]$
<hr/>	
OTHER / EXCEPTIONS	
<hr/>	
H, M, F	reserved to denote the collections of timesteps grouped as the history $\{t_1, \dots, t_{k-\ell-1}\}$, memory $\{t_{k-\ell}, \dots, t_{k-1}\}$ and the future $\{t_k, \dots, t_n\}$ at arbitrary timestep t_k for a memory of length ℓ
U, V	reserved to denote the unitary matrices associated to unitary maps, <i>i. e.</i> , $\mathcal{U}(\bullet) := U(\bullet)U^\dagger$
$\mathbf{W}, \mathbf{W}^\perp$	boldface capital letter represents a vector space, with the \perp superscript denoting its orthogonal complement
$\hat{\rho}$	a caret is used to denote elements of a fixed basis of a vector space
\mathcal{I}	an instrument is a collection of completely-positive maps that sums to a completely-positive and trace-preserving map (these can be higher order objects defined across multiple timesteps, referred to as instrument sequences or testers)
Υ	the Choi operator of the process tensor whose mapping is represented by \mathcal{T}
$\tilde{\Upsilon}^{(x)}$	a tester element of the process tensor Υ , <i>i. e.</i> , $\sum_x \tilde{\Upsilon}^{(x)} = \Upsilon$, with Υ a proper process tensor

NOTATION SUMMARY

$\underline{\Upsilon}$	the projection of $\Upsilon \in \mathbf{W} \oplus \mathbf{W}^\perp$ on the \mathbf{W}^\perp subspace
$\psi^+, \psi^-, \phi^+, \phi^-$	normalised maximally entangled two-party Bell states
$\Psi^+, \Psi^-, \Phi^+, \Phi^-$	unnormalised maximally entangled two-party Bell operators, which are respectively the Choi operators corresponding to the channels that implement the four Pauli rotations $\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$
$X, x, \mathcal{X}; Y, y, \mathcal{Y}$	reserved to denote random variables (capital Roman) and realisations of said random variable (lowercase), which take values from some set (capital script)
H, T	heads and tails, symbolic of the possible outcomes of any binary experiment
δ	the Dirac-delta distribution

Table A.1: *Notational conventions.* A summary of the notational conventions employed.

CLASSICAL AND QUANTUM DYNAMICS WITH NOISE

B.1 CLASSICAL MASTER EQUATION

Given a Markovian process, we can consider the first three timesteps and immediately marginalise over the intermediate variable of the joint distribution to derive

$$\begin{aligned}\mathbb{P}_{3,1}(x_3, x_1) &= \sum_{x_2} \mathbb{P}_{3:1}(x_3, x_2, x_1) \\ &= \sum_{x_2} \mathbb{P}_3(x_3|x_2) \mathbb{P}_2(x_2|x_1) \mathbb{P}_1(x_1).\end{aligned}\tag{B.1}$$

Dividing both sides by $\mathbb{P}_1(x_1)$ yields the *Chapman-Kolmogorov* equation [4]

$$\mathbb{P}_3(x_3|x_1) = \sum_{x_2} \mathbb{P}_3(x_3|x_2) \mathbb{P}_2(x_2|x_1).\tag{B.2}$$

This equation expresses that for a Markovian process that begins with value x_1 at t_1 and reaches x_3 at t_3 , it must do so in a manner that is specified given knowledge of the value x_2 at the intermediate timestep t_2 .

We can recast this property into an integro-differential equation that the transition probabilities for any Markovian process must satisfy, known as the [ME](#). To do so, for the moment we consider time as a continuous parameter. We further assume that the process is homogeneous in time, meaning that the conditional distributions connecting adjacent timesteps do not explicitly depend on the times themselves, rather, only their difference. This allows us to define

$$\mathbb{P}_2(x_2|x_1) =: S_{\tau'}(x'|x'') \quad \text{and} \quad \mathbb{P}_3(x_3|x_2) =: S_{\tau}(x|x'),\tag{B.3}$$

where $\tau' := t_2 - t_1$, $\tau := t_3 - t_2$, and we have dropped the (now redundant) subscript labels on the outcomes, with the understanding that in each transition probability, the conditioning argument always represents a value measured prior to the remaining argument by the time difference denoted, *i. e.*, $S_{\tau}(x|x')$ refers to the probability that x will be

observed given that x' was the outcome τ units of time ago. The Chapman-Kolmogorov equation is then expressed as

$$S_{\tau+\tau'}(x|x'') = \sum_{x'} S_{\tau}(x|x') S_{\tau'}(x'|x''). \quad (\text{B.4})$$

Considering the interval τ to be small and expanding $S_{\tau}(x|x')$ about $\tau = 0$ gives

$$S_{\tau}(x|x') = (1 - a(x)\tau)\delta(x, x') + \tau W(x|x') + \mathcal{O}(\tau^2), \quad (\text{B.5})$$

where $(1 - a(x)\tau)$ represents the probability that the state does not transition from x' to x during the short time interval τ , $W(x|x') := \frac{\partial S_{\tau}(x|x')}{\partial \tau}|_{\tau=0} \geq 0$ is the instantaneous rate for this transition to occur, and $\mathcal{O}(\tau^2)$ represents higher-order terms that eventually vanish upon taking the limit $\tau \rightarrow 0$. By the normalisation constraint on $S_{\tau}(x|x')$, it follows that

$$a(x) = \sum_{x'} W(x'|x). \quad (\text{B.6})$$

Substituting Eqs. (B.5) and (B.6) into Eq. (B.4) and subsequently taking the limit $\tau \rightarrow 0$ yields the classical [ME](#)

$$\frac{\partial S_t(x|x'')}{\partial t} = \sum_{x'} \{W(x|x') S_t(x'|x'') - W(x'|x) S_t(x|x'')\}. \quad (\text{B.7})$$

B.2 TOMOGRAPHY OF A DYNAMICAL MAP

The tomographic representation of a dynamical map relies on the concept of duals. Considering the tomographic protocol outlined in Subsection 2.2.2, we begin with a basis set of states $\{\hat{\rho}^{(i)}\}$. Although these need not be orthonormal, they are linearly independent, and hence there exists a dual set of objects $\{\hat{D}^{(i)}\}$ such that

$$\text{tr} [\hat{D}^{(i)\dagger} \hat{\rho}^{(j)}] = \delta_{ij} \quad \forall i, j. \quad (\text{B.8})$$

These can be constructed explicitly as follows [51, 74]. Begin by writing $\hat{\rho}^{(i)} = \sum_j h_{ij} \Gamma^{(j)}$, where $h_{ij} \in \mathbb{C}$ and $\{\Gamma^{(j)}\}$ are a Hermitian, self-dual linearly independent set of operators satisfying $\text{tr} [\Gamma^{(i)} \Gamma^{(j)}] = 2\delta_{ij}$; for instance, these could be chosen as the generalised Pauli basis. Since $\{\hat{\rho}^{(i)}\}$ are linearly independent, the columns of the matrix $H = \sum_{ij} h_{ij} |i\rangle\langle j|$ are linearly independent vectors, which means that H is invertible. Let $F^{\dagger} = H^{-1}$ so that $HF^{\dagger} = \mathbb{1}$, implying that the columns of F^* are orthonormal to those of H . Finally, define $\hat{D}^{(i)} = \frac{1}{2} \sum_j f_{ij} \Gamma^{(j)}$, where f_{ij} are the entries of F .

For example, in the case of a qubit system, we can use

$$\begin{aligned}\hat{\rho}^{(1)} &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, & \hat{\rho}^{(2)} &= \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}, \\ \hat{\rho}^{(3)} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, & \hat{\rho}^{(4)} &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.\end{aligned}\tag{B.9}$$

Although they are not orthonormal, these matrices are linearly independent and span the space of qubits. The dual set to the basis defined in Eq. (B.9) is given by

$$\begin{aligned}\hat{D}^{(1)} &= \frac{1}{2} \begin{bmatrix} 0 & 1+i \\ 1-i & 2 \end{bmatrix}, & \hat{D}^{(2)} &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \\ \hat{D}^{(3)} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & \hat{D}^{(4)} &= \frac{1}{2} \begin{bmatrix} 0 & -1+i \\ -1-i & 2 \end{bmatrix}.\end{aligned}\tag{B.10}$$

MEMORY LENGTH

C.1 QUANTUM MARKOV ORDER CONSTRAINT ON PROCESS TENSOR

Here we prove that Def. 4.1 is equivalent to Eq. (4.4). Explicitly writing out the conditioning over outcomes stipulated in Eq. (4.1) gives

$$\begin{aligned} & \frac{\mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H)}{\sum_{x_F} \mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H)} \\ &= \frac{\sum_{x_H} \mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H)}{\sum_{x_F x_H} \mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H)}. \end{aligned} \quad (\text{C.1})$$

On the l.h.s, we can immediately make use of the causal structure of the process tensor to simplify the denominator: since the choice of instruments in the future cannot overall influence the statistics on the history and memory blocks, we have, for any \mathcal{J}_F ,

$$\begin{aligned} \sum_{x_F} \mathbb{P}_{FMH}(x_F, x_M, x_H | \mathcal{J}_F, \mathcal{J}_M, \mathcal{J}_H) &= \text{tr} \left[\left(\mathcal{O}_F \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right] \\ &= \text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{MH} \right], \end{aligned} \quad (\text{C.2})$$

where $\Upsilon_{MH} := \frac{1}{d_{F^\circ}} \text{tr}_F [\Upsilon_{FMH}]$ with $d_{F^\circ} := \prod_{j=k}^n d_{j^\circ}$ denoting the joint dimension of the output spaces associated to F and we employ the previously introduced notation for the **CPTP** map corresponding to an instrument $\mathcal{O}_X := \sum_{x_X} \mathcal{O}^{(x_X)}$ (or, more precisely, the overall deterministic comb corresponding to the tester implemented on the future). Note that we have dropped the explicit labelling of the instrument that the **CPTP** map \mathcal{O}_X is associated with for compactness; of course, different instruments (in general) correspond to different overall **CPTP** maps. The fact that Υ_{MH} is a proper, *i. e.*, causally-ordered, process tensor can be seen by simply following the hierarchy of trace conditions assumed to be satisfied by Υ_{FMH} from the future backwards (see Eq. (3.23)). We cannot use a similar trick on the numerator of the r.h.s, since, in general, the statistics observed over the memory and future timesteps depend upon the choice of instrument implemented

across the history; although the denominator can be simplified in a likewise manner. Then, expressing Eq. (C.1) in terms of the process tensor and making use of Eq. (C.2), we have

$$\begin{aligned} & \frac{\text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right]}{\text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{MH} \right]} \\ &= \frac{\text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H \right)^T \Upsilon_{FMH} \right]}{\text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H \right)^T \Upsilon_{MH} \right]}. \end{aligned} \quad (\text{C.3})$$

The tensor product structure of Eq. (4.4) is clearly a sufficient condition for Eq. (4.1). Note that the l.h.s of Eq. (C.3) represents $\mathbb{P}(x_F | \mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H)$; considering performing the trace over M first yields

$$\begin{aligned} & \mathbb{P}(x_F | \mathcal{J}_F; x_M, \mathcal{J}_M; x_H, \mathcal{J}_H) \\ &= \frac{\text{tr}_{FH} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right] \right]}{\text{tr}_H \left[\mathcal{O}_H^{(x_H)T} \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{MH} \right] \right]} \\ &= \frac{\text{tr}_{FH} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_F^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)} \right]}{\text{tr}_H \left[\mathcal{O}_H^{(x_H)T} \tilde{\Upsilon}_H^{(x_M)} \right]} \\ &= \text{tr}_F \left[\mathcal{O}_F^{(x_F)T} \Upsilon_F^{(x_M)} \right] = \mathbb{P}(x_F | \mathcal{J}_F; x_M, \mathcal{J}_M). \end{aligned} \quad (\text{C.4})$$

We now consider the converse direction. Eq. (C.3) must hold for all instruments \mathcal{J}_H , and since we can vary the CP maps while keeping the overall CPTP map of the instrument fixed, this implies that we must have

$$\begin{aligned} & \frac{\text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right]}{\text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{MH} \right]} \\ &= \frac{\text{tr} \left[\left(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H'^{(x_H')} \right)^T \Upsilon_{FMH} \right]}{\text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H'^{(x_H')} \right)^T \Upsilon_{MH} \right]}, \end{aligned} \quad (\text{C.5})$$

for all CP maps $\mathcal{O}_H^{(x_H)}, \mathcal{O}_H'^{(x_H')}$.

We can simplify the numerator on both sides by defining the conditional future process tensor

$$\Upsilon_F^{(x_M, x_H)} := \frac{\text{tr}_{MH} \left[\left(\mathbb{1}_F \otimes \mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{FMH} \right]}{\mathbb{P}(x_M, x_H | \mathcal{J}_M, \mathcal{J}_H)}, \quad (\text{C.6})$$

where

$$\mathbb{P}(x_M, x_H | \mathcal{J}_M, \mathcal{J}_H) = \text{tr} \left[\left(\mathcal{O}_M^{(x_M)} \otimes \mathcal{O}_H^{(x_H)} \right)^T \Upsilon_{MH} \right]. \quad (\text{C.7})$$

This leads to

$$\text{tr} \left[\mathcal{O}_F^{(x_F)T} \Upsilon_F^{(x_M, x_H)} \right] = \text{tr} \left[\mathcal{O}_F^{(x_F)T} \Upsilon_F^{(x_M, x'_H)} \right]. \quad (\text{C.8})$$

Since this must hold true for all future instruments \mathcal{J}_F and the trace corresponds to an inner product on the space on which the Choi operators are defined, we finally have

$$\Upsilon_F^{(x_M, x_H)} = \Upsilon_F^{(x_M, x'_H)} \quad \forall x_H, x'_H, \quad (\text{C.9})$$

which implies that the conditional future process tensor defined in Eq. (C.6) is independent of the historic outcomes, *i. e.*, $\Upsilon_F^{(x_M, x_H)} = \Upsilon_F^{(x_M, x'_H)} = \Upsilon_F^{(x_M)}$.

Returning to Eq. (C.6) and considering first taking the partial trace over M , we have

$$\frac{\text{tr}_H \left[\left(\mathbb{1}_F \otimes \mathcal{O}_H^{(x_H)} \right)^T \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right] \right]}{\text{tr}_H \left[\mathcal{O}_H^{(x_H)T} \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{MH} \right] \right]}. \quad (\text{C.10})$$

It is clear from the expression above that in order for it to be independent of the outcome x_H , it must be the case that $\text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right]$ splits into a tensor product, *i. e.*, $\text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right] = \Upsilon_F^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)}$, where

$$\Upsilon_F^{(x_M)} := \frac{\text{tr}_{MH} \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right]}{\text{tr} \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{MH} \right]}, \quad (\text{C.11})$$

and

$$\tilde{\Upsilon}_H^{(x_M)} := \text{tr}_{FM} \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right]. \quad (\text{C.12})$$

Note that, by construction, the conditional future process defined in Eq. (C.11) is a proper process tensor for each x_M , *i. e.*, a positive semi-definite operator satisfying the hierarchy of trace conditions of Eq. (3.23). Indeed, it is easy to see that the denominator normalises the expression such that for each x_M we have $\text{tr} \left[\Upsilon_F^{(x_M)} \right] = d_{F^*}$ (which is necessary for satisfaction of Eq. (3.23)). Conversely satisfaction of Eq. (3.23) for the process tensor Υ_{FMH} means that Υ_{MH} is a well-defined, proper process tensor. On the other hand, such a normalisation for the historic part of the process cannot be achieved, since an object such as $\text{tr}_H \left[\Upsilon_{FMH} \right]$ does not represent a well-defined process tensor, as it implicitly dictates a specific choice of instrument on the history (namely, the sequence of information trashing identity operators) that influence the future statistics. This is

in contrast to the definition of the conditional future process, where causality ensures that any such implicit choice of instrument on the future is irrelevant to the statistics measured on the history, which permits for the appropriate normalisation to ensure that each $\Upsilon_F^{(x_M)}$ is a proper process tensor. Thus, $\tilde{\Upsilon}_H^{(x_M)}$ is not necessarily a proper process tensor. Nonetheless, from this object we can calculate the probability to measure certain outcomes on the memory block given a choice of instrument on the memory and the history

$$\mathbb{P}(x_M | \mathcal{J}_M, \mathcal{J}_H) = \text{tr} \left[\mathbf{O}_H^T \tilde{\Upsilon}_H^{(x_M)} \right]. \quad (\text{C.13})$$

By the normalisation of total probability, we have that $1 = \sum_{x_M} \mathbb{P}(x_M | \mathcal{J}_M, \mathcal{J}_H) = \sum_{x_M} \text{tr} \left[\mathbf{O}_H^T \tilde{\Upsilon}_H^{(x_M)} \right]$, which must hold for all CPTP \mathbf{O}_H . This implies that when summed over outcomes, the conditional history process forms a proper process tensor, *i.e.*, $\sum_{x_M} \tilde{\Upsilon}_H^{(x_M)} = \Upsilon_H^{(x_M)}$ satisfies Eq. (3.23). In other words, each $\tilde{\Upsilon}_H^{(x_M)}$ is a post-selected tester element of the process tensor describing the history, which is precisely what is indicated by the tilde.

C.2 DEMANDING FINITE QUANTUM MARKOV ORDER FOR ALL INSTRUMENTS IMPLIES MARKOVIANITY

Here we prove Theorem 4.2. We begin with the following Lemma:

Lemma C.1. *The only operators Υ_{FMH} that satisfy Eq. (4.4) for all possible instruments \mathcal{J}_M are those where the M subsystem is in tensor product with F or H (or both).*

Choose a linearly independent, IC set of operators $\mathcal{J}_M = \{\mathbf{O}_M^{(x)}\}$ as the instrument on M . Any linearly independent set has an associated dual set of operators $\{\Delta_M^{(y)}\}$ such that $\text{tr} \left[\mathbf{O}_M^{(x)} \Delta_M^{(y)\dagger} \right] = \delta_{xy} \forall x, y$. Thus, we can write any tripartite state satisfying Eq. (4.4) for each measurement outcome as follows

$$\Upsilon_{FMH} = \sum_x \Upsilon_F^{(x)} \otimes \Delta_M^{(x)*} \otimes \tilde{\Upsilon}_H^{(x)}. \quad (\text{C.14})$$

Now, consider a different instrument comprising a set of projectors defined via a linear expansion of the original set $\mathcal{J}'_M = \{\mathbf{Q}_M^{(y)} := \sum_x q_{xy} \mathbf{O}_M^{(x)}\}$, with $\{q_{xy}\}$ some non-trivial coefficients. The conditional process upon realisation of any outcome of this instrument is

$$\begin{aligned}
 \tilde{\Upsilon}_{FH}^{(y)} &= \text{tr}_M \left[\mathbf{Q}_M^{(y)\text{T}} \Upsilon_{FMH} \right] \\
 &= \sum_x \Upsilon_F^{(x)} \otimes \tilde{\Upsilon}_H^{(x)} \text{tr} \left[\mathbf{Q}_M^{(y)\text{T}} \Delta_M^{(x)*} \right] = \sum_x q_{xy} \Upsilon_F^{(x)} \otimes \tilde{\Upsilon}_H^{(x)}.
 \end{aligned} \tag{C.15}$$

This gives a conditional product state iff either $q_{xy} = \delta_{xy} \forall x, y$, which is false by construction; or, either $\Upsilon_F^{(x)}$ or $\tilde{\Upsilon}_H^{(x)}$ (or both) are independent of x . Since the original choice of linearly independent CP maps was arbitrary the proof holds for arbitrary instruments on M . The only remaining way to satisfy Eq. (4.4) is if either the F or H (or both) parts of the process tensor are in tensor product with the part on M .

The proof of Theorem 4.2 is immediate from Lemma C.1, once we consider the fact that the Markov order condition must hold for any block M of length ℓ . Recall that any such memory block can begin and end on either input or output Hilbert spaces separated by ℓ timesteps; for concreteness, here we consider the scenario where M begins and ends on an output Hilbert space, with the proofs for the other cases following the same logic. Consider first the the block M to begin at timestep $t_{k-\ell}$ and end at timestep t_{k-1} . Without loss of generality, suppose that, by Lemma C.1, the process tensor factorises into the product $\Upsilon_{n:1} = \Upsilon_{n:k-\ell^\circ} \otimes \tilde{\Upsilon}_{k-\ell^i:1}$. Had we chosen the block M to begin one timestep later, the same condition leads to the product $\Upsilon_{n:1} = \Upsilon_{n:k-\ell+1^\circ} \otimes \tilde{\Upsilon}_{k-\ell+1^i:1}$. The only way for a single process to satisfy both of these conditions is if there is a CPTP channel $\mathbf{C}_{k-\ell+1^i:k-\ell^\circ}$ taking whatever an experimenter feeds into the process at timestep $t_{k-\ell}$ to the subsequent output from the process at the next timestep $t_{k-\ell+1}$: $\Upsilon_{n:1} = \Upsilon_{n:k-\ell+1^\circ} \otimes \mathbf{C}_{k-\ell+1^i:k-\ell^\circ} \otimes \tilde{\Upsilon}_{k-\ell^i:1}$. Repeating this argument for all timesteps of the process immediately leads to the Markovian (product) process tensor structure of Eq. (3.36).

C.3 CLASSICAL MARKOV ORDER WITH FUZZY MEASUREMENTS

The fact that such coarse-graining can increase the memory length observed by an experimenter arises from the well-known property that the space of Markovian processes is not convex, as exhibited in the following example.

Example C.1 (Fuzzy measurements can increase classical Markov order). Consider the classical process depicted in the left panel of Fig. C.1. At each timestep t_k , the system of interest is described by one of three distinct states $x_k \in \{a_k, b_k, c_k\}$. Between each step of dynamics, the time-invariant transition probabilities are given by $\mathbb{P}_k(b_k|a_{k-1}) = \mathbb{P}_k(c_k|b_{k-1}) = 1$, $\mathbb{P}_k(a_k|c_{k-1}) = p$, $\mathbb{P}_k(b_k|c_{k-1}) = 1 - p$ (with $p \in (0, 1)$) and all

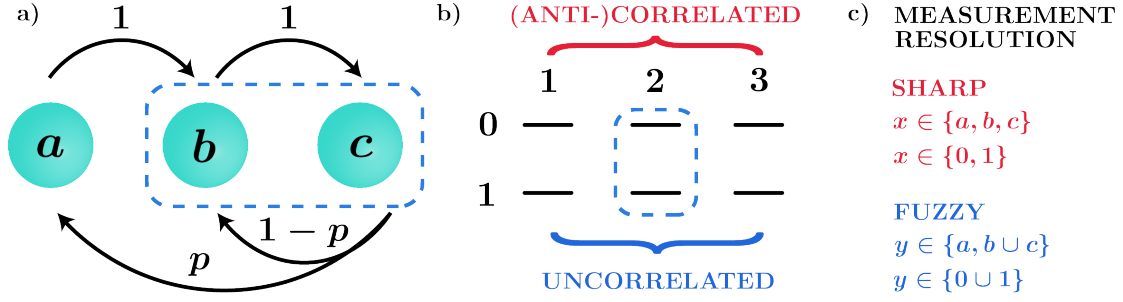


Figure C.1: *Instrument-dependence of classical Markov order with fuzzy measurements.* Here we depict two classical processes to highlight the instrument-dependence of Markov order when sharp measurements are not assumed. In panel **a)**, the process of Example C.1 is shown, defined by the transition probabilities depicted at each timestep. Here, if one is able to record observations sharply, *i.e.*, measure the values $x \in \{a, b, c\}$, the process is Markovian; however, if one cannot measure at that resolution and, *e.g.*, the measurement apparatus only records fuzzy statistics of $y \in \{a, b \cup c\}$, as depicted by the blue dashed box, the process would be classified as non-Markovian. In panel **b)**, the process of Example C.2 is shown. Here, three bits are initially prepared as described in the text, and each bit fed out of the process at successive timesteps. The preparation is such that if the second bit is sharply measured to be in the state 0, bits 1 and 3 are perfectly correlated; if the second bit is in state 1, bits 1 and 3 are perfectly anti-correlated; whilst on average, *i.e.*, with respect to the fuzzy measurement coarse-graining over everything in the dashed blue box, bits 1 and 3 are completely uncorrelated. Note that a legend is provided in the rightmost panel **c)**.

other transitions are forbidden. Such a process is clearly Markovian, as knowledge of any current state suffices to deduce the probability of the next. Now suppose that, an experimenter could not distinguish between outcomes b_k and c_k , *i.e.*, instead of x_j , they observe $y_k \in \{a_k, d_k = b_k \cup c_k\}$. In this case, when the state at some time is a , the next state is for sure d ; while if the state is d , with probability p it will transition next to a or with probability $1 - p$ it will remain d (alternating between b and c deterministically, although the experimenter is ignorant of this fact). Conditioned on any consecutive sequence of j observations of outcome d following an observation of a , we have

$$\mathbb{P}_k(a_k | d_{k-1}, \dots, d_{k-j}, a_{k-j-1}) = \begin{cases} 0 & j \text{ odd} \\ p & j \text{ even,} \end{cases} \quad (\text{C.16})$$

which is different from $\mathbb{P}_k(a_k | d_{k-1}) = p$. Thus, with only the fuzzy measurement apparatus at hand, the experimenter would consider the process to be non-Markovian. Lastly note that given a faulty instrument that alternatively measures x and then y at each consecutive pair of timesteps, the experimenter would determine the Markov order to be $\ell = 2$.

Interestingly, we can also have the opposite scenario occur, *i. e.*, a process can display finite Markov order with respect to a fuzzy measurement sequence, but given access to the system at a finer resolution, the experimenter would attribute a longer memory length to the process, as we now show.

Example C.2 (Fuzzy measurements can decrease classical Markov order). Consider the classical process depicted in the middle panel of Fig. C.1. Here, three bits $x_j = \{0, 1\}$ are output by some process in succession over three timesteps $\{t_1, t_2, t_3\}$. Suppose that these bits are initially prepared according to the probability distribution $\mathbb{P}_{3:1}(x_3, x_2, x_1)$ which is such that $\mathbb{P}_{3:1}(0, 0, 0) = \mathbb{P}_{3:1}(1, 0, 1) = \mathbb{P}_{3:1}(0, 1, 1) = \mathbb{P}_{3:1}(1, 1, 0) = \frac{1}{4}$, and the rest of the possibilities vanish. The process thus constructed is such that if the bit output at the second step is measured to be 0, then the first and third bits are perfectly (classically) correlated; whilst if bit at the second step is measured to be 1, then the first and third bits are perfectly (classically) anti-correlated. Thus, the process is perceived to be non-Markovian with respect to sharp measurements of the second bit value. On the other hand, on average, there is no correlation between the first and third bits; thus, with respect to a coarse-grained measurement that sums over outcomes of the second bit value, the process is perceived to be Markovian.

In either of the above cases, the perceived memory length of the process is instrument-dependent: the first example is a process that is Markovian, but exhibits non-Markovian statistics to the experimenter; whilst the second example is a non-Markovian process that looks Markovian on average, *i. e.*, with respect to the coarse-graining instrument.

C.4 MEMORY LENGTH OF A GENERALISED COLLISION MODEL WITH MEMORY VIA REPEATED SYSTEM-ANCILLA INTERACTIONS

In Section 4.3 we introduced a type of underlying system-environment dynamics that arises from a generalised collision model where the system interacts ℓ times with each ancilla in the order depicted in Fig. 4.3. We claimed that the state of the system subject to such dynamics interspersed with the application of ℓ trash-and-prepare operations can be expressed as a function of only the last ℓ preparations. Here, we explicitly prove this statement.

Consider, without loss of generality, the case for $\ell = 2$ (the extension to larger ℓ is straightforward). Writing out the specific form of the collision model dynamics considered here explicitly, the final output state of the system following two trash-and-prepare

instruments, with the re-preparations of the system state at time t_j represented by σ_j^S , is given by

$$\begin{aligned}\rho_3^S &= \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \text{tr}_{SA_2} \left[\mathcal{U}_{2:1}^{SA_2} \mathcal{U}_{2:1}^{SA_3} \sigma_1^S \text{tr}_{SA_1} \left[\mathcal{U}_{1:0}^{SA_1} \mathcal{U}_{1:0}^{SA_2} \rho_0^S \otimes \tau^E \right] \right] \right] \\ &= \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \dots \text{tr}_{SA_1} \left[\mathcal{U}_{1:0}^{SA_1} \tau^{A_1} \otimes \mathcal{U}_{1:0}^{SA_2} \rho_0^S \otimes \tau^{A_2} \right] \right], \quad (\text{C.17})\end{aligned}$$

where the initial environment is composed of ancillas, $\tau^E = \tau^{A_1} \otimes \tau^{A_2} \otimes \tau^{A_3} \otimes \tau^{A_4}$, and in the second line we pulled these individual ancillas through the maps that do not act upon them.

Now note that we can write the joint SA_2 state after the first interaction, *i. e.*, $\mathcal{U}_{1:0}^{SA_2} \rho_0^S \otimes \tau^{A_2}$, as $\bar{\rho}_0^S(\rho_0^S, \tau^{A_2}) \otimes \bar{\tau}^{A_2}(\rho_0^S, \tau^{A_2})$, where $\bar{\rho}_0^S(\rho_0^S, \tau^{A_2}) := \text{tr}_{A_2} \left[\mathcal{U}_{1:0}^{SA_2} \rho_0^S \otimes \tau^{A_2} \right]$ and similarly for $\bar{\tau}^{A_2}(\rho_0^S, \tau^{A_2})$. This simply expresses the post-interaction marginal states (marked with the overline) as a linear map acting on the pre-interaction states. Importantly, despite the tensor product, this notation does not imply a product state of SA_2 , because of the cross-dependency of the input states; to make this clear, we clearly track this dependency of states through the process with respect to arbitrary unitary interactions, as we are interested in understanding how far into the future their influence can persist. Continuing from above and repeatedly applying this method, we yield

$$\begin{aligned}\rho_3^S &= \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \text{tr}_{SA_2} \left[\mathcal{U}_{2:1}^{SA_2} \bar{\tau}^{A_2}(\rho_0^S, \tau^{A_2}) \otimes \mathcal{U}_{2:1}^{SA_3} \sigma_1^S \otimes \tau^{A_3} \right] \right] \\ &= \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \bar{\tau}^{A_3}(\sigma_1^S, \tau^{A_3}) \otimes \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \right] \\ &= \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \text{tr}_S \left[\mathcal{U}_{2:1}^{SA_3} \sigma_1^S \otimes \tau^{A_3} \right] \otimes \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \right] \\ &= \mathcal{M}(\sigma_1^S, \sigma_2^S).\end{aligned} \quad (\text{C.18})$$

Here, in the penultimate line, we re-expanded $\bar{\tau}^{A_3}(\sigma_1^S, \tau^{A_3})$ to make explicit the fact that ρ_3^S is a function of only the previously 2 prepared states, which can be written as a linear map $\mathcal{M} : \text{BL}(\mathcal{H}_1^S \otimes \mathcal{H}_2^S) \rightarrow \text{BL}(\mathcal{H}_3^S)$ as in the final line, with no dependency on prior historic states such as ρ_0^S . Through time-translational invariance, the proof method holds for arbitrary timesteps and the extension to longer ℓ is immediate. Indeed, the process depicted in Fig. 4.3 has a length- ℓ memory with respect to the trash-and-prepare protocol.

If, on the other hand, one were to apply a different instrument, then the output state here, denoted $\rho_3'^S$, would in general show dependence on the historic state ρ_0^S . Consider for concreteness that an experimenter were to first apply a trash-and-prepare instrument and then at the second timestep a measurement on the system of some outcome m followed by an independent re-preparation of the system into the state σ_2^S . Changing the second operation to a measurement and re-preparation amounts to introducing the

local system measurement operator, $\Pi_2^{(m)}$, into Eq. (C.18) directly after the joint unitary dynamics $\tilde{\mathcal{U}}_{2:1}$ as follows

$$\rho_3'^S = \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \text{tr}_{SA_2} \left[\Pi_2^{(m)} \mathcal{U}_{2:1}^{SA_2} \bar{\tau}^{A_2}(\rho_0^S, \tau^{A_2}) \otimes \mathcal{U}_{2:1}^{SA_3} \sigma_1^S \otimes \tau^{A_3} \right] \right]. \quad (\text{C.19})$$

However, since the system and ancillas A_2 and A_3 , in general, build up correlations during the interactions $\mathcal{U}_{2:1}^{SA_2}$ and $\mathcal{U}_{2:1}^{SA_3}$, the ancillary state of A_3 that feeds forward into the next step of dynamics will be conditioned upon the measurement outcome m , which implicitly depends upon the initial system state ρ_0^S ; indeed, the future dynamics proceeds differently for distinct histories. Explicitly, the furthest we can proceed is to write

$$\rho_3'^S = \text{tr}_{A_4 A_3} \left[\mathcal{U}_{3:2}^{SA_3} \bar{\tau}^{A_3}(m; \rho_0^S, \sigma_1^S, \tau^{A_2}, \tau^{A_3}) \otimes \mathcal{U}_{3:2}^{SA_4} \sigma_2^S \otimes \tau^{A_4} \right], \quad (\text{C.20})$$

where

$$\begin{aligned} & \bar{\tau}^{A_3}(m; \rho_0^S, \sigma_1^S, \tau^{A_2}, \tau^{A_3}) \\ & := \text{tr}_{SA_2} \left[\Pi_2^{(m)} \mathcal{U}_{2:1}^{SA_2} \bar{\tau}^{A_2}(\rho_0^S, \tau^{A_2}) \otimes \mathcal{U}_{2:1}^{SA_3} \sigma_0^S \otimes \tau^{A_3} \right]. \end{aligned} \quad (\text{C.21})$$

Without knowledge of ρ_0^S , the output state $\rho_3'^S$ when this instrument sequence is applied cannot be specified and hence the process displays memory effects that persist longer than ℓ timesteps when a specific measurement, rather than an averaging over such measurements, is recorded.

C.5 OTHER GENERALISED COLLISION MODELS WITH MEMORY

The example introduced in Section 4.3 presents a generalisation of a collision model to include the possibility of memory effects; however, its construction provides by no means the only way to build memory into collision models, which we now briefly explore for the curious reader. A discrete-time, n -step memoryless collision model consists of a system S interacting with an environment E that has a particular structure: it is made up of a number of constituent ancillary subsystems, A_j , with the dynamics proceeding through successive unitary collisions between the system and ancillas (see the top panel in Fig. 4.2). A memoryless collision model assumes the following three key points

1. The system only interacts with each ancilla once.
2. There are no ancilla-ancilla interactions.
3. The ancillas are initially uncorrelated.

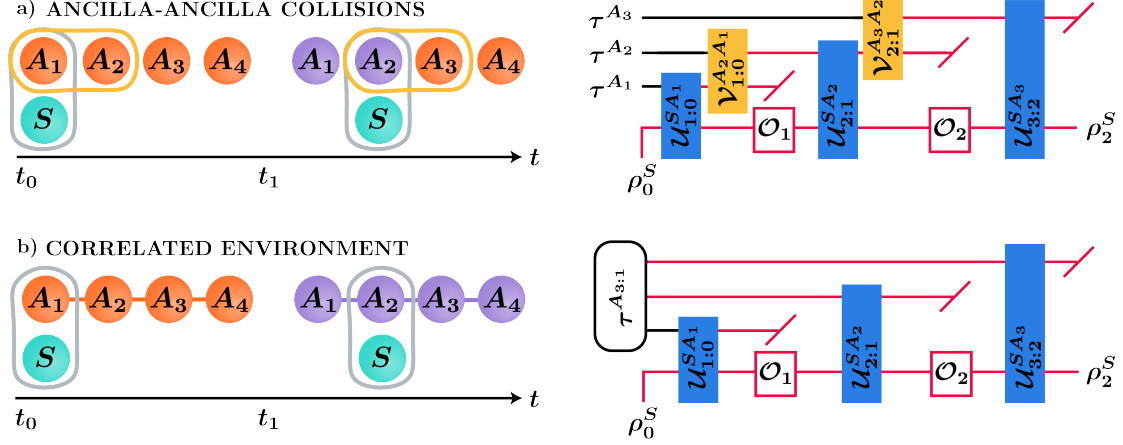


Figure C.2: *Generalised collision models with memory.* Memory can also be built into collision models by allowing for: **a)** ancilla-ancilla interactions (top row) and **b)** an initially correlated environment (bottom row). (The legend is as per Figs. 4.2 and 4.3). The top-left panel depicts a schematic of the dynamics where ancilla-ancilla interactions (yellow boundary) are interleaved between the system-ancilla collisions (grey boundary). After t_1 , the ancilla A_2 already has knowledge of the state of the initial system state mediated via the A_1A_2 interaction, and so the future dynamics is conditioned on the initial system state. The top-right panel displays the corresponding circuit diagram. Here, for arbitrary operations on the system \mathcal{O}_j , it is clear that the ancilla-ancilla interactions provide a possible path of influence from the history to the future state; hence, such a process generically displays infinite Markov order with respect to any instrument sequence (as shown by the red path). The bottom-left panel depicts a schematic of the dynamics where the ancillas constituting the environment begin in a correlated state (represented by the orange line connecting them). As soon as the system interacts with a part of the correlated environment state, all other ancillas can store information about the initial system state, and therefore can influencing the future dynamics. The bottom-right panel displays the corresponding circuit diagram for this case. Again, the initial correlations in the environment provide a mechanism for the history to influence the future over an infinite length of time.

Such a model has surprising power in describing dynamics which, in the continuous-time limit, are governed by a Lindbladian master equation as per Eq. (2.39) [159, 161]. Breaking any one of the above assumptions, whilst maintaining satisfaction of the other two, endows the process with a different type of memory mechanism [173] (see Figs. 4.2, 4.3 and C.2 for illustration). We now examine such memory effects in terms of the structure of the underlying dilation, without any assumptions on the action of the unitaries.

Case 1: Repeated System-Ancilla Interactions. As shown in Section 4.3 and Appendix C.4, in the case where one allows for repeated system-ancilla interactions, as in Refs. [169, 170], the memory effect depends on the nature of these repeated collisions. For example, if they occur in the nested order depicted in Fig. 4.3 then the process has

Markov order ℓ with respect to the trash-and-prepare protocol. If the interactions are simply repeated between the system and a given ancilla multiple times between each timestep, then the process is Markovian on an appropriate coarse-graining of timesteps, *i. e.*, grouping together blocks of ℓ timesteps as one. In general, however, repeated system-ancilla interactions give rise to infinite-length memory (even with respect to the trash-and-prepare protocol). This can be seen by considering the dynamics depicted in Fig. 4.3 with the order of any pair of joint unitary operations flipped: now, a continuous path can be drawn from the history to the future across a length- ℓ trash-and-prepare protocol, indicating a possible historic influence on the future dynamics.

Case 2: Ancilla-Ancilla Interactions. This includes the scenarios considered in Refs. [164–168] and is depicted in the top row of Fig. C.2. In the case where ancilla-ancilla interactions are allowed, the historic influence can, in principle, last forever, since it can permeate continuously through the environment by ancilla-ancilla interactions. Consider specifically the case where at the first step, S is swapped with A_1 through the swap map $\mathcal{U}_{1:0}^{SA_1} = \mathcal{G}^{SA_1}$, then during each successive ancilla-ancilla interaction, the initial system state is continually swapped into the next ancilla via $\mathcal{G}^{A_j A_{j-1}}$, before, finally, A_n , which now stores the initial system state, is swapped back to the system level through $\mathcal{U}_{n:n-1}^{SA_n} = \mathcal{G}^{SA_n}$. Suppose that all but the first and last system-ancilla interactions are identity transformations and we allow for the application of arbitrary probing operations on the system at each timestep in between. These are represented by \mathcal{O}_j , which could, *e. g.*, be trash-and-prepare operations. It is clear that the output system is (trivially) a function of its initial state, regardless of whatever intermediary operations an experimenter applies on the system

$$\begin{aligned} \rho_n^S &= \text{tr}_{A_n} \left(\mathcal{G}^{SA_n} \mathcal{O}_{n-1}^S \text{tr}_{A_{n-1}} \left(\mathcal{G}^{A_n A_{n-1}} \mathcal{O}_{n-2}^S \dots \right. \right. \\ &\quad \left. \left. \dots \text{tr}_{A_2} \left(\mathcal{G}^{A_3 A_2} \mathcal{O}_1^S \left(\mathcal{G}^{A_2 A_1} \mathcal{G}^{SA_1} \rho_0^S \otimes \tau^{A_1} \otimes \dots \otimes \tau^{A_n} \right) \right) \right) \right) \\ &= \text{tr}_{A_n:A_1} \left[\mathcal{G}^{SA_n} \mathcal{G}^{A_n A_{n-1}} \dots \mathcal{G}^{A_2 A_1} \mathcal{O}_{n-1}^S \dots \mathcal{O}_1^S \mathcal{G}^{SA_1} \rho_0^S \otimes \tau^{A_1} \otimes \dots \otimes \tau^{A_n} \right] \\ &= \text{tr}_{A_1} \left[\mathcal{G}^{SA_1} \rho_0^{A_1} \mathcal{O}_{n-1:1}^S \bar{\tau}^S \right] = \rho_0^S. \end{aligned} \tag{C.22}$$

Here, we made use of the composition property of the swap map $\mathcal{G}^{AB} \mathcal{G}^{BC} = \mathcal{G}^{AC}$, compressed the description of the operation sequence as $\mathcal{O}_{n-1:1}^S := \mathcal{O}_{n-1}^S \dots \mathcal{O}_1^S$ and defined $\bar{\tau}^S := \text{tr}_{A_1} \left[\mathcal{G}^{SA_1} \rho_0^S \otimes \tau^{A_1} \right]$ is the initial state of A_1 that is swapped into the system space during the first joint interaction.

Despite the generally infinite-length memory, from the perspective of simulation, this type of memory is not complex: here, given control over part of the environment, one only needs to track one additional ancilla to efficiently simulate such processes, hence

the classification of a *memory depth* of 1 [168], even though the memory length here is infinite. Memory depth is the number of additional ancillas required to embed a non-Markovian process as a Markovian one; in other words, a process with a single ancilla-ancilla interaction between timesteps evolves in a Markovian fashion with respect to treating the system and the ancilla it interacts with at each timestep together as a single larger system of interest. In distinction, memory length concerns the number of timesteps back one needs to store information about the state of the system that could influence future dynamics. The notion of memory depth is indeed interesting for further pursuit in regards to understanding the complexity of the underlying memory at hand.

Case 3: Initially Correlated Environment. Lastly, consider the case of an initially correlated environment, as is studied in Refs. [162, 163] and is depicted in the bottom row of Fig. C.2. Again, there is no generic way to erase the influence of the state’s history on its future evolution by action on the system alone: this is because the ancillary states in the environment begin correlated, and so as soon as the system interacts with the first ancilla, in principle all of the ancillas that *will* interact with the system at some time in the future already store knowledge of the initial system state. Thus, through later interactions, this information can feed-back to dictate the future evolution of the system, giving rise to non-Markovian behaviour.

In the case of an initially correlated environment, one requires control over the entire collection of ancillas to simulate general processes. Again, consider the situation where, A_1 and A_{n-1} begin correlated, and at the first interaction S and A_1 are swapped. Due to the A_1 – A_{n-1} correlation, A_{n-1} also stores knowledge of the initial system state, which can be swapped back to the system level at the final interaction to give the final output. At all intermediate timesteps the dynamics looks like the initial state of A_1 interacting with each other ancilla pairwise in succession. It is clear that, as in Case 2 above, the final state of the system will be identical to its initial state, regardless of the operations one might perform. However, in contrast, simulation of such processes is generically highly complex, as it requires control over a large number of ancillary subsystems in the environment.

D

PROCESSES WITH FINITE MEMORY LENGTH

D.1 PROCESS WITH FINITE MARKOV ORDER AND NON-ZERO QUANTUM CMI

Here we construct the process tensor for a simple dynamics and show that, whilst it has Markov order 1 with respect to a [POVM](#) measurement, the quantum [CMI](#) is non-vanishing.

Example D.1 (Finite Markov order does not imply vanishing quantum conditional mutual information). Consider the case of a three-step process on a qubit, where Alice and Bob have access to the first and second steps respectively, and the final output state is accessible to Charlie (depicted in Fig. [D.1](#)). Initially, the following tripartite state is constructed

$$\rho_{ABC} = \sum_b \frac{1}{4} \rho_A^{(b)} \otimes \Delta_B^{(b)} \otimes \rho_C^{(b)}, \quad (\text{D.1})$$

where, for each value of $b = \{1, 2, 3, 4\}$, $\Delta_B^{(b)} := \frac{1}{2}(\mathbb{1} + \sqrt{3} \sum_i \beta_i^{(b)} \sigma^i)$ is defined in terms of Pauli matrices $\{\sigma_i\}$ with tetrahedral coefficient vectors $\{\beta^{(b)}\} = \{(1, 1, 1), (1, -1, -1), (-1, 1, -1), (-1, -1, 1)\}$. These objects forms the dual set to the following [POVM](#) Π_B , comprising elements $\Pi^{(b)} := \frac{1}{4}(\mathbb{1} + \frac{1}{\sqrt{3}} \sum_i \beta_i^{(b)} \sigma^i)$. We define the states $\rho_X^{(b)} = \frac{3}{8}\mathbb{1} + \frac{1}{2}\Pi^{(b)}$, with $X = \{A, C\}$ in terms of these [POVM](#) elements, before finally normalising the overall tripartite state. The process is such that the A subsystem of the state constructed in Eq. [\(D.1\)](#) is first given to Alice, who can make any operation that she likes. After this, Alice's part is discarded, and the B part of the state above is given to Bob, who can make any operation that he likes. Lastly, Bob's part is discarded, and the C part of the state is given to Charlie. The process tensor is thus given by $\Upsilon_{ABC} = \rho_{ABC}^i \otimes \mathbb{1}_{AB}^o$, where the identity operators on the output spaces of Alice and Bob signify that whatever they feed back into the process is discarded.

Now, suppose Bob chooses to measure the [POVM](#) Π_B as his instrument (potentially with an arbitrary update to the state after the measurement). Then, Eq. [\(4.4\)](#) holds for

Nonetheless, such processes have finite Markov order with respect to the instrument \mathcal{J}_M . This observation directly leads to the proposition of Theorem 5.3.

D.2 FINITE MARKOV ORDER DOES NOT IMPLY VANISHING QUANTUM CMI

Proof of Theorem 5.3. From the structure of Eq. (5.9) it is clear that there exists a history-blocking instrument sequence, namely that comprising the projectors onto each of the m orthogonal subspaces. Begin by rewriting Eq. (5.9) as a regular sum by projecting onto the constituent orthogonal subspaces of the decomposition

$$\begin{aligned}\Upsilon_{FMH}^{\text{CMI}=0} &= \bigoplus_m p_m \tilde{\Upsilon}_{FML}^{(m)} \otimes \tilde{\Upsilon}_{MRH}^{(m)} \\ &= \sum_m p_m \Pi_{ML}^{(m)} \tilde{\Upsilon}_{FML}^{(m)} \Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)} \tilde{\Upsilon}_{MRH}^{(m)} \Pi_{MR}^{(m)}.\end{aligned}\tag{D.2}$$

Consider now the instrument made up of the projectors in the above decomposition, *i.e.*, $\mathcal{J}_M = \{\Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)}\}$. This constitutes a valid instrument sequence as it sums to an identity operator on $\text{BL}(\mathcal{H}_M)$ which is CPTP. It also constitutes a history-blocking sequence for the process described by $\Upsilon_{FMH}^{\text{CMI}=0}$, as for each realisation of the instrument, the future and history are conditionally independent

$$\begin{aligned}\text{tr}_M \left[\left(\Pi_{ML}^{(m')} \otimes \Pi_{MR}^{(m')} \right)^T \Upsilon_{FMH}^{\text{CMI}=0} \right] \\ &= \text{tr}_M \left[\sum_m p_m \Pi_{ML}^{(m)} \tilde{\Upsilon}_{FML}^{(m)} \Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)} \tilde{\Upsilon}_{MRH}^{(m)} \Pi_{MR}^{(m)} \delta_{mm'} \right] \\ &= p_{m'} \text{tr}_M \left[\tilde{\Upsilon}_{FML}^{(m')} \Pi_{ML}^{(m')} \otimes \tilde{\Upsilon}_{MRH}^{(m')} \Pi_{MR}^{(m')} \right] \\ &= p_{m'} \text{tr}_{ML} \left[\tilde{\Upsilon}_{FML}^{(m')} \right] \otimes \text{tr}_{MR} \left[\tilde{\Upsilon}_{MRH}^{(m')} \right] = \Upsilon_F^{(m')} \otimes \tilde{\Upsilon}_H^{(m')},\end{aligned}\tag{D.3}$$

where we use the orthogonal projector identity $\Pi^{(m)} \Pi^{(m')} = \delta_{mm'} \Pi^{(m)}$ and the trace properties of cyclicity and linearity, and in the final line absorbed the probability into $\tilde{\Upsilon}_H^{(m')}$ to yield the exact form of Eq. (4.4).

We now examine the structure of vanishing quantum CMI processes in further detail: this serves to illuminate the connection between processes with finite Markov order with respect to instruments comprising only orthogonal projectors and those with vanishing quantum CMI, which we explore in Section 5.2. Continuing from Eq. (D.2), note that the projectors in the decomposition are not necessarily rank-1; we can thus expand the conditional process tensor parts in terms of a basis within each m subspace as

$$\begin{aligned} \tilde{\Upsilon}_{FML}^{(m)} \otimes \tilde{\Upsilon}_{MRH}^{(m)} &= \Upsilon_F^{(m)} \otimes \Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)} \otimes \tilde{\Upsilon}_H^{(m)} \\ &+ \sum_{ss'} \tilde{\Upsilon}_F^{(m,s)} \otimes \xi_{ML}^{(s)} \otimes \xi_{MR}^{(s')} \otimes \tilde{\Upsilon}_H^{(m,s')}. \end{aligned} \quad (\text{D.4})$$

The $\xi_{ML/R}$ encode the off-diagonal elements within each m subspace (since the projector $\Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)}$ encodes all of the diagonal elements), and can therefore be chosen such that $\text{tr} [\xi_{ML/R}^{(s)}] = 0$ and $\Pi_{ML/R}^{(m)} \xi_{ML/R}^{(s)} = \delta_{ms} \xi_{ML/R}^{(m)} \forall m, s$. In this expansion, neither $\tilde{\Upsilon}_F^{(m,s)}$ nor $\tilde{\Upsilon}_H^{(m,s')}$ are required to be proper process tensors, since the $\xi_{ML/R}$ do not necessarily represent physical operators. We therefore have

$$\begin{aligned} \Upsilon_{FMH}^{\text{CMI}=0} &= \sum_m p_m \Upsilon_F^{(m)} \otimes \Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)} \otimes \tilde{\Upsilon}_H^{(m)} \\ &+ \sum_{m,s,s'} p_m \tilde{\Upsilon}_F^{(m,s)} \otimes \xi_{ML}^{(m,s)} \otimes \xi_{MR}^{(m,s')} \otimes \tilde{\Upsilon}_H^{(m,s')}. \end{aligned} \quad (\text{D.5})$$

Note that if the M subspaces in the decomposition of Eq. (5.9) (see also Eq. (D.2)) are all 1-dimensional, *i. e.*, the projectors in Eq. (D.2) are all rank-1, we only have the first term in the above equation

$$\Upsilon_{FMH}^{\text{CMI}=0} = \sum_m p_m \Upsilon_F^{(m)} \otimes \Pi_{ML}^{(m)} \otimes \Pi_{MR}^{(m)} \otimes \tilde{\Upsilon}_H^{(m)}. \quad (\text{D.6})$$

Regarding the converse statement of Theorem 5.3, we have shown examples of processes with finite Markov order with non-vanishing quantum CMI (Examples 5.1, 5.2 and the generalised collision model of Section 4.3 all display this feature, and for an explicit calculation see Ex. D.1). The analysis above shows that the structural constraint required to guarantee vanishing quantum CMI is strict; processes with finite Markov order must only satisfy the more relaxed structure outlined in Theorem 5.1, and it is therefore insufficient to conclude that such processes have vanishing quantum CMI. Furthermore, even if a given process has finite Markov order with respect to an instrument sequence comprising only rank-1, orthogonal projectors, the process can still have non-vanishing quantum CMI. In this case, since any such set of projectors form a self-dual set, we can reconstruct the process via Eq. (5.2) as

$$\Upsilon_{FMH} = \sum_{x=1}^d \Upsilon_F^{(x)} \otimes \Pi_M^{(x)} \otimes \tilde{\Upsilon}_H^{(x)} + \sum_y \tilde{\Upsilon}_{FMH}^{(y)}, \quad (\text{D.7})$$

with $\text{tr} [\Pi_M^{(x)} \tilde{\Upsilon}_M^{(y)}] = 0 \forall x, y$. Even though the projectors in the history-blocking instrument are not necessarily the same as those that project onto the subspaces defined in the decomposition of Eq. (5.9), this condition does not imply that the process tensor

is block-diagonal in some basis of $\text{BL}(\mathcal{H}_M)$; rather, the process can have off-diagonal elements with respect to the subspaces defined by $\{\Pi_M^{(x)}\}$ and satisfy Eq. (D.7). This implies that there are processes with non-vanishing quantum CMI but finite Markov order. \square

To summarise, the salient points from this analysis are as follows. Firstly, suppose that a process has finite Markov order with respect to an instrument sequence comprising only orthogonal projectors that are not rank-1: in this case, there is no reason that the future-history correlations within each m subspace must obey the product structure outlined in Eq. (D.6), and hence the process can have non-vanishing quantum CMI. This is shown explicitly in Example D.2 of Appendix D.3. However, similar behaviour also arises in an operational interpretation of classical stochastic processes, as discussed in Section 4.2: if an experimenter cannot measure realisations of the process sharply, *i. e.*, with sequences of rank-1 projectors, then the statistics observed do not necessarily have vanishing classical CMI, even if the true underlying process is one of finite Markov order (see the examples of Appendix C.3).

Secondly, suppose that a process has finite Markov order with respect to an instrument sequence comprising only sharp, orthogonal projectors. The condition $\text{tr} \left[\Pi_M^{(x)} \tilde{\Upsilon}_M^{(y)} \right] = 0 \forall x, y$ of Eq. (D.7) does not imply that the process must be block-diagonal in some basis of $\text{BL}(\mathcal{H}_M)$, as is required for the quantum CMI to vanish (see Eq. (D.2)), and it follows that there exist such processes with non-vanishing quantum CMI. In contrast to the earlier point regarding instrument sequences comprising higher-rank projectors, the present statement is indeed a fundamentally quantum mechanical phenomenon. In the classical setting, finite Markov order with respect to sharp realisations of the process and the classical CMI vanishing are equivalent statements (see Subsection 2.1.6).

It is lastly interesting to consider why these two notions are equivalent in the classical setting but not for quantum processes. Suppose that a classical process has finite Markov order with respect to the sequence of sharp projectors $\{\Pi_M^{(m)}\}$; then, the process can be written of the form in Eq. (D.7). However, in the classical setting, where there can be no off-diagonal terms, $\text{tr} \left[\Pi_M^{(x)} \tilde{\Upsilon}_M^{(y)} \right] = 0 \forall x, y$ indeed implies that $\tilde{\Upsilon}_M^{(y)} = 0$. Alternatively, d orthogonal projectors are informationally-complete in the classical setting; thus, the process must be of the form of Corollary 5.2, with the projectors on the M block. In either case, the process is then of the form of Eq. (D.6) (by choosing either \mathcal{H}_{ML} or \mathcal{H}_{MR} to be trivial), meaning the quantum CMI vanishes.

D.3 FUZZY ORTHOGONAL PROJECTIVE MEASUREMENTS ON A QUANTUM PROCESS

For the sake of completeness, here we provide a quantum mechanical analog of Example C.1. As in the classical case, when fuzzy projective measurements are allowed and such a sequence can block the effect of history on the future, the CMI over the statistics observed does not necessarily vanish.

Example D.2 (Process with non-vanishing quantum conditional mutual information but finite Markov order for a sequence of fuzzy, orthogonal projectors). Consider the process depicted in Fig. D.2. Begin with the four two-qubit Werner states defined as

$$\rho_{3^i 1^i}^{(x)}(r) := r\beta^{(x)} + (1-r)\frac{\mathbb{1}}{2}, \quad (\text{D.8})$$

where $r \in (0, 1)$ and each $\beta^{(x)} \in \text{BL}(\mathcal{H}_{3^i} \otimes \mathcal{H}_{1^i})$ is the projector of one of the four Bell pairs

$$|\psi^\pm\rangle := (|00\rangle \pm |11\rangle)/\sqrt{2} \quad \text{and} \quad |\phi^\pm\rangle := (|01\rangle \pm |10\rangle)/\sqrt{2}. \quad (\text{D.9})$$

Now take some symmetric, IC qubit POVM $\{\Pi_{2^i}^{(x)}\}$, such as the tetrahedral measurement defined in Example D.1. In terms of its dual set $\{\Delta_{2^i}^{(x)}\}$, construct the following state

$$\mu_{3^i 2^i 1^i}(r) := \sum_x \frac{1}{4} \rho_{3^i 1^i}^{(x)}(r) \otimes \Delta_{2^i}^{(x)}. \quad (\text{D.10})$$

This object is positive, and therefore a valid quantum state, only for $r \in (0, 1/3]$, which correspond to the values for which the Werner states defined in Eq. (D.8) are separable. Now, suppose that the system associated to \mathcal{H}_{2^i} represents a qutrit: the first two levels are described by Eq. (D.10), the state of which is mixed with probability $q \in (0, 1)$ with an arbitrary tensor product state $\sigma_{3^i} \otimes \sigma_{1^i}$ in product with the third level basis state $|2\rangle$, giving the overall initial system-environment state

$$\rho_{3^i 2^i 1^i}(q, r) = q\mu_{3^i 2^i 1^i}(r) + (1-q)\sigma_{3^i} \otimes |2\rangle\langle 2|_{2^i} \otimes \sigma_{1^i}. \quad (\text{D.11})$$

The process proceeds by initially preparing this state and feeding out the ρ_{j^i} marginal state at each timestep $t_j = \{t_1, t_2, t_3\}$. No matter what operations are implemented on the system at these timesteps, the process acts to discard whatever is fed back into it; therefore, it has trivial output spaces and the corresponding process tensor is

$$\Upsilon_{3^i:1^i}(q, r) = \rho_{3^i 2^i 1^i}(q, r) \otimes \mathbb{1}_{2^0 1^0}. \quad (\text{D.12})$$

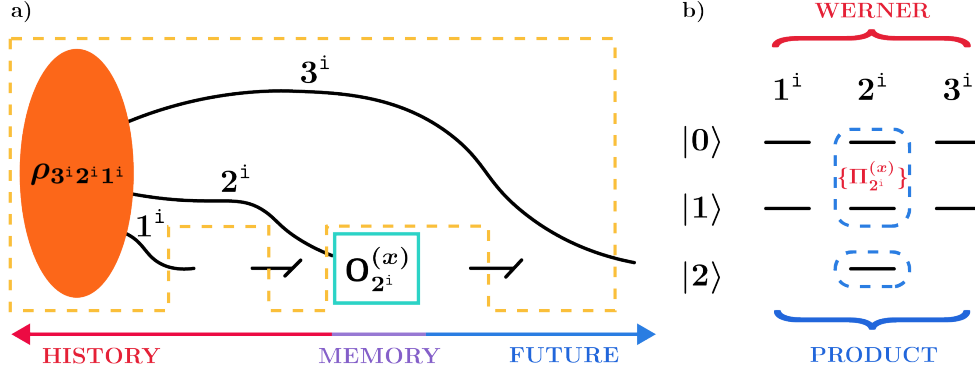


Figure D.2: Process with non-vanishing quantum CMI but finite Markov order with respect to fuzzy, orthogonal projectors. The tripartite state $\rho_{3^i 2^i 1^i}$ as defined in Eq. (D.11) is depicted on the left. Here, if an experimenter cannot distinguish between measurement outcomes in the $\{|0\rangle, |1\rangle\}$ subspace of \mathcal{H}_{2^i} , represented by the dashed, blue boxes on the right, then the conditional state $\rho_{3^i 1^i}^{(x)}$ for each outcome is product. If, on the other hand, the experimenter can resolve sharp measurements in the $\{|0\rangle, |1\rangle\}$ subspace and implement, *e.g.*, the operations $\{O_{2^i}^{(x)}\} = \{\Pi_{2^i}^{(x)}\}$, then for each outcome realised, the conditional state $\rho_{3^i 1^i}^{(x)}$ is a correlated Werner state, defined in Eq. (D.8). The fuzzy orthogonal measurement at timestep 2^i blocks the influence of history on the future, although a sharp measurement resolving all three outcomes does not. Lastly, the quantum CMI for this process does not vanish.

Now, consider the instrument made up of the following two fuzzy, orthogonal operations $O_{2^i}^{(1)} = (\mathbb{1} - |2\rangle\langle 2|)_{2^i}$ and $O_{2^i}^{(2)} = |2\rangle\langle 2|_{2^i}$. With respect to this instrument, the conditional process tensors for each outcome are

$$\Upsilon_{3^i 2^i 1^i}^{(1)} = \frac{\mathbb{1}_{3^i}}{2} \otimes \mathbb{1}_{2^i 1^i} \otimes \frac{\mathbb{1}_{1^i}}{2} \quad \text{and} \quad \Upsilon_{3^i 1^i}^{(2)} = \sigma_{3^i} \otimes \mathbb{1}_{2^i 1^i} \otimes \sigma_{1^i}. \quad (\text{D.13})$$

Thus, Eq. (4.4) is satisfied and the process has Markov order 1 with respect to this instrument comprising only fuzzy orthogonal projectors. (Note that this process is not Markovian, as an IC instrument of causal breaks does not block the history.) However, had the experimenter been able to resolve measurements in the $\{|0\rangle, |1\rangle\}$ subspace of \mathcal{H}_{2^i} , *e.g.*, apply the instrument comprising the operations $O_{2^i}^{(x)} = \Pi_{2^i}^{(x)}$ for $x \in \{1, 2, 3, 4\}$ and $O_{2^i}^{(5)} = |2\rangle\langle 2|_{2^i}$, then the conditional process tensors for each outcome are

$$\Upsilon_{3^i 1^i}^{(x)} = \psi_{3^i 1^i}^{(x)} \otimes \mathbb{1}_{2^i 1^i} \quad \text{and} \quad \Upsilon_{3^i 1^i}^{(5)} = \sigma_{3^i} \otimes \mathbb{1}_{2^i 1^i} \otimes \sigma_{1^i}. \quad (\text{D.14})$$

For each outcome x observed in the $\{|0\rangle, |1\rangle\}$ subspace, the conditional future and history processes exhibit correlations via one of the four Werner states, which are separable, but not product, and therefore correlated. Similarly, if the experimenter applied the sharp projectors that make up the fuzzy history-blocking instrument, *i.e.*, measure the three outcomes associated to $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |2\rangle\langle 2|\}$ at t_2 , the conditional states for outcomes

0 and 1 are again correlated. Lastly, note that this process has non-vanishing quantum [CMI](#): a straightforward calculation shows that $I(F : H|M) = q$ for $\Upsilon_{3^i;1^i}(q, r)$ defined in Eq. [\(D.12\)](#).

BIBLIOGRAPHY

- [1] P. TARANTO, F. A. POLLOCK, S. MILZ, M. TOMAMICHEL, and K. MODI, *Phys. Rev. Lett.* **122**, 140401 (2019).
- [2] P. TARANTO, S. MILZ, F. A. POLLOCK, and K. MODI, *Phys. Rev. A* **99**, 042108 (2019).
- [3] C. GARDINER, *Stochastic Methods (4th Ed.)*, Springer-Verlag, Berlin Heidelberg (2009).
- [4] N. VAN KAMPEN, *Stochastic Processes in Physics and Chemistry*, Elsevier, New York (2011).
- [5] H. WIO, R. DEZA, and J. LÓPEZ, *An Introduction to Stochastic Processes and Nonequilibrium Statistical Physics (Rev. Ed.)*, World Scientific, Singapore (2012).
- [6] E. B. DAVIES, *Quantum Theory of Open Systems*, Academic Press, London (1976).
- [7] H.-P. BREUER and F. PETRUCCIONE, *The Theory of Open Quantum Systems*, Oxford University Press, New York (2002).
- [8] M. NIELSEN and I. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).
- [9] P. ZANARDI and M. RASETTI, *Phys. Rev. Lett.* **79**, 3306 (1997).
- [10] E. KNILL, R. LAFLAMME, and W. H. ZUREK, *Science* **279**, 342 (1998).
- [11] L. VIOLA, E. KNILL, and S. LLOYD, *Phys. Rev. Lett.* **82**, 2417 (1999).
- [12] A. J. LEGGETT and A. GARG, *Phys. Rev. Lett.* **54**, 857 (1985).
- [13] A. J. LEGGETT, *Rep. Prog. Phys.* **71**, 022001 (2008).
- [14] S. MILZ, F. SAKULDEE, F. A. POLLOCK, and K. MODI, *arXiv:1712.02589* (2017).
- [15] A. N. KOLMOGOROV, *Foundations of the Theory of Probability*, Chelsea, New York (1956).
- [16] J. PEARL, *Causality*, Oxford University Press, New York (2000).
- [17] F. COSTA and S. SHRAPNEL, *New J. Phys.* **18**, 063032 (2016).

BIBLIOGRAPHY

- [18] J.-M. A. ALLEN, J. BARRETT, D. C. HORSMAN, C. M. LEE, and R. W. SPEKKENS, *Phys. Rev. X* **7**, 031021 (2017).
- [19] G. MILBURN and S. SHRAPNEL, *Entropy* **20**, 687 (2018).
- [20] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Wiley, New York (1971).
- [21] E. T. JAYNES, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge (2003).
- [22] Á. RIVAS, S. F. HUELGA, and M. B. PLENIO, *Rep. Prog. Phys.* **77**, 094001 (2014).
- [23] H.-P. BREUER, E.-M. LAINE, J. PIILO, and B. VACCHINI, *Rev. Mod. Phys.* **88**, 021002 (2016).
- [24] I. DE VEGA and D. ALONSO, *Rev. Mod. Phys.* **89**, 015001 (2017).
- [25] N. VAN KAMPEN, *Braz. J. Phys.* **28**, 90 (1998).
- [26] G. LINDBLAD, *Commun. Math. Phys.* **65**, 281 (1979).
- [27] L. ACCARDI, A. FRIGERIO, and J. T. LEWIS, *Publ. Rest. Inst. Math. Sci.* **18**, 97 (1982).
- [28] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *EPL* **83**, 30004 (2008).
- [29] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [30] G. CHIRIBELLA, G. M. D’ARIANO, and P. PERINOTTI, *Phys. Rev. A* **80**, 022339 (2009).
- [31] G. CHIRIBELLA, G. M. D’ARIANO, P. PERINOTTI, and B. VALIRON, *Phys. Rev. A* **88**, 022318 (2013).
- [32] O. ORESHKOV, F. COSTA, and Č. BRUKNER, *Nat. Commun.* **3**, 1092 (2012).
- [33] O. ORESHKOV and C. GIARMATZI, *New J. Phys.* **18**, 093020 (2016).
- [34] S. SHRAPNEL, F. COSTA, and G. MILBURN, *arXiv:1702.01845* (2017).
- [35] S. MILZ, F. A. POLLOCK, T. P. LE, G. CHIRIBELLA, and K. MODI, *New J. Phys.* **20**, 033033 (2018).
- [36] S. SHRAPNEL, F. COSTA, and G. MILBURN, *Int. J. Quantum Inf.* **16**, 1840010 (2018).

- [37] F. COSTA, M. RINGBAUER, M. E. GOGGIN, A. G. WHITE, and A. FEDRIZZI, *Phys. Rev. A* **98**, 012328 (2018).
- [38] M. RINGBAUER and R. CHAVES, *Quantum* **1**, 35 (2017).
- [39] L. HARDY, *Philos. Trans. Royal Soc. A* **370**, 3385 (2012).
- [40] L. HARDY, [arXiv:1608.06940](#) (2016).
- [41] J. COTLER, C.-M. JIAN, X.-L. QI, and F. WILCZEK, *J. High Energy Phys.* **2018**, 93 (2018).
- [42] G. GUTOSKI and J. WATROUS, “Toward a General Theory of Quantum Games”. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, San Diego, USA (2007).
- [43] G. GUTOSKI, *J. Math. Phys.* **53**, 032202 (2012).
- [44] C. PORTMANN, C. MATT, U. MAUERER, R. RENNER, and B. TACKMANN, *IEEE Trans. Inf. Theory* **65**, 3277 (2017).
- [45] P. FIGUEROA-ROMERO, K. MODI, and F. A. POLLOCK, *Quantum* **3**, 136 (2018).
- [46] P. STRASBERG, [arXiv:1810.00698](#) (2018).
- [47] D. KRETSCHMANN and R. F. WERNER, *Phys. Rev. A* **72**, 062323 (2005).
- [48] F. CARUSO, V. GIOVANNETTI, C. LUPO, and S. MANCINI, *Rev. Mod. Phys.* **86**, 1203 (2014).
- [49] F. A. POLLOCK, C. RODRÍGUEZ-ROSARIO, T. FRAUENHEIM, M. PATERNOSTRO, and K. MODI, *Phys. Rev. Lett.* **120**, 040405 (2018).
- [50] F. A. POLLOCK, C. RODRÍGUEZ-ROSARIO, T. FRAUENHEIM, M. PATERNOSTRO, and K. MODI, *Phys. Rev. A* **97**, 012127 (2018).
- [51] S. MILZ, F. A. POLLOCK, and K. MODI, *Open Sys. Info. Dyn.* **24**, 1740016 (2017).
- [52] F. SAKULDEE, S. MILZ, F. A. POLLOCK, and K. MODI, *J. Phys. A* **51**, 414014 (2018).
- [53] C. GIARMATZI and F. COSTA, [arXiv:1811.03722](#) (2018).
- [54] S. MILZ, M. S. KIM, F. A. POLLOCK, and K. MODI, [arXiv:1901.05223](#) (2019).
- [55] S. MILZ, “On the Operational Theory of General (Quantum) Stochastic Processes,” PhD thesis, Monash University, Melbourne (2019).
- [56] D. P. S. MCCUTCHEON, *Phys. Rev. A* **93**, 022119 (2016).

- [57] T. IKEDA, H. ITO, and Y. TANIMURA, *J. Chem. Phys.* **142**, 212421 (2015).
- [58] S. L. SALZBERG, A. L. DELCHER, S. KASIF, and O. WHITE, *Nucleic Acids Res.* **26**, 544 (1998).
- [59] G. THIJS, M. LESCOT, K. MARCHAL, S. ROMBAUTS, B. DE MOOR, P. ROUZE, and Y. MOREAU, *Bioinformatics* **17**, 1113 (2001).
- [60] M. ROSVALL, A. V. ESQUIVEL, A. LANCICHINETTI, J. D. WEST, and R. LAMBIOTTE, *Nat. Commun.* **5**, 4630 (2014).
- [61] K. BANASZEK, A. DRAGAN, W. WASILEWSKI, and C. RADZEWICZ, *Phys. Rev. Lett.* **92**, 257901 (2004).
- [62] N. EREZ, G. GORDON, M. NEST, and G. KURIZKI, *Nature* **452**, 724 (2008).
- [63] D. M. REICH, N. KATZ, and C. P. KOCH, *Sci. Rep.* **5**, 12430 (2015).
- [64] J. MORRIS, F. A. POLLOCK, and K. MODI, *arXiv:1902.07980* (2019).
- [65] K. MODI and E. C. G. SUDARSHAN, *Phys. Rev. A* **81**, 052119 (2010).
- [66] Á. RIVAS, S. F. HUELGA, and M. B. PLENIO, *Phys. Rev. Lett.* **105**, 050403 (2010).
- [67] B. VACCHINI, A. SMIRNE, E.-M. LAINE, J. PILO, and H.-P. BREUER, *New J. Phys.* **13**, 093004 (2011).
- [68] R. ALICKI, *arXiv:0205188* (2002).
- [69] P. HÄNGGI and H. THOMAS, *Z. Phys. B* **26**, 85 (1977).
- [70] H. GRABERT, P. TALKNER, and P. HÄNGGI, *Z. Phys. B* **26**, 389 (1977).
- [71] P. HÄNGGI, H. THOMAS, H. GRABERT, and P. TALKNER, *J. Stat. Phys.* **18**, 155 (1978).
- [72] H. GRABERT, P. TALKNER, P. HÄNGGI, and H. THOMAS, *Z. Phys. B* **29**, 273 (1978).
- [73] K. MODI, *Open Sys. Info. Dyn.* **18**, 253 (2011).
- [74] K. MODI, *Sci. Rep.* **2**, 581 (2012).
- [75] K. MODI, C. RODRÍGUEZ-ROSARIO, and A. ASPURU-GUZI, *Phys. Rev. A* **86**, 064102 (2012).
- [76] L. LI, M. J. HALL, and H. M. WISEMAN, *Phys. Rep.* **759**, 1 (2018).
- [77] V. GIOVANNETTI and G. M. PALMA, *Phys. Rev. Lett.* **108**, 040401 (2012).

- [78] S. LORENZO, F. CICCARELLO, and G. M. PALMA, *Phys. Rev. A* **96**, 032107 (2017).
- [79] S. LORENZO, F. CICCARELLO, G. M. PALMA, and B. VACCHINI, *Open Sys. Info. Dyn.* **24**, 1740011 (2017).
- [80] K. KRAUS, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Springer, Berlin (1983).
- [81] R. ALICKI and K. LENDI, *Quantum Dynamical Semi-groups and Applications*, Springer, Berlin (1987).
- [82] H. M. WISEMAN and G. J. MILBURN, *Quantum Measurement and Control*, Cambridge University Press, Cambridge (2010).
- [83] M. PARIS and J. REHACEK, *Quantum State Estimation*, Springer, Berlin (2004).
- [84] M. HORODECKI, P. HORODECKI, and R. HORODECKI, *Phys. Lett. A* **223**, 1 (1996).
- [85] R. HORODECKI, P. HORODECKI, M. HORODECKI, and K. HORODECKI, *Rev. Mod. Phys.* **81**, 865 (2009).
- [86] W. F. STINESPRING, *Proc. Am. Math. Soc.* **6**, 211 (1955).
- [87] V. GORINI, A. KOSSAKOWSKI, and E. C. G. SUDARSHAN, *J. Math. Phys.* **17**, 821 (1976).
- [88] G. LINDBLAD, *Commun. Math. Phys.* **48**, 119 (1976).
- [89] V. A. FRANKE, *Theor. Math. Phys.* **27**, 406 (1976).
- [90] K. MODI, “A Theoretical Analysis of Experimental Open Quantum Dynamics,” PhD thesis, University of Texas, Austin (2009), [arXiv:0903.2724](https://arxiv.org/abs/0903.2724).
- [91] S. C. HOU, X. X. YI, S. X. YU, and C. H. OH, *Phys. Rev. A* **83**, 062115 (2011).
- [92] H.-P. BREUER, E.-M. LAINE, and J. PILO, *Phys. Rev. Lett.* **103**, 210401 (2009).
- [93] C. A. RODRÍGUEZ-ROSARIO, K. MODI, A.-M. KUAH, A. SHAJI, and E. C. G. SUDARSHAN, *J. Phys. A* **41**, 205301 (2008).
- [94] E.-M. LAINE, J. PILO, and H.-P. BREUER, *EPL* **92**, 60010 (2010).
- [95] M. GESSNER and H.-P. BREUER, *Phys. Rev. Lett.* **107**, 180402 (2011).
- [96] L. MAZZOLA, C. A. RODRÍGUEZ-ROSARIO, K. MODI, and M. PATERNOSTRO, *Phys. Rev. A* **86**, 010102 (2012).

- [97] C. A. RODRÍGUEZ-ROSARIO, K. MODI, L. MAZZOLA, and A. ASPURU-GUZI, *EPL* **99**, 20010 (2012).
- [98] M. GESSNER, M. RAMM, T. PRUTTIVARASIN, A. BUCHLEITNER, H.-P. BREUER, and H. HÄFFNER, *Nature Physics* **10**, 105 (2013).
- [99] M. M. WOLF, J. EISERT, T. S. CUBITT, and J. I. CIRAC, *Phys. Rev. Lett.* **101**, 150402 (2008).
- [100] A. R. USHA DEVI, A. K. RAJAGOPAL, and SUDHA, *Phys. Rev. A* **83**, 022109 (2011).
- [101] D. CHRUSCIŃSKI and S. MANISCALCO, *Phys. Rev. Lett.* **112**, 120404 (2014).
- [102] S. LUO, S. FU, and H. SONG, *Phys. Rev. A* **86**, 044101 (2012).
- [103] Z. HE, H.-S. ZENG, Y. LI, Q. WANG, and C. YAO, *Phys. Rev. A* **96**, 022106 (2017).
- [104] X.-M. LU, X. WANG, and C. P. SUN, *Phys. Rev. A* **82**, 042103 (2010).
- [105] F. F. FANCHINI, G. KARPAT, B. ÇAKMAK, L. K. CASTELANO, G. H. AGUILAR, O. J. FARÍAS, S. P. WALBORN, P. H. S. RIBEIRO, and M. C. DE OLIVEIRA, *Phys. Rev. Lett.* **112**, 210402 (2014).
- [106] B. BYLICKA, D. CHRUSCIŃSKI, and S. MANISCALCO, *Sci. Rep.* **4**, 5720 (2014).
- [107] C. PINEDA, T. GORIN, D. DAVALOS, D. A. WISNIACKI, and I. GARCÍA-MATA, *Phys. Rev. A* **93**, 022117 (2016).
- [108] B. BYLICKA, M. JOHANSSON, and A. ACIN, *arXiv:1603.04288* (2016).
- [109] P. HAIKKA, J. D. CRESSER, and S. MANISCALCO, *Phys. Rev. A* **83**, 012112 (2011).
- [110] D. CHRUSCIŃSKI, A. KOSSAKOWSKI, and Á. RIVAS, *Phys. Rev. A* **83**, 052128 (2011).
- [111] P. PECHUKAS, *Phys. Rev. Lett.* **73**, 1060 (1994).
- [112] R. ALICKI, *Phys. Rev. Lett.* **75**, 3020 (1995).
- [113] P. PECHUKAS, *Phys. Rev. Lett.* **75**, 3021 (1995).
- [114] C. RODRÍGUEZ-ROSARIO, K. MODI, and A. ASPURU-GUZI, *Phys. Rev. A* **81**, 012313 (2010).
- [115] M. A. NIELSEN, E. KNILL, and R. LAFLAMME, *Nature* **396**, 52 (1998).

- [116] A. M. CHILDS, I. L. CHUANG, and D. W. LEUNG, *Phys. Rev. A* **64**, 012314 (2001).
- [117] M. W. MITCHELL, C. W. ELLENOR, S. SCHNEIDER, and A. M. STEINBERG, *Phys. Rev. Lett.* **91**, 120402 (2003).
- [118] Y. S. WEINSTEIN, T. F. HAVEL, J. EMERSON, N. BOULANT, M. SARACENO, S. LLOYD, and D. G. CORY, *J. Chem. Phys.* **121**, 6117 (2004).
- [119] J. L. O'BRIEN, G. J. PRYDE, A. GILCHRIST, D. F. V. JAMES, N. K. LANGFORD, T. C. RALPH, and A. G. WHITE, *Phys. Rev. Lett.* **93**, 080502 (2004).
- [120] F. MASILLO, G. SCOLARICI, and L. SOLOMBRINO, *J. Math. Phys.* **52**, 012101 (2011).
- [121] F. BUSCEMI, *Phys. Rev. Lett.* **113**, 140502 (2014).
- [122] G. ARGENTIERI, F. BENATTI, R. FLOREANINI, and M. PEZZUTTO, *J. Stat. Phys.* **159**, 1127 (2015).
- [123] S. MILZ, F. A. POLLOCK, and K. MODI, *Phys. Rev. A* **98**, 012108 (2018).
- [124] D. S. ABRAMS and S. LLOYD, *Phys. Rev. Lett.* **81**, 3992 (1998).
- [125] A. PERES, *Phys. Rev. Lett.* **63**, 1114 (1989).
- [126] N. GISIN, *Phys. Lett. A* **143**, 1 (1990).
- [127] A. PERES, *Found. Phys.* **33**, 1543 (2003).
- [128] F. HAAKE and R. REIBOLD, *Phys. Rev. A* **32**, 2462 (1985).
- [129] D. SCHMID, K. RIED, and R. W. SPEKKENS, *arXiv:1806.02381* (2018).
- [130] M. RINGBAUER, C. J. WOOD, K. MODI, A. GILCHRIST, A. G. WHITE, and A. FEDRIZZI, *Phys. Rev. Lett.* **114**, 090402 (2015).
- [131] E. C. G. SUDARSHAN, P. M. MATHEWS, and J. RAU, *Phys. Rev.* **121**, 920 (1961).
- [132] K. KRAUS, *Ann. Phys.* **64**, 311 (1971).
- [133] A. JAMIOŁKOWSKI, *Rep. Math. Phys.* **3**, 275 (1972).
- [134] M.-D. CHOI, *Linear Algebra Its Appl.* **10**, 285 (1975).
- [135] D. GENKINA, G. CHIRIBELLA, and L. HARDY, *Phys. Rev. A* **85**, 022330 (2012).
- [136] D. PETZ, *Rev. Math. Phys.* **15**, 79 (2003).
- [137] P. HAYDEN, R. JOZSA, D. PETZ, and A. WINTER, *Commun. Math. Phys.* **246**, 359 (2004).

- [138] B. IBINSON, N. LINDEN, and A. WINTER, *Commun. Math. Phys.* **277**, 289 (2008).
- [139] O. FAWZI and R. RENNER, *Commun. Math. Phys.* **340**, 575 (2015).
- [140] M. M. WILDE, *Proc. Royal Soc. A* **471**, 2182 (2015).
- [141] D. SUTTER, O. FAWZI, and R. RENNER, *Proc. Royal Soc. A* **472**, 20150623 (2016).
- [142] D. SUTTER, M. BERTA, and M. TOMAMICHEL, *Commun. Math. Phys.* **352**, 37 (2017).
- [143] M. B. RUSKAI, *J. Math. Phys.* **43**, 4358 (2002).
- [144] D. PETZ, *Comm. Math. Phys.* **105**, 123 (1986).
- [145] M. M. WOLF, F. VERSTRAETE, M. B. HASTINGS, and J. I. CIRAC, *Phys. Rev. Lett.* **100**, 070502 (2008).
- [146] S. SIEGERT, R. FRIEDRICH, and J. PEINKE, *Phys. Lett. A* **243**, 275 (1998).
- [147] J. P. M. HEALD and J. STARK, *Phys. Rev. Lett.* **84**, 2366 (2000).
- [148] M. SIEFERT, A. KITTEL, R. FRIEDRICH, and J. PEINKE, *EPL* **61**, 466 (2003).
- [149] M. WAECHTER, F. RIESS, T. SCHIMMEL, U. WENDT, and J. PEINKE, *Eur. Phys. J. B* **41**, 259 (2004).
- [150] F. BÖTTCHER, J. PEINKE, D. KLEINHANS, R. FRIEDRICH, P. G. LIND, and M. HAASE, *Phys. Rev. Lett.* **97**, 090603 (2006).
- [151] D. KLEINHANS, R. FRIEDRICH, M. WÄCHTER, and J. PEINKE, *Phys. Rev. E* **76**, 041109 (2007).
- [152] B. LEHLE, *Phys. Rev. E* **83**, 021113 (2011).
- [153] J. G. KEMENY and J. L. SNELL, *Finite Markov Chains*, Springer-Verlag, New York (1976).
- [154] E. J. KOSTELICH and T. SCHREIBER, *Phys. Rev. E* **48**, 1752 (1993).
- [155] H. KANTZ and T. SCHREIBER, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge (1997).
- [156] S. GUDDER, *Adv. Appl. Math.* **20**, 169 (1998).
- [157] J. RAU, *Phys. Rev.* **129**, 1880 (1963).
- [158] M. ZIMAN, P. ŠTELMACHOVIČ, V. BUŽEK, M. HILLERY, V. SCARANI, and N. GISIN, *Phys. Rev. A* **65**, 042105 (2002).

- [159] T. A. BRUN, *Am. J. Phys.* **70**, 719 (2002).
- [160] V. SCARANI, M. ZIMAN, P. ŠTELMACHOVIČ, N. GISIN, and V. BUŽEK, *Phys. Rev. Lett.* **88**, 097905 (2002).
- [161] M. ZIMAN, P. ŠTELMACHOVIČ, and V. BUŽEK, *Open Sys. Info. Dyn.* **12**, 81 (2005).
- [162] T. RYBÁR, S. N. FILIPPOV, M. ZIMAN, and V. BUŽEK, *J. Phys. B* **45**, 154006 (2012).
- [163] N. K. BERNARDES, A. R. R. CARVALHO, C. H. MONKEN, and M. F. SANTOS, *Phys. Rev. A* **90**, 032111 (2014).
- [164] F. CICCARELLO, G. M. PALMA, and V. GIOVANNETTI, *Phys. Rev. A* **87**, 040103 (2013).
- [165] R. MCCLOSKEY and M. PATERNOSTRO, *Phys. Rev. A* **89**, 052120 (2014).
- [166] S. LORENZO, F. CICCARELLO, and G. M. PALMA, *Phys. Rev. A* **93**, 052111 (2016).
- [167] B. ÇAKMAK, M. PEZZUTTO, M. PATERNOSTRO, and Ö. E. MÜSTECAPLIOĞLU, *Phys. Rev. A* **96**, 022109 (2017).
- [168] S. CAMPBELL, F. CICCARELLO, G. M. PALMA, and B. VACCHINI, *Phys. Rev. A* **98**, 012142 (2018).
- [169] A. L. GRIMSMO, *Phys. Rev. Lett.* **115**, 060402 (2015).
- [170] S. J. WHALEN, A. L. GRIMSMO, and H. J. CARMICHAEL, *Quantum Sci. Technol.* **2**, 044008 (2017).
- [171] S. KRETSCHMER, K. LUOMA, and W. T. STRUNZ, *Phys. Rev. A* **94**, 012106 (2016).
- [172] A. STRATHEARN, P. KIRTON, D. KILDA, J. KEELING, and B. W. LOVETT, *Nat. Commun.* **9**, 3322 (2018).
- [173] F. CICCARELLO, *Quantum Meas. Quantum Metrol.* **4**, 53 (2017).
- [174] K. WIESNER and J. P. CRUTCHFIELD, *arXiv:0611143* (2006).
- [175] D. E. MAKAROV and N. MAKRI, *Chem. Phys. Lett.* **221**, 482 (1994).
- [176] N. MAKRI and D. E. MAKAROV, *J. Chem. Phys.* **102**, 4600 (1995).
- [177] N. MAKRI and D. E. MAKAROV, *J. Chem. Phys.* **102**, 4611 (1995).
- [178] F. A. POLLOCK and K. MODI, *Quantum* **2**, 76 (2018).

- [179] S. PANG, T. A. BRUN, and A. N. JORDAN, [arXiv:1712.10109](#) (2017).
- [180] V. VEDRAL, [Rev. Mod. Phys.](#) **74**, 197 (2002).
- [181] J. CERRILLO and J. CAO, [Phys. Rev. Lett.](#) **112**, 110401 (2014).
- [182] R. ROSENBACH, J. CERRILLO, S. F. HUELGA, J. CAO, and M. B. PLENIO, [New J. Phys.](#) **18**, 023035 (2016).
- [183] A. A. KANANENKA, C.-Y. HSIEH, J. CAO, and E. GEVA, [J. Phys. Chem. Lett.](#) **7**, 4809 (2016).
- [184] M. R. JØRGENSEN and F. A. POLLOCK, [arXiv:1902.00315](#) (2019).
- [185] J. P. CLEMENS, S. SIDDIQUI, and J. GEA-BANACLOCHE, [Phys. Rev. A](#) **69**, 062313 (2004).
- [186] G. CHIRIBELLA, M. DALL’ARNO, G. M. D’ARIANO, C. MACCHIAVELLO, and P. PERINOTTI, [Phys. Rev. A](#) **83**, 052305 (2011).
- [187] A. BEN-AROYA and A. TA-SHMA, [IEEE Trans. Inf. Theory](#) **57**, 3982 (2011).
- [188] C. LUPO, L. MEMARZADEH, and S. MANCINI, [Phys. Rev. A](#) **85**, 012320 (2012).
- [189] C. L. EDMUNDS, C. HEMPEL, R. HARRIS, H. BALL, V. FREY, T. M. STACE, and M. J. BIERCUK, [arXiv:1712.04954](#) (2017).

DECLARATION

I hereby declare that 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 that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis. This thesis includes ideas presented in 2 original articles that have been accepted for publication, namely Ref. [1]: “Quantum Markov Order”, which is to be published in Physical Review Letters and Ref. [2]: “The Structure of Quantum Stochastic Processes with Finite Markov Order”, which is to be published in Physical Review A. The core theme of the thesis is Memory Effects in Quantum Stochastic Processes. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Monash Quantum Information Science group under the supervision of Dr. Kavan Modi and Dr. Felix A. Pollock. (The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.) I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Melbourne, Australia, May 6, 2019.

Philip Taranto

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student’s and co-authors’ contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Melbourne, Australia, May 6, 2019.

Dr. Kavan Modi

Dr. Felix A. Pollock

COLOPHON

Philip Taranto: *Memory Effects in Quantum Processes*, © May 6, 2019.