A second-quantised Shannon theory

Hlér Kristjánsson

St Anne's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2022

Contents

Abstract

Shannon's theory of information laid the groundwork for the rapid developments in information and communications technologies over the last century. Yet, it assumed that information carriers were described by the laws of classical physics, whilst at the most fundamental level, nature obeys the laws of quantum physics. *Quantum Shannon theory*, which describes information carriers as quantum states, has led to a new era of possibilities, such as perfectly secure cryptography without pre-established keys. Yet, there is a sense in which this transition from classical to quantum is incomplete. Traditionally, quantum Shannon theory has focused on scenarios where the internal states of information carriers are quantum, whilst their trajectories in spacetime have still been assumed to be classical.

This work presents a second level of quantisation where both the information itself and its propagation in spacetime are treated in a quantum fashion. The second-quantised Shannon theory describes the possibility of a single particle being simultaneously transmitted through multiple communication channels in a quantum superposition. The framework is developed using the tools of *higher-order transformations* and *quantum resource theories*, formally quantifying the resources required for communication between a sender and receiver in this setting.

The advantages of the second-quantised theory are illustrated in a series of examples, showcasing various counterintuitive phenomena that occur when information is simultaneously transmitted through multiple communication channels. In particular, when a single particle travels in a quantum superposition through two alternative transmission lines, the noisy processes in the two lines can destructively interfere, leading to a cleaner communication channel overall. Various different scenarios are encompassed in the framework, including transmission through a superposition of both independent and correlated channels, as well as through large-scale communication networks. This work concludes with a study of the robustness of these protocols to errors and a discussion of recent experimental demonstrations of their associated communication advantages.

Acknowledgements

First of all, I would like to thank my supervisor Prof Giulio Chiribella, with whom the majority of the work presented in this thesis was completed, for guiding me in my research over the last four years. I have learnt so much from Giulio, from mathematical techniques and paper writing, to understanding how best to develop novel research ideas. In particular, I am extremely grateful to Giulio for putting in a great amount of time into working closely together with me during the first two years of my DPhil, making sure that I learned as much as possible, whilst also giving me increasing freedom to pursue ideas of my own in later years, enabling me to develop as an independent researcher. Moreover, I would like to thank Giulio for inviting me to spend several months in his group at the Department of Computer Science, the University of Hong Kong, in 2019 and early 2020, as part of the HKU-Oxford Joint Laboratory for Quantum Information and Computation.

Second, I would like to thank my supervisor Prof Jonathan Barrett for adopting me into his group in Oxford before I was officially his student, and for later offering his guidance on two major research projects during the latter two years of my DPhil. Jon has given me insights into some of the most fundamental problems in quantum foundations, and helped me understand what causality really means in the quantum world. I am particularly grateful to Jon for setting up the weekly Lobster Lunch group meetings, where everyone can present their latest ideas in a friendly and welcoming manner.

Third, I would like to thank Dr Sina Salek and Dr Daniel Ebler for helping me get started with research in the first year of my DPhil. I got a lot of invaluable advice from them, and eventually we collaborated on a paper together.

I would also like to thank Prof Bob Coecke for interesting discussions and support for various things during the first two years of my DPhil. I am grateful to Bob for leading the Quantum Group in Oxford during that time and for fostering an open and inclusive environment.

Next, I would like to thank my fellow PhD students in the Quantum Group in Oxford for making my time in the group both happy and lively. In particular, I would like to thank my collaborators Augustin Vanrietvelde, Matt Wilson and Nick Ormrod for fruitful work together, as well as my officemate Nicola Pinzani for all the discussions on the literature – I have learnt a lot from all of you! I would also like to thank the rest of the past and present students in the group, including James, Tein, Isaac, Kathleen, Lia, Luca, Robin, Vojta, Subhayan, Marietta, and everyone else.

At the University of Hong Kong, I would like to thank all of the PhD students in Giulio's group for making me feel at home during my research visits in Hong Kong. Special mention goes to Bai Ge, Mo Yin, Zhao Xiaobin, Meng Fei, Liu Zixuan and Swati for many enjoyable conversations and meals together.

I would moreover like to thank Mike Wenxu Mao, Zhong Yan and Anthony Munson, whom I co-supervised as undergraduate or Master's students with Prof Giulio Chiribella, for our exciting collaborations. It was a pleasure to work with you all!

I would like to thank all of my experimental collaborators, in particular Rob Gardner and Santiago Sempere Llagostera, for agreeing to implement our results in practice, and for their patience and enthusiasm in building the experiments.

Furthermore, I would like to thank Dr Sergii Strelchuk for inviting me to the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, and for our interesting discussions and collaboration that followed. Similarly, I would like to thank Prof Dr Andreas Buchleitner for hosting me at the Department of Physics, University of Freiburg, several times during my DPhil, to continue our collaboration from my Master's work there. In addition, I would like to thank Prof Lucien Hardy and Prof Rob Spekkens for inviting me to the Perimeter Institute for Theoretical Physics in Waterloo, Ontario, for a very stimulating research visit. I would also like to thank St John's College, The University of Hong Kong, for letting me stay in their accommodation during my research visits in Hong Kong.

On the financial side, I would first of all like to thank the UK Engineering and Physical Sciences Research Council (EPSRC) for fully funding my DPhil, through grant EP/R513295/1. I would like to thank various institutions for providing funding that enabled me to travel to international conferences: St Anne's College, University of Oxford; the Department of Computer Science, University of Oxford; the Department of Computer Science, The University of Hong Kong; the Annual Conference on Quantum Information Processing (QIP) 2022, and the Quantum Information Structure of Spacetime (QISS) consortium. I would like to thank the Department of Computer Science, The University of Hong Kong, for funding my research visits to Hong Kong, via the Hong Kong Research Grant Council through grant 17307719 and though the Senior Research Fellowship Scheme SRFS2021-7S02, and the John Templeton Foundation through grant 61466, The Quantum Information Structure of Spacetime (qiss.fr).

Importantly, I would like to thank all the people and departments that enabled me to complete my DPhil, despite suffering from long-term health conditions that make it difficult for me to type on a computer. In particular, I would like to thank the Disability Advisory Service (DAS) and the Department of Computer Science, University of Oxford, for providing me with the highest quality ergonomic equipment and dictation software. This thesis was typeset to a large extent using the Dragon dictation software. I would like to thank Gregor Matl for developing high-level dictation commands in Dragon for mathematical commands in LaTeX and giving me access to them. I would like to thank Ben Solomons for helping me type up and make final corrections to this thesis, in the time when my condition was at its worst.

A big thank you goes to Augustin, Matt, James, Nick, Nicola, Tein, and Rob, as well as my supervisors Jon and Giulio, for proofreading different parts of this thesis. I would also like to thank my examiners Prof Aleks Kissinger and Prof Eric Chitambar for their helpful comments and suggestions, and a lively discussion during the viva.

One of my biggest problems during my DPhil was that I was constantly hungry. As a result, my heartfelt gratitude goes to the catering teams at St Anne's College, Oxford, Green Templeton College, Oxford, and St John's College, Hong Kong, for providing me with affordable, healthy and delicious meals throughout my DPhil – it certainly wouldn't have been the same without lunch at these places.

In addition, I would like to thank my personal trainer Daniel Collado Rodriguez for pushing me to achieve my full potential in fitness, and for his guidance on exercise and high-protein diets.

I would also like to thank my friends from the Arcadian Singers and Wadham College Chapel Choir for all the nice music we sang together, and my singing teacher Stephen Taylor for his musical and technical advice.

On the personal side, I would like to thank all of the friends I have made in Oxford and Hong Kong during my DPhil, as well as at conferences – all of you have made a difference! In particular, I would like to thank my housemates over the years in Oxford: Guillaume, Karl, Sarah, Santiago, Meredith, Anna, Gloria, Wilf and Luana. It has been a real pleasure living with you all. Special mention goes to Karl for all our evenings of cooking humongous meals together and many hours in the gym. I would also like to thank the fellow international students at St John's College, Hong Kong, for making me feel welcome, including Arina, Dominika, Philippine, Victor, Daniel, Daniel, Monty, Mezgeb, Honour and everyone else. I will miss all of our fun adventures in Hong Kong! Moreover, I would like to thank my best friends from Fisher Hall, Imperial College London, for their continued friendship over the last few years and all the fun activities we have done together: James, Alwin, Callum, Giovanni, Merissa, Cynthia, Edita, Shahira, Ed, Bálint, Monika and Sabina.

I am very grateful to Mizu for her love and support over the last two years, and for all the happy times we have had together.

Finally, I would like to thank my parents for supporting and encouraging me throughout my education and career all these years. From violinist to quantum physicist, my parents have always been there for me, and I would not have got to this stage without their academic, financial and personal support and guidance.

Preface

This thesis is based on the following works completed during my DPhil in Computer Science at the University of Oxford, under the supervision of Prof Giulio Chiribella and Prof Jonathan Barrett.

Four theoretical works, together with a smaller number of previously unpublished results, constitute the main part of this thesis:

1. Giulio Chiribella and Hlér Kristjánsson, 'Quantum Shannon theory with superpositions of trajectories', *Proceedings of the Royal Society A*, vol. 475, no. 2225, p. 20180903, 2019. (Ref. [50])

2. Hlér Kristjánsson, Giulio Chiribella, Sina Salek, Daniel Ebler and Matthew Wilson, 'Resource theories of communication', *New Journal of Physics*, vol. 22, p. 073014, 2020. (Ref. [119])

3. Hlér Kristjánsson, Wenxu Mao and Giulio Chiribella, 'Witnessing latent time correlations with a single quantum particle', *Physical Review Research*, vol. 3, no. 4, p. 043147, 2021. (Ref. [121])

4. Hlér Kristjánsson, Yan Zhong, Anthony Munson and Giulio Chiribella, 'Quantum networks with coherent routing of information through multiple nodes', *in preparation*. (Ref. [118])

Additionally, elements of two experimental collaborations, which also include some new theoretical results, are briefly reviewed:

5. Giulia Rubino, Lee Rozema, Daniel Ebler, Hlér Kristjánsson, Sina Salek, Philippe Allard Guérin, Alastair A Abbott, Cyril Branciard, Časlav Brukner, Giulio Chiribella and Philip Walther, 'Experimental quantum communication enhancement by superposing trajectories', *Physical Review Research*, vol. 3, no. 1, p. 013093, 2021. (Ref. [166])

6. Santiago Sempere Llagostera, Robert Gardner, Kwok Ho Wan, Raj B. Patel, Hlér Kristjánsson, Giulio Chiribella and Ian A. Walmsley, 'Experimental perfect correction of noisy channels via quantum superposition of paths', *in preparation*. (Ref. [173])

Specifically, Chapter 3 is based on the framework sections of 1. and 3., Chapter 4 is mostly based on the main part of 2. with some additional previously unpublished results, Chapter 5 is based on a part of the results of 1., Chapter 6 is based on the results section of 3. and the new theoretical results in 6., Chapter 7 is based on short discussions in 2. and 3., Chapter 8 is based on the main part of the results in 4., and Chapter 9 is based on the last part of the results in 3. and 4., some previously unpublished results, and the conceptual discussions in 5. and 6.

The following four works were also completed during my DPhil but are not included as part of the results of this thesis. However, some parts of Chapter 2 (Background) and Chapter 10 (Discussions and conclusions) are based on the results of these works:

7. Hlér Kristjánsson, Robert Gardner and Giulio Chiribella, 'Quantum Communications Report for Ofcom', *Ofcom*, 2021. (Ref. [120])

8. Augustin Vanrietvelde, Hlér Kristjánsson and Jonathan Barrett, 'Routed quantum circuits', *Quantum*, vol. 5, p. 503, 2021. (Ref. [191])

9. Augustin Vanrietvelde, Hlér Kristjánsson, Nick Ormrod and Jonathan Barrett, 'Consistent circuits for indefinite causal order', *arXiv preprint arXiv:2206.10042*, 2022. (Ref. [192])

10. Richard Howl, Ali Akil, Hlér Kristjánsson, Xiaobin Zhao and Giulio Chiribella. 'Quantum gravity as a communication resource', *arXiv preprint arXiv:2203.05861*, 2022. (Ref. [110])

The results in this thesis based on the works 1.–4. were primarily completed by me under the supervision of Prof Giulio Chiribella, with smaller contributions from the other authors. Major results in the works 1.–4. that were completed entirely or initially by other authors are generally omitted from this thesis. The exception is when those results are integral to the flow of the thesis. A list of such results is given in the following paragraph:

Proposition 15 and Theorem 18 in Chapter 6 were both proven by my supervisor Prof Giulio Chiribella in 3. (indeed, the main part of the proof of the latter is through Lemma 24, the proof of which is omitted from this thesis). In Chapter 8, the derivations of Eq. (8.7) in §8.3 and (8.11) in §8.5 were initially found by Anthony Munson, and also appear in his Master's Thesis [143].

Additionally, credit should go to Wenxu Mao for performing numerical simulations, presented in his Bachelor's thesis [137], which were a first step towards the results presented in Figure 6.4a. The same goes for Yan Zhong, who came up with the initial idea of point (3) in Theorem 21.

The works 5.-6. are primarily experimental works, where my contribution lay in formulating the precise tasks to be experimentally implemented, translating some of the mathematical subtleties into physical terms, and interpreting the data obtained. In 6., I was moreover responsible for the new theoretical results.

1
Introduction

Since the foundations of information theory were laid by Claude Shannon in the early 20th century [174], the age of information has revolutionised modern science and technology. The discipline which studies information transfer over noisy communication channels became known as Shannon theory, which underlies all practical communication technologies. Around the same time, the development of quantum theory has led to the discovery of new forms of counterintuitive phenomena, such as superposition, entanglement [109] and single-particle interference [197]. In the second half of the century, researchers began investigating the potential of quantum systems for information processing: first for classical information [92, 100, 101, 128] and later for quantum information [83]. These investigations culminated in the seminal works on perfectly secure cryptography without pre-established keys [28, 81]. In 1994, Shor presented a quantum algorithm to factorise numbers in polynomial time [176], enabling the possibility of breaking RSA encryption. Shor's result initiated the field of quantum computation and the quest to build practical quantum computers.

Despite these promising theoretical works, realising quantum computation or quantum communication in practice is a highly non-trivial task, which needs to overcome the inherent obstacles of decoherence and noise present in any quantum system [124, 189]. This led to Shor's proposal of error correcting codes [177] and fault-tolerant quantum computation [178], which laid the groundwork for both theoretical and experimental research in quantum information theory to the present day. In particular, the idea of a quantum capacity of a noisy quantum channel was proposed [177], defining the problem of achieving optimal communication rates in the new field of *quantum* Shannon theory. An important result by Smith and Yard [182], building on the work of Ref. [183], highlighted a striking non-classical effect: when two noisy channels which individually cannot transmit any information are combined in parallel, it is possible that they can be activated to transmit information. These works on Shannon theory have applications not only in genuine communication scenarios, but also in any implementation of quantum computation in the presence of noise [122], where information needs to be transferred from one part of the computer to another [38], or between multiple different computers.

At a fundamental level, we can understand this transition from classical to quantum Shannon theory in terms of a transition in the physical description of information carriers. In Shannon's original formulation of information theory, the information carriers were assumed to obey the laws of classical physics. That is, they were assumed to have perfectly distinguishable internal states, and, to travel along definite trajectories in spacetime. However, the ultimate description of nature is given by the laws of quantum physics, where many of the fundamental assumptions in classical physics no longer hold. Quantum Shannon theory can be viewed as the generalisation of Shannon's original theory where the internal states of information carriers can be *quantum* states. That is, an information carrier no longer needs to be in a perfectly distinguishable internal state, and can be in a superposition of different classical states.

Although significant advances in technological capabilities have been made possible by the transition from classical to quantum Shannon theory, there is still a sense in which the transition to a fully quantum description of information is incomplete. Despite the possibility of encoding messages in quantum states, the trajectories of the information carriers in spacetime have, until now, still been assumed to be classical. Yet, as the double-slit experiment illustrates, quantum particles can travel simultaneously in a quantum superposition of multiple trajectories.

This ability to simultaneously propagate along multiple trajectories in a superposition allows a single quantum particle to undergo superpositions of multiple different evolutions [9, 146]. This leads to an interference between the different evolutions on alternative trajectories. In particular, the 2005 work of Gisin et al. [91] showed that when the processes occurring on different trajectories are noisy, they can interfere destructively and lead to a less noisy process overall. This was demonstrated in experiment in Ref. [123].

More exotically, research in quantum foundations over the last two decades has shown that a single particle can experience two alternative processes in a superposition of their two possible orderings [53, 57]. When the processes correspond to noisy communication channels, the interference between the two orders can lead to an increase in the capacity to communicate both classical and quantum bits [58, 80, 168]. These effects have also been demonstrated in a series of experiments [93, 94, 97, 160, 165]. Whether these experiments genuinely demonstrated an *indefinite causal order* of events, as per the original formulations in Refs. [53, 57, 149], or rather, simply more elaborate versions of a superposition of trajectories, as per Gisin et al., has been a subject of considerable debate [147, 152]. In any case, they highlight the practical potential of a communication model which includes the possibility of information carriers experiencing (some form of) superposition of evolutions.

The above examples of communication advantages arising from superpositions of multiple different evolutions showcase the need for a broader framework of information theory, where both the messages themselves and their propagation in spacetime are treated as quantum states. Such an extension can be regarded as a second level of quantisation of Shannon theory. Whilst the first level of quantisation promoted information carriers' internal degrees of freedom to quantum states, the second level of quantisation should additionally allow their external degrees of freedom to be quantum, thus enabling the possibility of transmitting messages in a coherent superposition of trajectories through alternative communication channels. The formalisation of such a quantisation, however, poses several challenges.

The first challenge is to formalise a communication paradigm in which a superposition of transmission lines can be built solely from the devices available to the communicating parties. A sender and receiver with access to two communication channels should be able to build a new device, corresponding to a superposition of the original channels. Mathematically, a transmission line is described by a quantum channel, however, the existing constructions of a superposition of two quantum channels depends not only on the channels themselves, but also on their physical realisation through an interaction with the environment [2, 9, 146]. This means that when a sender and receiver only know the input-output description of their transmission lines, it is not clear how they should construct a superposition of those devices. In particular, within a Shannon-theoretic description, an important task is to specify precisely what the communication resources available to the communicating parties are, and to develop such a specification in a way that enables communication channels to be naturally combined in a superposition.

The second challenge is to specify a clear-cut separation between the internal degrees of freedom (used to encode the message) and external degrees of freedom (along which the message propagates). Such a separation is not guaranteed by simply specifying which are the external and internal degrees of freedom: the possibility of phase kickback enables information to flow from an information-carrying system to an ancillary system if the two systems interact [62]. If the sender uses phase kickback to encode information in the path of the particle, then the path simply becomes part of the message, in which case the entire communication scenario can be described by standard quantum Shannon theory with a larger information-carrying system. In contrast, a *bona fide* second level of quantisation should pertain only to the study of communication scenarios where the message and path are assigned fundamentally different information-theoretic roles.

In this work, we construct a framework for the development of a second-quantised Shannon theory, which satisfies the above desiderata. The framework uses the tools of *higher-order transformations* (also known as supermaps) [51, 52, 54] to define an operation that combines two communication devices in a superposition. To ensure this is well defined, the framework relies on the abstract notion of vacuum, which describes the absence of an input to a communication device. Accordingly, communication devices are modelled as quantum channels that can act on the internal degree of freedom used as the message, on the vacuum, and on a coherent superposition of the message and vacuum. A superposition of two communication channels is then described by coherently controlling which communication channel receives the message as input and which receives the vacuum as input. Such a construction based on the coherent routing of the message and vacuum provides a well-defined way to construct a superposition of channels from purely the devices already available to the sender and receiver.

The corresponding shift in our description from the states of individual particles to the occupation of a set of modes (that is, whether or not a particle is present or the vacuum), resembles the shift from quantum mechanics to its second-quantisation [75, 84]. This gives further meaning to the title of this work.

To formally specify the information-theoretic distinction between the internal and external degrees of freedom, we use the tools of *resource theories* [60, 67]. We construct a general family of resource theories of communication, specifying a minimal requirement that all meaningful communication models must satisfy. We cast standard quantum Shannon theory as a resource theory in this way. We define an extension of quantum Shannon theory that includes superpositions of alternative trajectories by appending the original resource theory with the operation that places two communication channels in a superposition, as described in the previous paragraph. Our minimum requirement specifies natural constraints on communication models involving (any form of) superpositions of alternative evolutions. We show that our framework of *quantum Shannon theory with superpositions of trajectories* satisfies this requirement, whilst another recently proposed framework [96] does not.

For completeness, we also construct a separate resource theory that describes the potential use of communication channels in a superposition of causal orders, based on Refs. [58, 80, 168]. We show that this framework of *quantum Shannon theory with superpositions of causal orders* also satisfies our minimal requirement. Our resource-theoretic description highlights several fundamental differences between the superposition of trajectories and the superposition of causal orders, which we discuss, following several recent comments on the similarities and differences between these two paradigms [2, 58, 133].

With our framework in hand, we proceed to explore its applications to a variety of communication scenarios, showcasing some of the communication advantages that can be attained by allowing information carriers to travel in a superposition of trajectories. We begin by considering the transmission of information on a superposition of alternative paths, in the spirit of Refs. [2, 3, 91, 146]. We find that certain channels, which are so noisy that no classical information can be transmitted through them when used on their own, can be activated to transmit classical information at a non-zero rate when used in a superposition of paths. The same is true for quantum information: channels which completely decohere the message when used on their own can, in specific cases, be transformed into a channel that transmits qubits at a non-zero rate. Yet, the superposition of two independent noisy transmission lines can never lead to a perfectly noiseless channel.

Building on the results for independent channels, we move to examine the, as of yet unexplored, possibility of placing communication channels with correlated noise between successive uses [117, 135] in a superposition. The presence of correlations in the noise experienced by particles sent through the same transmission line at different times arises naturally in many physical scenarios. For example, an optical fibre induces random changes in the polarisation of photons passing through it, and since these changes happen on a finite timescale, photons sent at close times will experience approximately the same noisy processes. In standard quantum Shannon theory, the presence of correlations is both a threat and an opportunity for communication. On the one hand, it can undermine the effectiveness of standard error correcting schemes, which assume independent errors on the transmitted particles. On the other hand, tailored codes that exploit the correlations among different particles can enhance the transmission of information [17, 18, 24, 39, 43, 45, 55, 70, 89, 90, 113, 135, 136, 141, 156, 167, 198].

In the context of a second level of quantisation of Shannon theory, just as a single quantum particle can travel in a superposition of two paths, a single quantum particle can also be transmitted at a superposition of alternative moments in time. This leads to a striking foundational result, which is interesting in its own right independently of information-theoretic considerations. Classically, if a communication channel exhibits correlations in noise between its use at two consecutive time steps, then in order to probe these correlations, a particle must be sent through each of these two time steps. A single classical particle, however, can only travel through the channel at one of these time steps or the other, and is therefore unable to probe the correlations between successive uses. In stark contrast, we find that a single quantum particle sent at a superposition of going through the channel at one time step or another can probe these correlations, even though the particle only traverses the channel once.

Taking advantage of the time-correlations in the noise, we show that it is possible to enhance the amount of information that a single particle can carry from a sender to a receiver, beating the ultimate limit achievable in the lack of correlations. We demonstrate this effect with two extreme examples. In the former, a single quantum particle can transmit one bit of classical information through a transmission line that completely erases information at every definite time step, while in the latter, a single quantum particle can transmit one qubit of quantum information through a transmission line that completely decoheres any quantum state at every definite time step. These phenomena witness the presence of correlations between different uses of the transmission line: in the lack of correlations, we show that the maximum number of bits or qubits that can be transmitted through the respective channels is bounded by a value strictly less than one.

Having investigated the enhancements in the transmission of both classical and quantum information through noisy channels, we move to study the potential of the second-quantised Shannon theory for long-distance communication networks, where a single particle can travel in a superposition of multiple different paths through the network. In real-life communication networks, noise in transmission degrades the information transmitted in the message in a way that scales exponentially with the length of the communication channels. Similarly, in optical communication networks, the probability of particle loss scales exponentially with distance [42]. These two sources of errors limit the distance over which information can be transmitted via quantum particles. An important step in bringing quantum Shannon theory with superpositions of trajectories to practical use is to understand how robust its protocols are to communication over a long sequence of channels, and whether it can, to some extent, suppress the exponential degradation of information. An additional source of errors arises in the context of the superposition of paths itself: crucially, the superposition of paths relies on coherence between the two trajectories to be maintained throughout the protocol, however, this is also expected to degrade over long distances.

We show, using a combination of analytical and numerical methods, that the superposition of channels is robust against both particle loss and decoherence on the external degree of freedom. That is, in the case of decoherence on the path affecting the communication protocols where the superposition of channels provided a communication advantage, we find that for every finite amount of decoherence, the superposition of channels still provides a finite communication advantage over the analogous communication scenario where the channels are combined in a purely classical manner. In the case of particle loss with a fixed probability on each path, we find that the probability of loss experienced by the superposition of paths is just the same as that for each path individually.

More striking, however, is our finding that the superposition of channels can suppress the exponential degradation of the message itself, when used in combination with repeaters acting locally on the intermediate nodes of the network. In an extreme example of a sequence of noisy channels, where every channel returns a fixed state and therefore cannot transmit any information on its own, we find that for certain parameters of the channel, a superposition of sequences of such channels can correct the overall error at a finite rate, even in the asymptotic limit of an infinite sequence of channels. Building upon this result, we provide a general theorem specifying the conditions on any noisy channel, such that the possibility of information transmission at a non-zero rate is maintained in the asymptotic limit of an infinite sequence of such channels. Of course, in realistic scenarios, these conditions are unlikely to be satisfied exactly. Accordingly, numerical studies of the communication capacity under deviations of these conditions are provided, showing that our scheme can still provide an enhancement in the communication capacity under small deviations from the ideal case.

To emphasise the practical applicability of our framework, we discuss two recent experimental implementations of the superposition of channels, which have demonstrated the communication advantages in experiment. We focus on the conceptual issues of the implementation and summarise the experimental setup and results.

Finally, we conclude with a discussion and outlook. We discuss the significance of our results and its potential in future communication technologies. We discuss a number of other research directions related to our current framework, including the formal study of quantum circuits, indefinite causal order, and information-theoretic perspectives on quantum gravity, highlighting how our techniques and methods have impacted these areas. We propose that our framework can be further extended to other areas of quantum information and foundations, such as the study of quantum algorithms and quantum causality. In preparation for these extensions, we present our initial theoretical framework on the superposition of channels in full generality, where the quantum channels can in principle represent any type of information-processing device, to be used in communication, computation or other scenarios.

The rest of this thesis is structured as follows. Chapter 2 presents the standard theory and notation used in the remainder of the work and summarises the relevant literature. Chapter 3 presents the framework of superposition of quantum channels. Chapter 4 presents the framework of resource theories of communication, casting standard quantum Shannon theory as well as its extensions to superpositions of trajectories and superpositions of causal orders in this form. Chapter 5 presents communication advantages of our framework for a superposition of independent channels, while Chapter 6 presents analogous communication advantages for a superposition of time-correlated channels. In Chapter 7, observations made in the previous chapters are used to discuss the distinction between the superposition of trajectories and the superposition of causal orders. Chapter 8 extends the framework to communication networks, highlighting the new possibilities arising when a single particle can travel in a superposition of multiple paths through different nodes in the network. Chapter 9 discusses practical considerations, such as the tolerance of the superposition of channels to particle loss and dephasing on the path, as well as recent experimental implementations of our protocols. Chapter 10 provides a discussion of related works, as well as a summary and outlook.

2

Background on selected topics in quantum information and foundations

In this chapter, we review the essential concepts and relevant literature from a variety of fields within quantum information and foundations, which form the basis on which the work of this thesis builds. First, the basic features of quantum theory are reviewed, together with the notation we shall use. Second, we provide an introduction the study of communication capacities in quantum Shannon theory. Third, the concepts of quantum circuits and their transformations are discussed, which are essential for understanding the ways in which communication channels can be composed. Fourth, we discuss the recent literature on an indefinite causal ordering of processes, which cannot be written in the form of a circuit. Fifth, we consider the superposition, interference, or coherent control of quantum channels, which forms the basis of the work of this thesis, with such constructions formally defined in Chapter 3. Sixth, we provide a brief overview of quantum resource theories, which are used in Chapter 4. Finally, we discuss time-correlated quantum channels and their applications to communication, which we use in Chapter 6.

2.1 Basic concepts and notation

This work uses the standard notation and language of quantum information theory, as presented in e.g. Ref. [196]. Additionally, some notation inspired by the process-theoretical approach to quantum theory, following Ref. [56], is introduced. The formal circuit diagrams used to illustrate the key ideas are inspired by the diagrammatic-categorical approach to quantum theory [5, 65] and closely follow the constructions in Refs. [54, 69].

Quantum systems In quantum Shannon theory, the degrees of freedom used to carry information are represented by *quantum systems*. A quantum system A is associated to a Hilbert space $\mathcal{H}_A$. For simplicity, this work only considers finite-dimensional systems.

Quantum states An information carrier is represented by a *quantum state* ρ, i.e. a positive semi-definite linear operator with unit trace on the Hilbert space $\mathcal{H}_A$ (also called a density matrix). The set of all quantum states on a physical system A is denoted $\mathsf{St}(A) \subset L(\mathcal{H}_A)$, where $L(\mathcal{H}_A)$ is the space of all linear operators on $\mathcal{H}_A$.

Pure states are quantum states that can be written in the form

$$\rho = |\psi\rangle\langle\psi| \in \mathsf{St}(A)\,, \tag{2.1}$$

for some $|\psi\rangle \in \mathcal{H}_A$ (where $\langle\psi|$ is the Hermitian conjugate of $|\psi\rangle$). The qubit computational basis states are denoted $|0\rangle\,, |1\rangle$, and the Fourier basis states are denoted

$$|\pm\rangle := \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}\,. \tag{2.2}$$

Quantum states which cannot be written in the form (2.1) are known as *mixed states*, which can always be written in the form

$$\rho = \sum_{i=0}^{r-1} p_i\, |\psi_i\rangle\langle\psi_i|\;, \tag{2.3}$$

where $\{|\psi_i\rangle\}_{i=0}^{r-1}$ is a set of pure states and $\{p_i\}_{i=0}^{r-1}$ are (non-negative) probabilities.

Quantum channels A (single use of a) quantum device, e.g. a computational gate or a transmission line, is described by a *quantum channel* $\mathcal{C}$, i.e. a completely positive trace-preserving (CPTP) map, from $L(\mathcal{H}_A)$ to $L(\mathcal{H}_B)$, where A and B are the input and output quantum systems, respectively. Quantum channels transform quantum states into quantum states. The set of all linear maps from $L(\mathcal{H}_A)$ to $L(\mathcal{H}_B)$ is denoted $\mathsf{Map}(A \to B)$. The set of all quantum channels is denoted $\mathsf{Chan}\,(A \to B) \subset \mathsf{Map}(A \to B)$, and when $A = B$, the shorthand $\mathsf{Chan}(A) := \mathsf{Chan}(A \to A)$ is used. When the input and output states are arbitrary, we simply write Chan.

The action of a quantum channel on a quantum state can be written in the *Kraus representation* as

$$\mathcal{C}(\rho) = \sum_i C_i \rho C_i^\dagger \,, \tag{2.4}$$

where $\{C_i\}$ is a (non-unique) set of *Kraus operators*, satisfying the normalisation condition

$$\sum_i C_i^\dagger C_i = I \,, \tag{2.5}$$

where I is the identity linear map.

Quantum channels are written in calligraphic fonts (e.g. $\mathcal{C}$) and the corresponding Kraus operators are written in standard italics (such as C_i). In particular, a unitary channel formed from a unitary gate U (i.e. a linear map satisfying $U^\dagger U = UU^\dagger = I$) is denoted $\mathcal{U}(\cdot) = U(\cdot)U^\dagger$, and the identity channel is written as $\mathcal{I}(\cdot) = I(\cdot)I^\dagger$.

The composite system consisting of subsystems A and B is denoted by $A \otimes B$. The use of two separate devices $\mathcal{C}, \mathcal{D}$ in parallel is described by their tensor product $\mathcal{C} \otimes \mathcal{D}$. Multiple (independent) uses of the same physical device $\mathcal{C}$ is also described by a tensor product of multiple copies of the channel $\mathcal{C}$. For example, if a transmission line $\mathcal{C}$ is used k times between a sender and receiver, then the overall process is described by the quantum channel $\mathcal{C}^{\otimes k}$.

A single quantum device can have multiple subsystems as inputs and outputs, which are known as *ports* of the device. A quantum device with k input/output ports is described by a k-partite quantum channel $\mathcal{C} \in \mathsf{Chan}\left(A^{(1)} \otimes \cdots \otimes A^{(k)} \to A'^{(1)} \otimes \cdots \otimes A'^{(k)}\right)$ where $\left(A^{(j)}, A'^{(j)}\right)$ denotes the j-th input-output pair. It can be shown that every map $\mathcal{M} \in \mathsf{Map}(A_1 \otimes A_2 \to B_1 \otimes B_2)$ can be decomposed into a sum of product maps, that is,

$$\mathcal{M} = \sum_{j=1}^{L} \mathcal{M}_{1,j} \otimes \mathcal{M}_{2,j} \,, \tag{2.6}$$

with $\mathcal{M}_{1,j} \in \mathsf{Map}(A_1 \to B_1)$ and $\mathcal{M}_{2,j} \in \mathsf{Map}(A_2 \to B_2)$ for all $j \in \{1, \ldots, L\}$.

Any quantum channel $\mathcal{C}$ can be recovered from a unitary channel on a larger Hilbert space via the *Stinespring dilation*. That, is for any channel $\mathcal{C} \in \mathsf{Chan}(A \to B)$, there exists a (not generally unique) unitary channel $\mathcal{U} \in \mathsf{Chan}(A \otimes E \to B \otimes F)$ and a pure state $|\eta\rangle \in \mathcal{H}_E$, where E, F are 'environment' quantum systems, such that

$$\forall \rho \in \mathsf{St}(A): \quad \mathrm{Tr}_F \mathcal{U}(\rho \otimes |\eta\rangle\langle\eta|) = \mathcal{C}(\rho)\,. \tag{2.7}$$

Here, the partial trace over system F corresponds to discarding the system F. Each Stinespring dilation gives rise to a choice of Kraus operators $\{K_i = \langle i|_F\, U\, |\eta\rangle_E\}$, for some basis $\{|i\rangle_F\}$.

Quantum supermaps An operation that transformations a set of quantum devices into another set of quantum devices is described by a quantum supermap [52]. A quantum supermap is a linear transformation from $\mathsf{Map}(A \to B)$ to $\mathsf{Map}(A' \to B')$, where A, A', B, B' are generic systems, such that it always maps quantum channels into quantum channels, i.e. it is required to be (i) linear, to preserve probabilities, and (ii) completely positive, to ensure that it is well-defined when acting locally on a bipartite quantum channel [52] [1].

The tensor product of two quantum supermaps $\mathcal{S} : \mathsf{Map}(A_1 \to B_1) \to \mathsf{Map}(A_1' \to B_1')$ and $\mathcal{T} : \mathsf{Map}(A_2 \to B_2) \to \mathsf{Map}(A_2' \to B_2')$ is the supermap $\mathcal{S} \otimes \mathcal{T} : \mathsf{Map}(A_1 \otimes A_2 \to B_1 \otimes B_2) \to \mathsf{Map}(A_1' \otimes A_2' \to B_1' \otimes B_2')$ defined by the condition

$$(\mathcal{S} \otimes \mathcal{T})(\mathcal{M}_1 \otimes \mathcal{M}_2) := \mathcal{S}(\mathcal{M}_1) \otimes \mathcal{T}(\mathcal{M}_2)\,, \tag{2.8}$$

for all $\mathcal{M}_1 \in \mathsf{Map}(A_1 \to B_1)$ and $\mathcal{M}_2 \in \mathsf{Map}(A_2 \to B_2)$. Since all of the maps in $\mathsf{Map}(A_1 \otimes A_2 \to B_1 \otimes B_2)$ are linear combinations of product maps, this condition uniquely defines the supermap $\mathcal{S} \otimes \mathcal{T}$.

Graphical representation The graphical representations of a quantum state and a quantum channel are shown in Figure 2.1. The action of a channel $\mathcal{C} \in \mathsf{Chan}(A \to B)$ on a state $\rho \in \mathsf{St}(A)$ is represented by connecting the two lines labelled by A.

[1] Note, that here we use the term *quantum supermap* to refer to what in Ref. [52] is called a *deterministic supermap*, as we do not consider non-deterministic supermaps in this work.

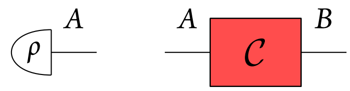


Figure 2.1: *Left: A quantum state* $\rho \in \mathsf{St}(A)$*. Right: A quantum channel* $\mathcal{C} \in \mathsf{Chan}(A \to B)$*.* In this thesis (with the exception of Chapters 4 and 7, where a more elaborate convention is used), quantum channels representing communication channels are drawn in colour to emphasise that they are the resource of interest, whilst quantum states and quantum channels representing other types of operations are kept in white.

2.2 Key concepts from quantum Shannon theory

In this section, we review some of the key definitions and results from classical and quantum Shannon theory. The purpose of the section is to explain the meaning of the classical channel capacity and its quantum analogues: the classical capacity of a quantum channel, and the quantum capacity of a quantum channel. These capacities quantify the amount of information that can be reliably transmitted through a (possibly noisy) communication channel. This brief review follows elements of Ref. [120].

First, we describe the physical setup. The simplest communication scenario between a sender and a receiver is where the two communicating parties are directly connected by a single transmission line. Each use of the transmission line is modelled as a quantum channel $\mathcal{N} \in \mathsf{Chan}(A \to B)$, where A denotes the input system accessible to the sender (Alice) and B denotes the output system accessible to the receiver (Bob). In this case, the question of how much information can be transmitted from the sender to the receiver is determined by the possible encoding and decoding operations performed locally by the sender and receiver, respectively, over some given number of channel uses k [196]. The capacities described below quantify the ability of a communication channel to transmit information between the communicating parties, where both Alice and Bob are assumed to be able to perform any local encoding or decoding operations. Other types of communication capacities corresponding to more complex communication scenarios, for example with multiple senders or receivers, or where the communicating parties have access to entangled states, can be defined in a similar way. For a more detailed discussion, the reader is referred to the standard textbook by Wilde [196].

In the following, we use 'communication' in the broadest meaning of the term, where Alice and Bob could either be distant communicating parties, or, more generally any two spatially separated points in an information-processing quantum device, such as a quantum computer.

2.2.1 Communication rate of classical information

We begin by reviewing the definition of the channel capacity for classical communications. Building up to this definition, we first define the concepts of entropy and mutual information.

Entropy Consider an experiment with a list of possible classical outcomes x of a random variable X, which each occur with probability $p(x)$. Classical Shannon theory tell us that the expected amount of information gained from the result of the experiment is given by the *entropy*

$$H(X) = -\sum_{x} p(x) \log p(x), \tag{2.9}$$

which is given in units of bits (here we take the logarithm to be base 2). For example, a fair coin has the probability of each of heads or tails equal to 1/2, giving an entropy of H=-0.5log0.5-0.5log0.5=1 bit.

Mutual information When information is transmitted through a noisy communication channel, then in addition to the randomness inherent in the choice of message (as quantified by the probability distribution $\{p(x)\}$), the errors occurring in the channel induce a second level of randomness. Alice encodes her message in a set of symbols $\{x\}$, which are instances of the random variable X, each sent with probability $p(x)$. Now, the possible outcomes Bob can receive are given by a set of symbols $\{y\}$, which are instances of the random variable Y, and correspond directly to the $\{x\}$. However, due to errors in transmission, Bob cannot decode the message directly with knowledge only of the sending probabilities $p(x)$. Instead, he needs to consider the conditional probabilities $p(x|y)$, which correspond to the probability that Alice actually sent x given that Bob received y. The probabilities $p(x|y)$ fully characterise the noisy channel [196]; we assume that it is possible for the communicating parties to obtain knowledge of this distribution. Mathematically, the amount of transmitted information that is error free is quantified by the *mutual information*:

$$I(X;Y) = \sum_{x}\sum_{y} p(x,y) \log \frac{p(x|y)}{p(x)} = H(X) - H(X|Y). \tag{2.10}$$

Channel capacity The fundamental limit on how much information can be transmitted through a noisy channel $\mathcal{N}$ is given by the *channel capacity* $C(\mathcal{N})$, defined as the maximum number of bits that can be transmitted through the channel per use of the channel. The channel capacity is equal to the maximum of the *mutual information* of the channel $\mathcal{N}$ over all possible choices of sending probabilities [174]:

$$C(\mathcal{N}) = \max_{\{p(x)\}} I(X;Y). \tag{2.11}$$

2.2.2 Communication rate of classical information through quantum channels

In this subsection, we quantify the classical capacity of a quantum channel. In order to understand the definition of classical capacity, we start by defining the quantum entropy, which enables the definition of the conditional quantum entropy, followed by the quantum mutual information and Holevo capacity.

Quantum entropy The entropy of a classical system corresponded to how much uncertainty there was in obtaining one of the possible outcomes. If we are sure what outcome is to come, then the entropy is zero; if all outcomes are equiprobable then the entropy is maximal. In the quantum case, a definition of entropy will have to capture both the classical uncertainty associated with Alice's choice of message, as well as the quantum uncertainty arising from the possible choices of measurement basis.

This is done by replacing the probability distribution of Alice's preparation in the classical case by the quantum state of Alice's preparation. The *quantum entropy*, also known as the *von Neumann entropy*, of a state ρ_A is given by:

$$H(A)_\rho := -\operatorname{Tr}(\rho_A \log \rho_A) \tag{2.12}$$

Operationally, suppose Alice produces states $|\psi_x\rangle$ in her lab with some probability distribution p_x. Then from Bob's point of view, before Alice sends him the state, his expected state is $\rho_A = \sum_x p_x |\psi_x\rangle\langle\psi_x|$. Now, if the states $\{|\psi_x\rangle\}$ form an orthonormal basis, then we recover the classical entropy of the probability distribution $\{p_x\}$.

Conditional quantum entropy In order to define the quantum analogue to the mutual information, we first need a definition of conditional quantum entropy. This turns out to be a non-trivial task, as there is no direct analogue of conditional probabilities in the quantum case. The following definition is chosen [196]: The *conditional quantum entropy* of the joint state ρ_{AB} of Alice and Bob's composite system AB is

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho \tag{2.13}$$

where $\rho_B := \mathrm{Tr}_A\, \rho_{AB}$ is the marginal state as seen by Bob. (Here, Tr_A is the partial trace over subsystem A only.)

Counterintuitively, the conditional quantum entropy can be negative, illustrating one of the most important differences between the classical and quantum worlds. For example, consider the maximally entangled Bell state between Alice and Bob $|\Phi^+\rangle := (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)/\sqrt{2}$. $H(AB)_\rho = 0$, but the marginal state as seen by Bob is $\rho_B = \mathrm{Tr}_A\, |\Phi^+\rangle\langle\Phi^+| = I/2$, i.e. the maximally mixed state, which has a quantum entropy $H(B)_\rho = 1$, leading to $H(A|B)_\rho = -1$. Conceptually, the negative conditional quantum entropy can be understood as quantifying the fact that we can know more about an entangled state as a whole than about any of its individual components.

Quantum mutual information The *quantum mutual information* can now be defined in complete analogy with the classical case:

$$I(A;B)_\rho := H(A)_\rho - H(A|B)_\rho, \tag{2.14}$$

or equivalently, $I(A;B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$. As in the classical case, the quantum mutual information describes the amount of classical correlations between two systems.

The most relevant communication scenario where this is useful is in the case of classical communication between two parties. Consider the scenario where Alice sends classical information through a noisy channel $\mathcal{N}$ to Bob. Alice prepares the states ρ_A^x which can be input into the channel, each with probability $p(x)$, and keeps a copy of the index of her chosen state in a classical register X with states $\{|x\rangle\}$. The expected density matrix of her prepared state is then

$$\rho_{XA} = \sum_x p(x)\, |x\rangle\langle x|_X \otimes \rho_A^x. \tag{2.15}$$

Now if she sends the state of the A system over to Bob via a quantum channel $\mathcal{N}_{A\to B} \in \mathsf{Chan}(A \to$

B), then the joint state of Bob's received system B with Alice's remaining classical register X is

$$\rho_{XB} = \sum_x p(x) \, |x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_A^x). \tag{2.16}$$

The quantum mutual information $I(X;B)_\rho$ of this state gives a measure of the classical information that Alice was able to transmit to Bob.

Holevo capacity of a quantum channel The fundamental question here, just as in the case of classical channels, is: what is the maximum amount of information that can be sent through a given quantum channel $\mathcal{N}$. If the input states to each use of the channel are not allowed to be entangled (as in the classical case), then the maximum number of bits that can be sent through a channel $\mathcal{N}$ per use of the channel is given by the *Holevo capacity*[2] (defined analogously to the channel capacity of a classical channel):

$$\chi(\mathcal{N}) := \max_{\rho_{XA}} I(X;B)_\rho \,, \tag{2.17}$$

where the maximisation is over all possible states of the form (2.15).

Classical capacity of a quantum channel In contrast to the classical case, the Holevo capacity is not the ultimate limit of classical information transfer through quantum channels. Crucially, the Holevo capacity is superadditive, meaning that $k\chi(\mathcal{N}) \leq \chi(\mathcal{N}^{\otimes k})$. That is, the Holevo capacity of k parallel uses of $\mathcal{N}$ can be larger than k times the Holevo capacity of $\mathcal{N}$ itself. In particular, the RHS can be made larger than the LHS for certain quantum channels, when successive input states are entangled. This leads to the following expression for the *classical capacity of a quantum channel* (Holevo–Schumacher–Westmoreland Theorem):

$$C_{\mathrm{C}}(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k}\chi(\mathcal{N}^{\otimes k}), \tag{2.18}$$

where $\chi(\mathcal{N}^{\otimes k})$ is the Holevo capacity of k parallel copies of $\mathcal{N}$ [102, 172]. The classical capacity of a quantum channel is the maximum number of bits that can be transmitted through the channel per channel use, in the asymptotic limit of infinitely many channel uses.

[2] In some texts, this quantity is also know as the Holevo *information* of a quantum channel, however, here we shall use the term Holevo capacity to unambiguously distinguish it from the Holevo information of a quantum state [196].

In general, the classical capacity of a quantum channel is intractable to calculate, as it involves the regularisation over infinitely many copies of the channel. In practice, the Holevo capacity is often a useful lower bound.

Example: depolarising channel The depolarising channel $\mathcal{N}_{\text{dep}}$ is an example of a process that describes common sources of noise that hinder the transfer of classical information. Given an input state ρ of a d-dimensional system, the depolarising channel returns the original state with probability $(1-p)$, and replaces it with the maximally mixed state I/d (i.e. white noise) with probability p:

$$\mathcal{N}_{\text{dep}}(\rho) = (1-p)\rho + p\frac{I}{d}\,. \tag{2.19}$$

The depolarising channel can arise via a randomisation over an orthogonal set of unitary matrices. For example, for qubits ($d=2$),

$$\mathcal{N}_{\text{dep}}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z), \tag{2.20}$$

where X, Y, Z are the Pauli matrices.

To illustrate the measures of information described above, consider the completely depolarising channel, i.e. $p = 1$, which is described by

$$\mathcal{N}_{dep}(\rho) = \text{Tr}(\rho)I/d\,. \tag{2.21}$$

This means that every possible output state $\mathcal{N}_{A\to B}(\rho_A^x)$ in the expression for the Holevo capacity (2.16) is independent of the classical register x, so that the quantum mutual information must always be zero. Therefore, both the Holevo capacity and the classical capacity of the completely depolarising channel are zero. A general depolarising channel with an error probability $0 < p < 1$ will have a non-zero Holevo capacity. It has been shown that the Holevo capacity of the depolarising channel is additive (i.e. does not exhibit superadditivity), and therefore the classical capacity is equal to the Holevo capacity [114]. The full expression is given by

$$C(\mathcal{N}_{\text{dep}}) = \log d - H_{\min}(\mathcal{N}_{\text{dep}}) \tag{2.22}$$

bits per channel use, where

$$H_{\min}(\mathcal{N}_{\text{dep}}) = -(1 - p + p/d) \log(1 - p + p/d) - (d - 1)(p/d) \log(p/d) \tag{2.23}$$

is a quantity known as the minimum output entropy of the channel [114].

2.2.3 Communication rate of quantum information through quantum channels

In this subsection, we quantify the quantum capacity of a quantum channel. We begin with the definition of the coherent information of a quantum state, followed by the coherent information of a quantum channel.

Coherent information A standard measure of the amount of quantum correlations between two systems is given by the *coherent information*:

$$I(A\rangle B)_\rho := H(B)_\rho - H(AB)_\rho, \tag{2.24}$$

which, interestingly, is equal to the negative of the conditional quantum entropy. This equivalence can be understood from the example above of the maximally entangled state which has maximal negative conditional quantum entropy. The essence of establishing quantum correlations is establishing entanglement, so it is reasonable that a measure of information which is maximised for a maximally entangled state corresponds to a quantification of quantum correlations.

Coherent information of a quantum channel The extent to which a quantum channel can preserve quantum correlations in the presence of noise is quantified by the coherent information of a quantum channel, defined similarly to the Holevo capacity in the case of classical information. Consider a pure bipartite state $\varphi_{AA'}$ in Alice's possession, where she sends the A' system to Bob through a channel $\mathcal{N}_{A' \to B} \in \mathsf{Chan}(A' \to B)$. If the initial state $\varphi_{AA'}$ is entangled, then the communication protocol would want the final state $\rho_{AB} = \mathcal{N}_{A' \to B}(\varphi_{AA'})$ to preserve this entanglement as much as possible. The *coherent information of a quantum channel* is defined as

$$Q(\mathcal{N}) := \max_{\varphi_{AA'}} I(A\rangle B)_\rho, \tag{2.25}$$

where the maximisation is with respect to all pure states $\varphi_{AA'}$. Just like the Holevo capacity of a quantum channel, the coherent information of a quantum channel is not in general additive, meaning that $kQ(\mathcal{N}) \leq Q(\mathcal{N}^{\otimes k})$.

Quantum capacity of a quantum channel The quantum capacity of a quantum channel is the maximum number of qubits that can be transmitted through the channel per channel use, in the asymptotic limit of infinitely many channel uses. This is the fully quantum analogue of the channel capacity in classical Shannon theory. The *quantum capacity of a quantum channel* is given by the *quantum capacity theorem*:

$$C_Q(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k} Q(\mathcal{N}^{\otimes k}). \tag{2.26}$$

where $Q(\mathcal{N})$ is the coherent information of channel $\mathcal{N}$. (This a result of the work of many different authors; a standard reference is the textbook by Wilde [196].)

The quantum capacity possesses a striking property: it is possible that two quantum channels, each with zero quantum capacity when used individually, have a non-zero quantum capacity when combined in parallel – a phenomenon known as superactivation [182].

Example: dephasing channel The paradigmatic error which affects the communication of quantum information, but not classical information, is the *dephasing error*. The dephasing error collapses a quantum state into a given classical basis, thus losing its essential quantum properties. Dephasing occurs naturally in most quantum systems due to their interaction with the environment. Mathematically, a dephasing channel $\mathcal{N}_{\text{deph}}$ has the following form:

$$\mathcal{N}_{\text{deph}}(\rho) = (1-p)\rho + p\sum_{i=0}^{d-1} |i\rangle\langle i|\,\rho\,|i\rangle\langle i|\,, \tag{2.27}$$

where $|i\rangle$ are the basis vectors of a given basis (of dimension d).

Now consider the qubit completely dephasing channel ($p = 1$). To see what it does to the coherent information and quantum capacity, we find that if one half of an entangled state is sent through a completely dephasing channel, then the entanglement is completely destroyed. This means that the coherent information is zero, and therefore also the quantum capacity. On the other hand, the computational basis states $|0\rangle$ and $|1\rangle$ can be transmitted error-free through the channel,

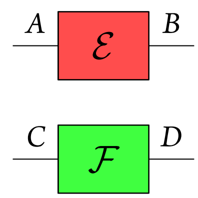


Figure 2.2: *The parallel composition of quantum channels* $\mathcal{E} \in \mathsf{Chan}(A \to B)$ *and* $\mathcal{F} \in \mathsf{Chan}(C \to D)$. In this thesis (except in Chapters 4 and 7), we shall use different colours to represent each independent communication device.

meaning that the classical capacity of a completely dephasing channel is 1 bit per channel use.

2.3 Quantum circuits and higher-order transformations

In real-life communication scenarios, there are typically multiple transmission lines connecting the sender and receiver, connected together in a communication network. The transmission lines are modelled as quantum channels with input and output systems which are accessible to the sender, intermediate parties, or the receiver. Similarly, in computation, an algorithm is usually performed by connecting multiple quantum gates together in a network. Both of these scenarios are formally described by a *quantum circuit* [5, 6, 65, 73, 145]. A quantum circuit is the most general object formed by combining quantum channels (and possibly quantum states) in parallel and sequence.

2.3.1 Connecting quantum circuits

Mathematically, the construction of a quantum circuit from elementary quantum channels can be described as follows, in the language of Refs. [51, 54].

To combine two quantum channels $\mathcal{E} \in \mathsf{Chan}(A \to B)$ and $\mathcal{F} \in \mathsf{Chan}(C \to D)$ in parallel, we simply construct their tensor product $\mathcal{E} \otimes \mathcal{F} \in \mathsf{Chan}(A \otimes C \to B \otimes D)$. In diagrams, this corresponds to just drawing the two channels side-by-side, with the inputs and outputs aligned, as shown in Figure 2.2.

Combining two quantum channels $\mathcal{E} \in \mathsf{Chan}(A \to B)$ and $\mathcal{F} \in \mathsf{Chan}(C \to D)$ in sequence is done by simple concatenation of the channels: $\mathcal{F} \circ \mathcal{E} \in \mathsf{Chan}(A \to D)$. In diagrams, this means that the output of the first channel is connected to the input of the second channel. When the channels have multiple input/output systems, only inputs and output of the same type (i.e.,

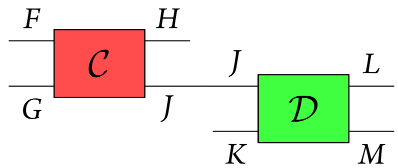


Figure 2.3: *The sequential composition of quantum channels* $\mathcal{C} \in \mathsf{Chan}(F \otimes G \to H \otimes J)$ *and* $\mathcal{D} \in \mathsf{Chan}(J \otimes K \to L \otimes M)$ *formed by contraction over system* J*.*

corresponding to the same Hilbert space) can be connected together. All inputs/outputs that are not involved in the concatenation are formally concatenated with an identity channel. That is, for two bipartite channels $\mathcal{C} \in \mathsf{Chan}(F \otimes G \to H \otimes J)$ and $\mathcal{D} \in \mathsf{Chan}(J \otimes K \to L \otimes M)$, the sequential composition over the system J is given by

$$(\mathcal{I}_H \otimes \mathcal{D}) \circ (\mathcal{C} \otimes \mathcal{I}_K) \,. \tag{2.28}$$

This is illustrated in Figure 2.3.

Formally, the sequential composition is constructed in the following way. A quantum channel $\mathcal{C} \in \mathsf{Chan}(A \to B)$ can be written in the Choi-Jamiołkowski representation [61, 112] as a quantum state $c \in \mathsf{St}(A \otimes B)$, with the isomorphism

$$\begin{aligned} c = \mathrm{Choi}(\mathcal{C}) &:= \mathcal{I} \otimes \mathcal{C}(|\Omega\rangle\langle\Omega|) \,, \\ \mathcal{C}(\rho) = \mathrm{Choi}^{-1}(c)(\rho) &:= \mathrm{Tr}_A \left[\left(\rho^T \otimes I_B \right) c \right] \,, \end{aligned} \tag{2.29}$$

where c is the Choi operator. Here, $\mathcal{I} \in \mathsf{Chan}(A \to A)$ is an identity channel, $|\Omega\rangle := \sum_n |n\rangle \otimes |n\rangle \in \mathcal{H}_A \otimes \mathcal{H}_A$ is the un-normalised maximally entangled state, and T is the transpose operation with respect to the orthonormal basis $\{|n\rangle\}$.

Two quantum channels taking multiple systems as input and output can be sequentially composed by connecting any number of outputs from one channel to inputs of the other channel. In this case, we write the systems to be connected using the same letter. For example, two channels with bipartite input and output systems and one connection are denoted $\mathcal{C} \in \mathsf{Chan}(F \otimes G \to H \otimes J)$ and $\mathcal{D} \in \mathsf{Chan}(J \otimes K \to L \otimes M)$. Mathematically, the composition is done by taking the link product $*$ of the Choi operators corresponding to the two channels. [51]. That is, the new channel

$\mathcal{E}$ is defined by its Choi operator e:

$$e = c * d := \mathrm{Tr}_J[(c_{FGHJ} \otimes I_{KLM})^{T_J}(I_{FGH} \otimes d_{JKLM})] \,, \tag{2.30}$$

where T_J is the partial transposition over the connected system J [51].

The above scheme captures the most general construction of quantum circuits formed from parallel and sequential compositions of quantum channels. Note, that quantum states can be connected in the same way, by treating states as channels with a trivial input system.

An equivalent purely graphical formulation of quantum circuits is given in Refs. [64, 65]. This diagrammatic calculus for representing quantum theory was introduced by Abramsky and Coecke in Ref. [5] and is grounded in category theory [27, 63, 66]. Specifically, Hilbert spaces and quantum states and channels are described by a dagger-symmetric monoidal category, which can be faithfully represented by two-dimensional diagrams obeying a set of rules.

2.3.2 Higher-order computation: transforming quantum circuits

In the above, we treated quantum channels as the highest level objects in our framework. That is, we connected quantum channels together to form a quantum circuit (mathematically, just a big quantum channel), after which the circuit was considered a fixed object. This fixed quantum circuit was then able to transform any choice of input quantum states into corresponding output states. Physically, this could correspond to connecting together either logical gates in computation, or transmission lines in communication.

However, we could take the idea of transforming quantum channels further than just simple combinations in parallel and sequence. Specifically, we could consider the possible transformations of quantum channels and quantum circuits as *higher-order transformations* [51, 52], which we study in their own right. We call these higher-order transformations of quantum channels *quantum supermaps* (see §2.1).

To illustrate the idea of supermaps, consider first a single quantum channel. One could ask, what is the most general form of a physically admissible transformation that maps a single quantum channel into a single quantum channel? To be physically admissible, we require the supermap to be linear and completely positive, which guarantees that the supermap transforms any input quantum channel (including when acting locally on part of a bipartite quantum channel) into another

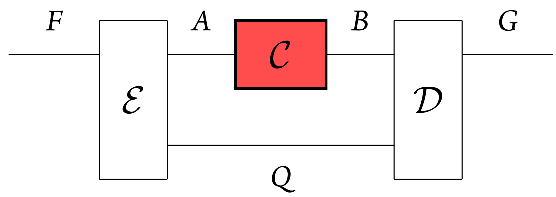


Figure 2.4: *The most general quantum supermap transforming a quantum channel* $\mathcal{C} \in \mathsf{Chan}(A \to B)$ *into a new quantum channel* $\mathcal{S}(\mathcal{C}) \in \mathsf{Chan}(F \to G)$. The supermap is implemented through pre and post-processing the channel $\mathcal{C}$ with quantum channels $\mathcal{E} \in \mathsf{Chan}(F \to A \otimes Q)$ and $\mathcal{D} \in \mathsf{Chan}(B \otimes Q \to G)$, respectively, via an interaction with an auxiliary system Q. Note that only the channel $\mathcal{C}$ is drawn in colour as it represents a communication (or computation) device, whilst $\mathcal{E}$ and $\mathcal{D}$ are considered part of the supermap.

quantum channel [52]. Ref. [52] showed that the most general supermap acting on a quantum channel $\mathcal{C}$ can be written as

$$\mathcal{S}(\mathcal{C}) = \mathcal{D} \circ (\mathcal{C} \otimes \mathcal{I}_Q) \circ \mathcal{E}\,, \tag{2.31}$$

where $\mathcal{E}$ and $\mathcal{D}$ are quantum channels and $\mathcal{I}_Q$ is the identity channel on an auxiliary system Q. This is illustrated in Figure 2.4.

Physically, this question could arise in a communication scenario, where the quantum channel in question represents a transmission line between the sender and receiver. In this case, channels $\mathcal{E}$ and $\mathcal{D}$ above can be thought of as local encoding and decoding operations performed by the sender and receiver, respectively. Of course, when performing communication with a given transmission line, typically one cannot assume that another transmission line is available. Therefore, in the communication scenario, the identity channel $\mathcal{I}_Q$ on the auxiliary system Q would be taken as trivial (i.e. one-dimensional, which is equivalent to not having Q at all). The question of how one can transform the original transmission line then reduces to the question of optimising the local encoding and decoding operations performed by the communicating parties, as described in §2.2.

Building on the idea of a supermap transforming a single quantum channel, one can ask: what is the most general form of a physically admissible transformation that maps quantum circuits into quantum circuits? In Ref. [51] it was shown that, in fact, any supermap which transforms a quantum circuit into a quantum circuit is itself another quantum circuit, known as a *quantum comb*. The name 'comb' is derived from the fact that by rearranging the configuration of inputs and outputs of a quantum circuit, it can be deformed into the shape of a comb, as illustrated in Figure 2.5, with k slots in sequence. Instead of all input systems taking states as inputs and all output systems return-

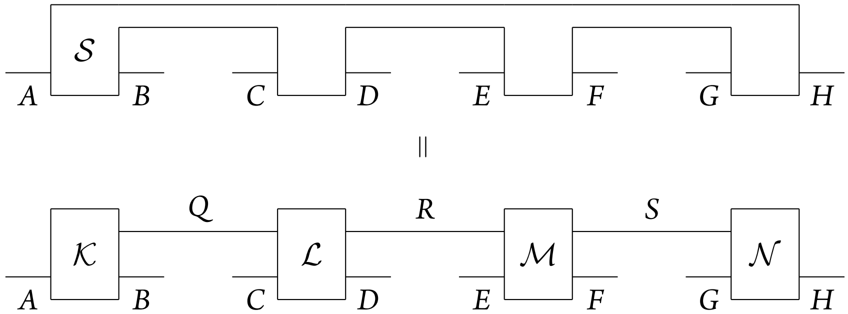


Figure 2.5: *A quantum comb* $\mathcal{S}$*, which takes a set of quantum channels in* $\mathsf{Chan}(B \to C)$, $\mathsf{Chan}(D \to E)$ *and* $\mathsf{Chan}(F \to G)$ *as input (in between its 'teeth') and returns a new quantum channel in* $\mathsf{Chan}(A \to H)$. This can be decomposed into a sequence of quantum channels $\mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}$ with ancillary systems Q, R, S.

ing states as outputs, a quantum comb can naturally take quantum channels as inputs, slotted in between the 'teeth' of the comb. Similarly to the most general form of a supermap transforming a single quantum channel, a quantum comb can always be decomposed into a sequence of quantum channels, one for each 'tooth', connected together via ancillary systems, as shown in Figure 2.5.

In the following, we shall restrict ourselves to quantum combs that take k quantum channels as input and return a single quantum channel as output, known as *quantum k-combs*. Mathematically, a quantum k-comb transforming system S_i into system S'_i at the i-th step is a k-partite quantum channel $\mathcal{C} \in \mathsf{Chan}(S_1 \otimes \cdots \otimes S_k \to S'_1 \otimes \cdots \otimes S'_k)$ satisfying the conditions

$$\begin{aligned}
\mathrm{Tr}_{S'_k}[\mathcal{C}(\rho)] &= \mathrm{Tr}_{S'_k}\left[\mathcal{C}\left(\mathrm{Tr}_{S_k}[\rho] \otimes \frac{I_{S_k}}{d_k}\right)\right] \\
\mathrm{Tr}_{S'_{k-1}S'_k}[\mathcal{C}(\rho)] &= \mathrm{Tr}_{S'_{k-1}S'_k}\left[\mathcal{C}\left(\mathrm{Tr}_{S_{k-1}S_k}[\rho] \otimes \frac{I_{S_{k-1}}}{d_{k-1}} \otimes \frac{I_{S_k}}{d_k}\right)\right] \\
&\vdots \\
\mathrm{Tr}_{S'_1 \cdots S'_k}[\mathcal{C}(\rho)] &= 1 \qquad \forall \rho \in \mathsf{St}(S_1 \otimes \cdots \otimes S_k)\,,
\end{aligned} \tag{2.32}$$

where the last condition is the condition of trace-preservation satisfied by all quantum channels. The set of all quantum k-combs with the input/output systems above is denoted by $\mathsf{Comb}[(S_1 \to S'_1), \ldots, (S_k \to S'_k)]$. (Note, that the choice of maximally mixed state in the conditions is arbitrary; any state could have been used.)

We have seen that the most general transformations of quantum channels connected in a quantum circuit have been characterised, namely, such transformations can always be realised as a quantum comb. However, given only a set of quantum channels without a predefined configuration, we may ask whether there exist transformations which are compatible with quantum theory but cannot be expressed as a quantum comb. Ref. [57] showed that the answer is positive: two quantum channels can be combined in an indefinite causal order, a composition of channels which goes beyond the framework of quantum circuits and combs. This brings us to the next section.

2.4 Indefinite causal order

The idea of performing computations without a definite causal order was first suggested by Hardy in 2009 [98], as potentially providing advantages over the conventional definite ordering of computations. This idea was further developed with the introduction of the quantum `SWITCH` [53, 57], a quantum supermap that combines two quantum channels in a superposition of their two possible causal orderings. In the years since, the study of indefinite causal order has firmly established itself as an active research area. The topic has been studied from a variety of perspectives, which can be broadly classified into two main strands.

The first strand of research is the direct study of its essential features, such as foundational research on causal structures [19, 116, 149, 155] and characterisations of the possible processes with indefinite causal order [13–15, 20–22, 36, 40, 44, 147, 148, 195]. An important concept in the first strand of research described above is *causal inequalities* [149]. These are inequalities designed to characterise processes which locally obey quantum theory but do not exhibit a global causal structure. Causal inequalities for causal order can be seen as analogous to the Bell inequalities for local realism.

The second strand of research on indefinite causal order relates to its potential information-processing advantages. This typically involves the quantum `SWITCH` as a method of combining quantum channels in a way that performs better than combining the channels in parallel or sequence. Such advantages have been found in areas such as quantum query complexity [11, 47, 68, 82], quantum communication complexity [95], and quantum metrology [201], as well as in quantum communication [58, 59, 80, 94, 133, 161, 162, 168]. (Note, that these lists of references are non-exhaustive.)

In the following, causal inequalities, together with the framework of process matrices typically used for their description, and the quantum `SWITCH` are discussed in turn.

2.4.1 Causal inequalities and process matrices

Causal inequalities were first proposed by Oreshkov, Costa and Brukner in 2011, in their seminal paper on quantum correlations without a predefined causal order [149]. Given that many of the fundamental notions of classical physics are subject to uncertainty and superposition in quantum mechanics, the authors suggested the possibility that even causality itself is subject to quantum indeterminism. The authors considered the situation where two experimenters operating in separate local laboratories, individually described by quantum mechanics, exchange messages without assuming the existence of a spacetime or definite causal order in which the individual laboratories are embedded. Strikingly, it was found that one could consistently define correlations between the two parties which are incompatible with a global causal structure, which can be characterised by the violation of causal inequalities. Causal inequalities test whether correlations are compatible with a global causal order, similar to the way in which Bell inequalities test whether correlations are compatible with local realism [25].

To illustrate the idea of a causal inequality, consider the following communication task between two communication parties, Alice and Bob. Each party receives a system into his/her respective laboratory exactly once, but no assumption is made on who receives the system first. When a party receives the system, he/she tosses a coin to obtain a random bit: a for Alice and b for Bob. Bob additionally generates a second random bit b' which determines their task: $b' = 0$ means Bob needs to communicate bit b to Alice, while $b' = 1$ means Bob has to guess bit a. Overall, their goal is to maximise the success probability

$$p_{\text{succ}} := \frac{1}{2}\left[P\left(x = b|b' = 0\right) + P\left(y = a|b' = 1\right)\right]. \tag{2.33}$$

Clearly, if the order in which Alice and Bob receive the system is fixed, then only one of the two alternative tasks can be successfully completed with certainty. Specifically, it can be shown that no strategy assuming a fixed causal structure can exceed the bound

$$p_{\text{succ}} \leq 3/4\,, \tag{2.34}$$

which is known as a causal inequality. However, by relaxing this assumption, the bound can be violated.

The description of processes consistent locally with quantum mechanics but inconsistent with a global order of events is typically described by the formalism of process matrices, which are a generalisation of quantum combs. Consider two parties Alice and Bob, with corresponding input and output systems A_I, A_O and B_I, B_O, respectively. Each party is free to perform any quantum operation from their input to output system. A quantum operation is described by a set of completely positive trace-non-increasing (CPTNI) maps $\{\mathcal{M}_i^{X_I X_O}\}$ labelled by possible outcomes i, with $X \in \{A, B\}$, where the sum of the possible maps $\sum_i \mathcal{M}_i^{X_I X_O}$ is a CPTP map. The joint probability of Alice obtaining result i and Bob obtaining result j is given by

$$P\left(\mathcal{M}_i^A, \mathcal{M}_j^B\right) = \mathrm{Tr}\left[W^{A_I A_O B_I B_O}\left(M_i^{A_I A_O} M_j^{B_I B_O}\right)\right], \tag{2.35}$$

where $M_i^{X_I X_O}$ is the Choi operator of $\mathcal{M}_i^{X_I X_O}$ and $W^{A_I A_O B_I B_O}$ is an operator on the tensor product system $A_I \otimes A_O \otimes B_I \otimes B_O$, called the *process matrix*. In order to obtain non-negative probabilities which sum up to 1, W is subject to the following constraints:

$$W^{A_I A_O B_I B_O} \geq 0, \tag{2.36}$$

$$\begin{aligned} \mathrm{Tr}\left[W^{A_I A_O B_I B_O}\left(M^{A_I A_O} M^{B_I B_O}\right)\right] &= 1, \\ \forall M^{A_I A_O}, M^{B_I B_O} \geq 0, \mathrm{Tr}_{A_O} M^{A_I A_O} = I^{A_I}, \mathrm{Tr}_{B_O} M^{B_I B_O} &= I^{B_I}. \end{aligned} \tag{2.37}$$

Any linear operator W on the appropriate spaces satisfying the constraints (2.36)–(2.36) is a valid process matrix.

Equations (2.36)–(2.37) are equivalent to the constraints on a quantum comb presented in [51], albeit without the condition of a definite causal order. In Ref. [181], it was shown that the process matrix formalism is moreover equivalent to the formalism of two-time quantum states [7, 8, 180]. This means that process matrices describe processes with independent initial and final boundary conditions, both of which occur deterministically.

The process matrix can be seen as a generalisation of a density matrix (i.e. quantum state), with Eq. (2.35) a generalisation of the Born rule. Indeed, the simplest example of a process matrix is a density matrix on the input systems, i.e. $W^{A_I A_O B_I B_O} = W^{A_I B_I} \otimes I^{A_O, B_O}$, corresponding to a quantum state shared by Alice and Bob.

Another example of a process matrix is when a state ρ^{A_I} initially possessed by Alice is sent to Bob via a quantum channel $\mathcal{C}$. This is described by the process matrix $W^{A_I A_O B_I B_O} = I^{B_O} \otimes (c^{A_O B_I})^T \otimes \rho^{A_I}$, where $c^{A_O B_I}$ is the Choi operator of $\mathcal{C}$. More generally, any quantum memory channel (§2.7) from Alice to Bob is described by a process matrix with the form $W^{A_I A_O B_I} \otimes I^{B_O}$. In this case, Alice can communicate to Bob but Bob cannot communicate to Alice. Process matrices of this form are denoted $W^{B \not\preceq A}$ Alternatively, the situation where Bob can communicate to Alice but not vice versa is described by a process matrix of the form $W^{B_I B_O A_I} \otimes I^{A_O}$, denoted $W^{A \not\preceq B}$. In a definite causal structure, the most general quantum process occurring between two parties is described by a convex combination of process matrices of the two above forms:

$$W^{A_I A_O B_I B_O} = qW^{B \not\preceq A} + (1-q)W^{A \not\preceq B}\,, \tag{2.38}$$

for some probability q, $0 \leq q \leq 1$.

Quantum processes satisfying Eq. (2.38) are *causally separable* and represent the most general bipartite quantum processes where the two parties are localised in closed laboratories within a definite causal structure.

Counterintuitively, Ref. [149] showed that there exist processes which are locally compatible with quantum mechanics, but are inconsistent with the existence of a global causal structure. These *causally non-separable* processes are described by process matrices not of the form of Eq. (2.38), which cannot be decomposed into a quantum circuit. An example is the process matrix

$$W^{A_I A_O B_I B_O} = \frac{1}{4}\left[I^{A_I A_O B_I B_O} + \frac{1}{\sqrt{2}}\left(Z^{A_O} Z^{B_I} + Z^{A_I} X^{B_I} Z^{B_O}\right)\right], \tag{2.39}$$

where the input and output systems are all two-dimensional and X, Z are the Pauli X and Z operators. Ref. [149] showed that the above process matrix defines a valid strategy (i.e. satisfies the constraints (2.36)–(2.36)) for the random bit game described above, leading to a success probability $p_{\text{succ}} = (2+\sqrt{2})/4$, which violates the causal inequality bound of $3/4$ for the case of a definite causal structure.

The possibility of accomplishing tasks that were impossible under the assumption of a fixed global causal order has given rise to a large body of works which have further elaborated on the characterisation of causally non-separable processes via causal inequalities [1, 20–22, 33, 40, 142]. Others works have employed the formalism of process matrices to investigate different aspects of indefinite causal order, for example, Ref. [44] examines the dynamics of causal structures, while Ref. [147] analyses the quantum `SWITCH`, which is described next.

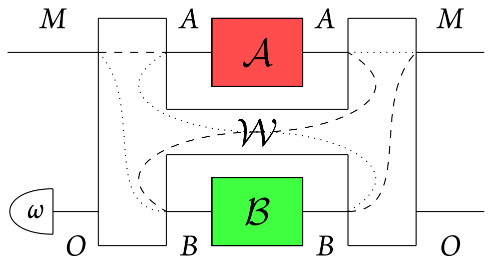


Figure 2.6: *The quantum* SWITCH *supermap transforming two quantum channels* $\mathcal{A} \in \mathsf{Chan}(A)$ *and* $\mathcal{B} \in \mathsf{Chan}(B)$ *into a new quantum channel* $\mathcal{W}(\mathcal{A},\mathcal{B}) \in \mathsf{Chan}(M \otimes O)$*, where* O *is an order qubit determining the causal order of the operations* $\mathcal{A}$ *and* $\mathcal{B}$*, and* $M \simeq A \simeq B$. In computational or communication protocols, O is initially fixed in a state ω and is inaccessible to the sender. The dashed and dotted lines depict the alternative orders corresponding to $\omega = |0\rangle\langle 0|$ and $\omega = |1\rangle\langle 1|$, respectively.

2.4.2 The quantum SWITCH

Perhaps the most prominent example of indefinite causal order is the supermap known as the quantum SWITCH, introduced by Chiribella, D'Ariano and Perinotti in 2009 [53, 57]. The quantum SWITCH takes two quantum channels $\mathcal{A}$ and $\mathcal{B}$ as input, and returns a new quantum channel which is a coherent superposition of both possible concatenations $\mathcal{B} \circ \mathcal{A}$ and $\mathcal{A} \circ \mathcal{B}$ of the channels.

Mathematically, the quantum SWITCH acting on two channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$ is defined by the supermap

$$
\begin{aligned}
&\mathcal{W} : \mathsf{Chan}(A) \times \mathsf{Chan}(B) \to \mathsf{Chan}(M \otimes O) \\
&\mathcal{W}(\mathcal{A},\mathcal{B})(\rho \otimes \omega) = \sum_{i,j} W_{ij}(\rho \otimes \omega) W_{ij}^{\dagger}, \\
&W_{ij} = B_j A_i \otimes |0\rangle\langle 0| + A_i B_j \otimes |1\rangle\langle 1| \ ,
\end{aligned}
\tag{2.40}
$$

where, $\{A_i\}$ and $\{B_j\}$ are *any* Kraus representations of the channels $\mathcal{A}$ and $\mathcal{B}$, respectively, $M \simeq A \simeq B$ is the message system, and ω is the control qubit O determining the causal order of the operations $\mathcal{A}$ and $\mathcal{B}$, which we call the order qubit. The quantum SWITCH is illustrated in Figure 2.6.

Ref. [57] showed that the quantum SWITCH cannot be realised as a quantum circuit. This proof relies on the fact that the quantum SWITCH is a supermap which takes as input only one use of each channel. Nevertheless, a more general model of computation than the circuit model, where the internal wires are movable, is suggested as potentially being able to implement the quantum SWITCH [57]. This has in fact now been achieved in the routed quantum circuits formalism, in parallel to the work of this thesis [19, 134, 191, 192]. In relation to the framework of process matrices discussed above, it is interesting to note that the quantum SWITCH does not violate any causal inequalities [13, 148], however it has been shown to exhibit genuine indefinite causal order in a device-independent way [13, 23, 40].

The quantum SWITCH has been proven to provide advantages for a wide range of information-processing tasks, when compared to analogous situations that assume a definite causal order, as described at the start of this section. In the context of quantum Shannon theory, the combination of two communication channels via the quantum SWITCH was first shown to offer communication advantages over standard quantum Shannon theory, where channels can only be combined in parallel and sequence, in Ref. [80]. This work showed that two completely depolarising channels, which ordinarily transform any input state into the maximally mixed state I/d and therefore have zero capacity, can be activated by the quantum SWITCH to transmit classical information. That is, the quantum SWITCH transformation of two completely depolarising channels results in a new quantum channel with non-zero classical capacity. The lower bound was found from the Holevo capacity to be 0.049 bits per channel use, which was later shown to be equal to the classical capacity [59]. Soon afterwards, similar advantages were found in the communication of quantum information: two dephasing channels, which each have zero quantum capacity individually, were shown to be transformed by the quantum SWITCH into a quantum channel with non-zero quantum capacity [168], and later for a different choice of dephasing channels, into a quantum channel with *perfect* quantum capacity of 1 qubit per channel use [57].

Various proposals for experimentally implementing the quantum SWITCH have been made, using photonics [160], trapped ions [85] or superconducting qubits [86]. This has resulted in a series of papers claiming to have realised the quantum SWITCH experimentally using photonic systems [93, 160, 165], most recently also demonstrating the aforementioned communication advantages through noisy channels [94, 97]. However, whether these optical demonstrations are genuine implementations of an indefinite causal order or rather simulations of such processes has been a matter of continued debate [147, 152]. Implementations of the quantum SWITCH in a quantum gravity scenario involving the superposition of spacetimes have been proposed in Refs. [152, 203]. These gravitational implementations have been claimed to avoid the criticisms directed at the photonic experiments, however, it is not yet clear whether such experiments are possible to implement in practice.

In the rest of this work, we shall not be concerned further with the experimental implementability of indefinite causal order. We shall consider indefinite causal order as motivated from a purely foundational perspective. This is in contrast to the superposition of channels, described below, which has a natural physical implementation, and is motivated from both a foundational and practical perspective.

Soon after the publication of the first results on communication advantages associated with the quantum `SWITCH`, the origin of these advantages became the subject of critical analysis. In Ref. [2] it was noted that the possibility of classical communication through the quantum `SWITCH` of two completely depolarising channels could be reproduced by the coherent control of two such channels within a *definite* causal order. This has led to a resurgence of interest in the superposition, or *coherent control*, of quantum channels, within the research community that had been studying indefinite causal order [2, 76, 96, 133, 190]. In these works, several similarities and differences between the coherent control of independent channels and the coherent control of causal orders have been argued for.

In Chapter 4, we attempt to set the scene for a formal answer to this question, which is finally dealt with in Chapter 7. This leads us to the next section, where the literature on coherent control, or superpositions, of quantum channels and unknown quantum operations is systematically explored, followed by a discussion on the comparisons between the superposition of causal orders and superposition of channels.

2.5 Superposition, interference or coherent control of quantum channels

The idea of extending superpositions of quantum states to superpositions of quantum *evolutions* was first proposed by Aharonov, Anadan, Popescu and Vaidman in 1990 [9]. In the literature that followed over the last 30 years, this phenomenon has been called various names, including *gluing* of completely positive maps [4], *interference* of quantum channels [146], channel *multiplexing* [91] and *coherent control* of quantum channels [2]. In this thesis, we shall use these terms interchangeably, but mostly stick with the original notion of *superposition* of quantum channels.

Superpositions of quantum channels involve several subtleties in definition, namely their apparent dependence on the choice of Kraus representations of the channels, and the impossibility of controlling an unknown unitary. Each of these provides a problem for constructing superpositions of channels in communication and/or computation. In the following, we discuss these two issues in turn, followed by applications to communication and finally comparisons with the quantum SWITCH.

2.5.1 Superposition of channels depends on the implementation of the channels

A concrete physical example of the superposition of channels was studied by Åberg [3, 4] and Oi [146] in 2003, where two independent quantum channels were placed on separate arms of an interferometer. By routing a single particle through a superposition of the two paths of the interferometer, a superposition of the two channels is induced. It was shown that the overall channel arising from such a superposition depends not only on the two original channels themselves, but also on their specific implementation through an interaction with the surrounding environment. That is, the particular way in which a noisy channel is realised via its extension to a unitary acting on the system and an environment (Stinespring dilation – see §2.1) affects the way in which it can be combined in a superposition. This has the consequence that given access only to a standard input-output description of a set of quantum channels, there is no unique way in which these channels can be combined in a superposition.

Mathematically, the interferometric scenario can be described in the following way. We start with a Mach-Zehnder interferometer, with 50:50 beamsplitters on both ends. Let the quantum channels on the upper and lower arms of the interferometer be denoted $\mathcal{A}, \mathcal{B} \in \mathsf{Chan}(M)$, respectively, where M is an internal degree of freedom of the photon. Let U, V be unitaries which implement respective Stinespring dilations of the two channels, with respective environment systems E, F, initialised in the states $|\eta\rangle_E$, $|\varphi\rangle_F$, such that

$$\begin{aligned} \mathcal{A}(\rho) &= \mathrm{Tr}_E\left[U_{ME}\left(\rho\otimes|\eta\rangle\langle\eta|\right)U_{ME}^\dagger\right] \\ \mathcal{B}(\rho) &= \mathrm{Tr}_F\left[V_{MF}\left(\rho\otimes|\varphi\rangle\langle\varphi|\right)V_{MF}^\dagger\right] . \end{aligned} \tag{2.41}$$

We take the spatial path degree of freedom of the photon to be P, where $|0\rangle$ corresponds to the photon being localised on the upper arm and $|1\rangle$ corresponds to the photon being localised on the lower arm. Since U only acts on the upper arm, and V acts only on the lower arm, we have that the overall action of U and V is given by the unitary

$$W = \left(U_{ME} \otimes I_F \otimes |0\rangle\langle 0|_P\right) + \left(V_{MF} \otimes I_E \otimes |1\rangle\langle 1|_P\right) . \tag{2.42}$$

We start the protocol by initialising a single photon with an internal degree of freedom M (e.g. polarisation) in the pure state $|\psi\rangle \in \mathcal{H}_M$. Upon application of the first beamsplitter, the state of the system is

$$|\psi\rangle_M \otimes |+\rangle_P . \tag{2.43}$$

After passing through the two channels, the state of the whole system is

$$\frac{U_{ME}(|\psi\rangle_M \otimes |\eta\rangle_E) \otimes |\varphi\rangle_F \otimes |0\rangle_P}{\sqrt{2}} + \frac{V_{MF}(|\psi\rangle_M \otimes |\varphi\rangle_F) \otimes |\eta\rangle_E \otimes |1\rangle_P}{\sqrt{2}} . \tag{2.44}$$

From the above, a superposition of the channels $\mathcal{A}$ and $\mathcal{B}$, specified by the Stinespring dilations $(U, |\eta\rangle)$ and $(V, |\varphi\rangle)$, can be defined as the channel

$$\mathcal{S}(\rho_M \otimes \omega_P) := \mathrm{Tr}_{EF}\left[W(\rho_M \otimes |\eta\rangle\langle\eta|_E \otimes |\varphi\rangle\langle\varphi|_F \otimes \omega_P)W^\dagger\right] , \tag{2.45}$$

taking as input a state ρ_M on the internal degree of freedom and a state ω_P on the path ($\omega_P \neq |+\rangle\langle+|$ is achieved by having a non 50:50 initial beamsplitter). Indeed, this construction of a superposition of two channels depends not only on the channels themselves, but also on the choice of unitary extension.

In the context of communication, it is important to be able to build any communication protocol from a set of given resources. In quantum Shannon theory, the available resources are communication devices described by quantum channels, specified by their input-output descriptions, but with the particular realisation through the environment unknown. This means that without further specification, a superposition of quantum channels is problematic in a theory of communication.

Even if perfect access to the environment is assumed, a further issue remains. Not only should the communication protocols be constructible from a given set of devices, but also from a given set of *operations* on those devices. The same issue holds for computation, where algorithms are constructed by combining a given set of gates into a circuit. Mathematically, this means that the operations that can be performed on quantum channels should be valid quantum supermaps [52], i.e. (in the case of definite causal order) described by a quantum comb with open slots that can take any quantum channels as input.

Assuming perfect access to the environment is equivalent to having access to unitary channels. Indeed, this particular issue was first presented in the context of computation with control over unknown unitary gates [12, 49, 85, 184, 188, 202]. The problem is that it is impossible to construct a supermap that is able to perform a universal controlled operation on any unknown unitary. Yet, given some additional information about the device, this has been shown to be possible. Universal control over unknown unitaries is presented next.

2.5.2 Controlling an unknown unitary

Controlled-unitary gates are a fundamental element of practically every protocol in quantum computation. They act on a target quantum system and a control quantum system, where the value of the control system determines whether or not a given unitary V is applied on the target system. A simple example is the CNOT gate

$$U_{\texttt{CNOT}} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X\,, \tag{2.46}$$

which acts as the identity on the target if the control is in state $|0\rangle$, and acts as the NOT gate X on the target if the control is in state $|1\rangle$. Here the unitary that is being controlled is the X gate. In the absence of such control, the X gate is defined only up to a global phase [145]. However, any controlled operation on the X gate enforces a particular phase to be chosen, in the case of the CNOT, simply the number 1. This means that the definition of any controlled-unitary gate depends not only on the unitary that is to be controlled, but also on a choice of phase. In the case of the X gate, any controlled-gate of the form

$$U_{X,\theta} = |0\rangle\langle 0| \otimes I + e^{i\theta}\,|1\rangle\langle 1| \otimes X, \quad \theta \in [0, 2\pi[\,, \tag{2.47}$$

is a valid gate that coherently controls X.

Consider now that we would like to have a procedure that universally constructs a controlled-unitary gate from any unknown unitary V that is given as a black box. Since the unitary is unknown, what we really have access to is the unitary $Ve^{i\varphi}$ for some unknown phase $\varphi \in [0, 2\pi[$. Given that the unitary itself has an unknown phase, it is impossible to universally choose a particular phase with which to control any such unitary. This simple analysis reveals that the coherent control of an unknown unitary is impossible in the standard circuit model. This is formally proven in Refs. [12, 184].

Nevertheless, if certain additional information about the devices modelled as unitary gates is known, then the no-go theorem for controlling an unknown unitary can be circumvented [85, 86, 188, 202]. This can be done, for example, by extending the Hilbert space of possible states to include a vacuum sector $|\text{vac}\rangle$, which is orthogonal to all other information-carrying states [3, 202], and is guaranteed to be preserved under the unknown unitaries.

For example, given a unitary gate $V \in \mathsf{Chan}(A)$ on some system A, one could define a new unitary gate

$$U := e^{i\varphi} V_A \oplus I_{\text{Vac}} \in \mathsf{Chan}(A \oplus \text{Vac}) \,, \tag{2.48}$$

where Vac is a one-dimensional vacuum system. Since the phase $e^{i\varphi}$ is now a relative phase with respect to the additional vacuum system, this phase becomes part of the specification of the extended unitary U. If every (unknown) unitary is guaranteed to decompose in this block-diagonal form, then a supermap can be defined to universally control an unknown unitary V with phase φ with the vacuum:

$$\begin{aligned} \mathcal{S}_{\text{control}}(\mathcal{V}, \varphi) &= S(\cdot)S^\dagger \\ S &:= |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Ve^{i\varphi} \,. \end{aligned} \tag{2.49}$$

The abstract notion of vacuum is directly inspired by the vacuum state in quantum optics, where such protocols have indeed been performed in experiment [125, 202].

2.5.3 Communication through a superposition of paths

The first concrete proposal for communication through two quantum channels combined in a superposition of paths was presented by Gisan, Linden, Massar and Popescu in 2005 [91]. In this work, it was shown that when a single particle is sent through a superposition of paths going through two different noisy channels, the overall amount of noise experienced by the particle is reduced, compared to going through only one of the two channels. This idea of *error filtration* has far-reaching implications for the possibility of communication through quantum noise. However, in the cases considered, particular forms of noise were always assumed, and the results depend on the particular physical implementation of the superposition. Thus, the aforementioned desiderata that any communication protocols should be constructible from only a given set of communication channels and fixed operations on those channels are not (yet) satisfied.

This leaves an important problem to be solved, namely, to construct a Shannon-theoretic framework for communication through a superposition of paths, where existing communication devices can be coherently controlled in a well-defined manner. A solution to this problem is the starting point of the results presented in Chapters 3, 6 and 8 of this thesis. In Chapter 5, we then present various Shannon-theoretic communication advantages that can be quantified from our definition of the superposition of channels.

We note that related communication schemes have been proposed using devices that act coherently with the vacuum, such as two-way classical communication using a single particle [72] and communication through multiple-access channels with a single particle [200]. However, a fundamental difference between these two frameworks and the superposition of channels is that there is no internal degree of freedom of the particle in those frameworks. These frameworks and their relationship to the work presented in this thesis are further discussed in Chapter 10.

2.5.4 Superposition of paths vs. superposition of causal orders

The framework of error filtration is similar in spirit to the results on communication enhancements from combining two quantum channels in a superposition of causal orders. In the former, the communication channels are combined through a superposition of paths, while in the latter, the communication channels are combined through a superposition of orders.

The formal similarities between the two paradigms led to the 2018 work of Abbott et al. [2], where the superposition of orders and superposition of paths is compared. In this work, the authors showed that the specific communication advantages arising from combining two noisy channels via the quantum `SWITCH`, presented in Ref. [80], could be reproduced or even surpassed by combining the noisy channels in a superposition of paths, as per the construction (2.45) of Oi [146]. Specifically, combining two completely depolarising channels on the two arms of an interferometer, with specific choices of Stinespring dilations, was shown to result in an overall channel with a Holevo capacity up 0.16 bits per channel use, which is greater than the value 0.049 bits per channel use from the quantum `SWITCH`. The authors argued that this suggests it is the common element of *coherent control*, rather than indefinite causal order, which gives rise to the communication advantages associated with the quantum `SWITCH`. A key difference, however, was noted, in that a superposition of causal orders depends only on the input-output description of the communication channels themselves, whilst a superposition of channels on alternative paths requires the additional specification of a choice of unitary dilation. This comparison was further elaborated in the 2018 work of Guérin, Rubino and Brukner [96], where a set of operations which can be considered superpositions of processes were compared with the quantum `SWITCH`.

Chapter 4 of this thesis is devoted to constructing a framework to enable the communication advantages associated with the two types of superpositions to be formally quantified. We do this by constructing a *resource theory* of communication. With this in hand, we argue in Chapter 7, contrary to Refs. [2, 96], that the communication advantages arising from the quantum `SWITCH` and superposition of paths are distinct to a certain extent, and, due to their dependence on different initial resources, cannot be readily compared on an equal footing. Resource theories are introduced in the next section.

2.6 Quantum resource theories

In this section, we introduce the notion of resource theories, and their application to quantum information. In this work, we use the framework of resource theories to describe communication. To motivate this framework, we begin by illustrating simple examples of resource theories in other fields, and then qualitatively describe the general framework in the context of quantum information, and specifically quantum communication. A quantitative description is left for Chapter 4.

In many areas of science, the concept of *resources* plays a fundamental role. For example, in industrial chemistry, an important problem is how to convert abundant raw materials into useful products. The resources come in two levels: first, there are the 'objects' that are available (e.g. the chemicals abundant in nature), and secondly there are 'transformations' which can be performed on the objects (e.g. chemical reactions). Objects which are abundant are considered 'free,' while those which require effort to produce are considered 'costly'. Similarly, transformations which are in some sense easy to perform are considered free. The relation between the objects and transformations is that costly objects are those which cannot be generated from free objects and free operations alone [60, 67]. For example, a pharmaceutical drug is costly as it requires a huge amount of processing to make from abundant raw materials.

In quantum information, the most prominent example of a resource theory is the resource theory of bipartite entanglement. In this theory, the 'objects' are quantum states and the 'transformations' are quantum channels. Entangled states between two parties Alice and Bob are considered as a costly resource, which are used to enable quantum communication between the two parties. Indeed, entangled states are generally difficult to produce in the laboratory. Separable states between the two parties are considered free. The free operations are local operations and classical communication (LOCC). Given a set of separable states between two parties, LOCC cannot generate an entangled state between the two parties [67, 196].

Formally, a resource theory can be defined in the language of category theory [27, 63, 66], following the work of Coecke, Fritz and Spekkens [67]. A resource theory is defined as a symmetric monoidal category, where the 'objects' are *objects* in the category and the 'transformations' are *morphisms* between the objects. The categorical structure is natively equipped with a parallel composition of objects and transformations, as well as a sequential composition of transformations. Other formal characterisations of resource theories have also been made [60]. A notable distinction between different approaches is that in the categorical formulation, a set of free transformations is defined first, and the set of free objects is defined with respect to the set of free transformations. Conversely, other works have first defined a set of free objects, and then defined free transformations as those which preserve the set of free objects [131, 132].

In this thesis, we broadly follow the categorical constructions, but do not directly employ any category theory, which will not be required for understanding any of the work presented in the thesis. For our purposes, it is sufficient to define a resource theory as follows: the set of all possible resources is a set of objects. These objects are equipped with a set of operations M that can act on them, which are closed under sequential and parallel composition. A subset of operations $\mathsf{M}_{\text{free}} \subseteq \mathsf{M}$, which are regarded as *free*, is chosen.

In typical applications of resource theories to quantum theory, the objects have been quantum states and the transformations have been quantum channels. However, in a theory of communication, the given resources are communication devices described by quantum channels, which then admit transformations such as encoding, decoding and other operations, described by quantum supermaps. The first application of a resource-theoretic framework to quantum Shannon theory was by Devetak, Harrow, and Winter in 2008 [74], which aimed to unify the underlying structures between the various coding theorems in quantum Shannon theory. In Chapter 4, we present a resource theory of communication which captures different choices of possible combinations of communication channels, including the superposition of causal orders and the superposition of quantum channels.

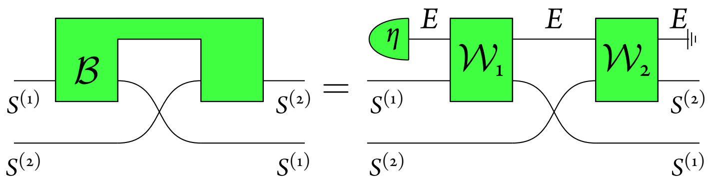


Figure 2.7: *A two-step quantum channel.* The left-hand side depicts a 2-step quantum channel $\mathcal{B}$ taking two input states on systems $S^{(1)}$ and $S^{(2)}$, in succession. The right-hand side shows the physical implementation of the 2-step channel via two unitary channels $\mathcal{W}_1$ and $\mathcal{W}_2$ [51, 54] where the memory between the two uses of the channel is realised by an environment E, which is inaccessible to the communicating parties. The ground symbol on a wire denotes the partial trace over the corresponding system. Note that the two unitary channels and the state η are all drawn in the same colour because they represent the same communication channel.

2.7 Quantum channels with correlated noise

2.7.1 Time-correlated processes: k-step quantum channels

Until now, we have always described the single use of a given quantum device, such as a computational gate or transmission line, by a single quantum channel. This could be a quantum channel with a single input/output port, or a k-partite quantum channel with k input/output ports, interpreted as acting simultaneously on k spatially separate subsystems. Now, we introduce a new possibility: we consider a quantum device that can act multiple times in succession, with possible correlations between successive uses. In this case, the quantum device is described by a k-partite quantum channel with k input/output ports, where each different port corresponds to a use of the device at a different time.

To illustrate this idea, consider a quantum device described by a random unitary gate chosen from a given set $\{V_m\}$, e.g. the set of Pauli gates, with some probability distribution $\{p(m)\}$. That is, each time this device is used, it is described by the quantum channel

$$\mathcal{R}_1(\rho) = \sum_m p_1(m) V_m \rho V_m^\dagger \in \mathsf{Chan}(S^{(1)})\,. \tag{2.50}$$

However, the specific choice of gate in any two consecutive applications is correlated. Specifically, two uses of the device at two time steps in succession, t_1 and t_2, is described by the bipartite quantum channel

$$\mathcal{R}(\rho_{12}) = \sum_{m,n} p(m,n)\,(V_m \otimes V_n)\,\rho_{12}\,(V_m \otimes V_n)^\dagger \in \mathsf{Chan}(S^{(1)} \otimes S^{(2)})\,, \tag{2.51}$$

where V_m and V_n are unitary gates in the given set and $p(m,n)$ is a joint probability distribution, with its marginal given by $p(m)$. Here, the system $S^{(1)}$ sent at time t_1 experiences the unitary gate V_m, while the system $S^{(1)}$ sent at time t_2 experiences the gate V_n. The quantum state ρ_{12} represents the joint state of the two systems sent at the two times t_1 and t_2, that is, ρ_{12} is a quantum state on the Hilbert space of the composite system $S^{(1)} \otimes S^{(2)}$. (Note, that in a quantum circuit, the description of the joint system sent at the two times using the tensor product $S^{(1)} \otimes S^{(2)}$ is justified by the wiring depicted in Figure 2.7. That is, we can start by considering two systems $S^{(1)}$ and $S^{(2)}$ in parallel, with $S^{(1)}$ sent through the input of the first port of the 2-step channel, whilst the system $S^{(2)}$ undergoes an identity channel. Then, the output of the first port of the 2-step channel is swapped with the output of the identity channel, with the latter then entering the input of the second port of the 2-step channel, and the former now undergoing an identity channel.) The probability distribution $p(m,n)$ specifies the correlations between the random unitary evolutions experienced by system $S^{(1)}$ and system $S^{(2)}$.

Generalising from the example of a bipartite random unitary channel, we can construct a general definition of time-correlated channels. A quantum device that can be used k times in succession is described by a *k-step (correlated) quantum channel* [135] (also known as a quantum k-comb [51, 54], quantum memory channel [39, 43, 117], or non-Markovian quantum process [41, 157]).

Formally, a *k-step quantum channel* is a special type of k-partite quantum channel $\mathcal{C}$ with the additional property that no signal propagates from an input $S^{(i)}$ to any group of outputs $S'^{(j)}$ with $j < i$ [51]. In this case, $S^{(j)}$ denotes the system sent at the j-th time. A k-step quantum channel is mathematically equivalent to a quantum k-comb (§2.3.2), although here it describes a device rather than a higher-order transformation as in §2.3.2. We shall denote the set of k-step quantum channels as $\mathsf{Chan}(S^{(1)} \to S'^{(1)}, \ldots, S^{(k)} \to S'^{(k)})$, or simply $\mathsf{Chan}(S^{(1)}, \ldots, S^{(k)})$ when the input and output of each pair coincide. When k is unspecified, we shall refer to these channels as *multi-step quantum channels*. For $k = 2$, an example of a 2-step quantum channel is illustrated in Figure 2.7.

Physically, a time-correlated random unitary channel of the form (2.51) can arise in a photonic setup where the systems $S^{(1)}$ and $S^{(2)}$ are modes of the electromagnetic field associated with two different time bins [77, 78, 111, 129]. The noisy channel can correspond, for example, to the action of an optical fibre, where the random unitary changes of the photon polarisation arise from random fluctuations in the birefringence. Correlations between the unitaries at different times can arise when the time difference $t_2 - t_1$ between successive uses of the channels is smaller than the timescale on which the birefringence fluctuates.

2.7.2 Spatially correlated processes: no-signalling quantum channels

The k-step quantum channels described in the previous subsection are a special case of k-partite quantum channels that represent processes where successive output systems are unable to causally influence previous input systems. Physically, such a restriction naturally occurs when successive pairs of input/output systems correspond to actions performed by the same party at successive moments in time.

Another special case of k-partite quantum channels are k-partite no-signalling quantum channels, which represent processes where each output system depends only on its corresponding input system and no others.

Formally, a *k-partite no-signalling quantum channel* is a special type of k-partite quantum channel $\mathcal{C}$ with the additional property that no signal propagates from an input $S^{(i)}$ to any group of outputs $S'^{(j)}$ with $j \neq i$ [57]. In the following, we denote the set of all k-partite no-signalling quantum channels by $\mathsf{NSChan}(S^{(1)} \to S'^{(1)}, \ldots, S^{(k)} \to S'^{(k)})$. When the inputs and outputs are arbitrary, we use the notation NSChan.

Physically, no-signalling channels naturally occur in the context of multi-partite channels, where each pair of input/output systems corresponds to causally disconnected parties.

Note, that no-signalling quantum channels are a special case of multi-step channels. In fact, the correlations in the example of Eq. (2.51) describe a no-signalling channel, and are not specific to time. The same expression can be used also to describe correlated channels acting on two spatially separated systems, or on any other type of independently addressable systems.

2.7.3 Communication advantages arising from time-correlations

Let us briefly examine an example of a communication protocol which makes use of correlated quantum channels. Consider a noisy transmission line which exhibits time-correlations in noise. With knowledge of the structure of the correlations, it can be possible for the sender to encode the message over a sequence of particles such that the correlations cancel out the noise overall [135].

The completely depolarising channel $\mathcal{N}_{\text{dep}} = \text{Tr}(\rho) I/d$ sends any input state to the maximally mixed state I/d, where d is the dimension of the output system, and hence cannot be used to transmit information. Now consider a transmission line, where each use in isolation is described by the completely depolarising channel, but where the noisy processes occurring on successive uses are correlated. Physically, the completely depolarising channel in dimension $d = 2$ can be realised as a uniform randomisation over the identity and Pauli channels, as in Eq. (2.50) with $U_0 = I, U_1 = X, U_2 = Y, U_3 = Z$ and $p(m) = 1/4$. If the time interval between successive applications of the device is small, then the application of the device on two consecutive particles, described by a (possibly entangled) quantum state ρ_{12}, could result in the choice of unitary being perfectly correlated. This results in the 2-step quantum channel (2.51), with

$$p(m,n) = \frac{\delta(m,n)}{d^4} \,. \tag{2.52}$$

In Ref. [135], Macchiavello and Palma showed that by encoding ρ_{12} as one of the four Bell states, it is possible to achieve perfect classical communication of 2 bits per use of the correlated channel. This example highlights the potential of more general communication scenarios than the traditional direct communication from a sender to a receiver using a single uncorrelated transmission line.

In Chapter 6, we extend the use of correlated quantum channels in communication to scenarios where a single particle is sent at a superposition of different times, in the spirit of the superpositions of channels discussed in §2.5.1.

3
Superposition of channels

In this chapter, we formally define a general notion of a superposition of two quantum channels, which is consistent with previous examples [2–4, 146]. We provide a characterisation of all possible superpositions of two channels in terms of both their Kraus representations and Hamiltonian realisations. We discuss how this characterisation clearly shows the non-uniqueness of a superposition of two channels, which is problematic in the context of communication, as discussed in §2.5.1. We circumvent these problems by defining the notion of a vacuum extension of a quantum channel, which enables the construction of supermap that generates a superposition of two independent channels. Finally, we extend our construction to a superposition of multiple independent channels, as well as to a superposition of multipartite channels. First, however, we introduce some important concepts relating to the action of channels on systems and sectors.

3.1 Systems and Sectors

In quantum information theory, the degrees of freedom used to carry information are represented by abstract quantum systems. However, in practice, an abstract quantum system A is only an effective description of some subset of degrees of freedom, which is accessible to an experimenter in a given spacetime region [48, 193, 199]. In this case, the Hilbert space $\mathcal{H}_A$ of the abstract quantum system A is a subspace of a larger Hilbert space $\mathcal{H}_S$, which describes all the degrees of freedom that could be accessed in principle. In this case, A is called a sector of S, and the states of system A satisfy the constraint

$$\mathrm{Tr}[\rho P_A] = 1 \qquad \forall \rho \in \mathsf{St}(A)\,, \tag{3.1}$$

where P_A is the projector onto $\mathcal{H}_A$.

For example, in quantum optics, a polarisation qubit is defined by the two orthogonal states $|1\rangle_{\mathbf{k},H} \otimes |0\rangle_{\mathbf{k},V}$ and $|0\rangle_{\mathbf{k},H} \otimes |1\rangle_{\mathbf{k},V}$, corresponding to a single photon with wavevector $\mathbf{k}$ in either the horizontal (H) or vertical (V) polarisation modes (here, $|n\rangle_X$ is a particle number state with n particles in mode X). Yet, this description only holds if the state of the electromagnetic field is constrained to lie within the subspace spanned by these two state vectors.

In general, the evolution of the larger system S is described by a quantum channel $\widetilde{\mathcal{C}} \in \mathsf{Chan}(S)$. The channel $\widetilde{\mathcal{C}}$ defines an effective evolution on sector A only if it maps all states in A onto states in A, that is, if it satisfies the *No Leakage Condition*

$$\mathrm{Tr}\left[P_A\,\widetilde{\mathcal{C}}(\rho)\right] = 1 \qquad \forall \rho \in \mathsf{St}(A)\,. \tag{3.2}$$

When equation (3.2) is satisfied, one can always define an effective channel $\mathcal{C} \in \mathsf{Chan}(A)$, simply given by

$$\mathcal{C}(\rho) := \widetilde{\mathcal{C}}(\rho) \qquad \forall \rho \in \mathsf{St}(A)\,. \tag{3.3}$$

In this case, we call $\mathcal{C}$ the *restriction* of $\widetilde{\mathcal{C}}$ to sector A and say that $\widetilde{\mathcal{C}}$ is an *extension* of $\mathcal{C}$. Clearly, given a channel $\widetilde{\mathcal{C}}$, $\mathcal{C}$ is uniquely defined. However, a given channel $\mathcal{C}$ does not have a uniquely determined extension $\widetilde{\mathcal{C}}$.

The relationship between a given channel $\widetilde{\mathcal{C}}$ and its restriction can be written in terms of the Kraus operators:

Lemma 1. *Let $\widetilde{\mathcal{C}}(\rho) = \sum_{i=1}^{r} \widetilde{C}_i \rho \widetilde{C}_i^{\dagger}$ be a Kraus representation of channel $\widetilde{\mathcal{C}} \in \mathsf{Chan}(S)$. Then, channel $\widetilde{\mathcal{C}}$ satisfies the No Leakage Condition with respect to sector A if and only if*

$$P_A \widetilde{C}_i P_A = \widetilde{C}_i P_A \qquad \forall i \in \{1, \ldots, r\} . \tag{3.4}$$

The proof is given in Appendix A.1.

Proposition 2. *Let $\widetilde{\mathcal{C}}(\rho) = \sum_{i=1}^{r} \widetilde{C}_i \rho \widetilde{C}_i^{\dagger}$ be a Kraus representation of a channel $\widetilde{\mathcal{C}} \in \mathsf{Chan}(S)$, which satisfies the No Leakage Condition with respect to sector A. Let $\mathcal{C}$ be the restriction of $\widetilde{\mathcal{C}}$ to sector A. Then a Kraus representation of $\mathcal{C}$ is*

$$\{C_i := \widetilde{C}_i P_A\}_{i=1}^{r} . \tag{3.5}$$

Proof. Follows directly from the definition of a restriction and Lemma 1. □

3.2 General superposition of quantum channels

Consider two abstract quantum systems A and B. If the Hilbert spaces corresponding to the systems A and B are orthogonal subspaces of a larger Hilbert space $\mathcal{H}_S$, we can construct a new system $S := A \oplus B$, corresponding to the direct sum Hilbert space $\mathcal{H}_S = \mathcal{H}_A \oplus \mathcal{H}_B$. Physically, the system $A \oplus B$ represents a quantum system that can be in sector A, in sector B, or in a coherent superposition of the two sectors.

The possible evolutions of system S given the evolutions of its sectors A and B, motivates a general definition of a superposition of channels:

Definition 3. *A superposition of two channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$ is any channel $\mathcal{S} \in \mathsf{Chan}(A \oplus B)$ such that*

$$\begin{aligned} \mathcal{S}(\rho) = \mathcal{A}(\rho) \qquad \forall \rho \in \mathsf{St}(A) , \\ \mathcal{S}(\rho) = \mathcal{B}(\rho) \qquad \forall \rho \in \mathsf{St}(B) . \end{aligned} \tag{3.6}$$

Physically, $\mathcal{S}$ represents an evolution that can take an input state in sector A, in sector B, or in a coherent superposition of the two sectors, and returns an output state in the corresponding sector.

Clearly, two channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$ do not define a unique superposition channel. To make this point, consider the example of a superposition of two unitary channels

$\mathcal{U}(\cdot) = U(\cdot)U^\dagger$ and $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$. One possible superposition is given by the unitary channel $\mathcal{S}(\cdot) = S(\cdot)S^\dagger$, with $S = U \oplus V$. Another possible superposition is given by the non-unitary channel $\mathcal{S}'(\cdot) = S'_1(\cdot)S'^\dagger_1 + S'_2(\cdot)S'^\dagger_2$, where $S'_1 = U \oplus 0_B$ and $S'_2 = 0_A \oplus V$, where 0_X denotes the null operator on system X. The latter example could be implemented in practice by first performing a non-demolition measurement that distinguishes the two sectors A and B whilst preserving coherence within each individual sector, followed by performing either the unitary channel $\mathcal{U}$ or $\mathcal{V}$ depending on the outcome.

Given two quantum channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$, with some Kraus representations $\{A_i\}_{i=1}^r$ and $\{B_i\}_{i=1}^r$ with the same number of Kraus operators, we can construct a superposition channel $\mathcal{S} \in \mathsf{Chan}(A \oplus B)$ specified by the Kraus operators

$$S_i := A_i \oplus B_i \qquad i \in \{1, \dots, r\}\,. \tag{3.7}$$

Note that two Kraus representations $\{A_i\}_{i=1}^{r_A}$ and $\{B_i\}_{i=1}^{r_B}$ with different numbers of Kraus operators can always be extended to Kraus representations with the same number of operators, for example, by appending null operators.

In fact, any superposition of channels can be formed with the above construction, as specified by the following theorem, which also provides a possible physical realisation of any superposition of channels:

Theorem 4. *The following are equivalent:*

1. *Channel $\mathcal{S} \in \mathsf{Chan}(A \oplus B)$ is a superposition of channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$.*

2. *The Kraus operators of $\mathcal{S}$ are of the form $S_i = A_i \oplus B_i$ for* some *Kraus representations $\{A_i\}$ and $\{B_i\}$ of channels $\mathcal{A}$ and $\mathcal{B}$, respectively.*

3. *There exists an environment E, a pure state $|\eta\rangle \in \mathcal{H}_E$, two Hamiltonians H_{AE} and H_{BE}, with supports in the orthogonal subspaces $\mathcal{H}_A \otimes \mathcal{H}_E$ and $\mathcal{H}_B \otimes \mathcal{H}_E$, respectively, and an interaction time T, such that $\mathcal{A}(\rho) = \mathrm{Tr}_E\left[U_{AE}(\rho \otimes \eta)U_{AE}^\dagger\right]$ with $U_{AE} = \exp\left[-iH_{AE}T/\hbar\right]$, and $\mathcal{B}(\rho) = \mathrm{Tr}_E\left[U_{BE}(\rho \otimes \eta)U_{BE}^\dagger\right]$ with $U_{BE} = \exp\left[-iH_{BE}T/\hbar\right]$, and $\mathcal{S}(\rho) = \mathrm{Tr}_E\left[U(\rho \otimes \eta)U^\dagger\right]$ with $U = \exp\left[-i(H_{AE} \oplus H_{BE})T/\hbar\right]$, where $\eta := |\eta\rangle\langle\eta|$.*

The proof is given in Appendix A.1.

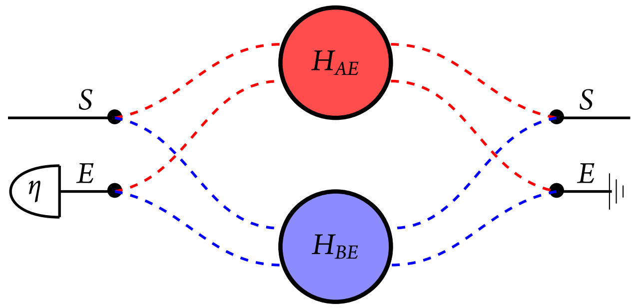


Figure 3.1: *Hamiltonian realisation of an arbitrary superposition of channels.* The system S and the environment E are jointly routed to one of two regions, R_A and R_B, depending on whether the system is in sector A or sector B. In region R_A (R_B) the system and the environment interact through a Hamiltonian H_{AE} (H_{BE}). After the interaction, the two paths of both the system and the environment are recombined, and finally the environment is discarded. In general, this realisation requires the ability to control the environment.

Condition 3 shows how any superposition of two channels can be implemented in practice: The system and the environment are jointly routed to two separate regions, R_A and R_B, depending on whether the system is in sector A or sector B. Local Hamiltonians H_{AE} and H_{BE} act jointly on the system and environment in each region. Finally, the two alternative trajectories are recombined and the environment is discarded. This is illustrated in Figure 3.1.

Theorem 4 highlights a fundamental problem with the superposition of channels from an operational point of view. In information processing or communication, one would like any combination of devices to depend on only the original devices themselves. That is, given two devices described by quantum channels $\mathcal{A}$ and $\mathcal{B}$, one would like to be able to construct the superposition of those channels given only the knowledge of the channels themselves. Mathematically, this amounts to saying that the superposition of two channels should be a quantum supermap (cf. §2.3.2). However, it is clear that the characterisation of the superpositions of channels in Theorem 4 does not uniquely specify the action of a superposition channel $\mathcal{S}$ given two constituent channels $\mathcal{A}$ and $\mathcal{B}$. Physically, this non-uniqueness can be attributed to two separate phenomena, as previously discussed in §2.5.1–2.5.2.

First, the dependence of the superposition of two noisy channels on their Kraus representations $\{A_i\}_{i=1}^r$ and $\{B_i\}_{i=1}^r$ signifies a dependence on the way in which the channels are realised via an interaction with their environment (cf. §2.5.1) [2, 146]. This is a problem in the context of information processing and communication because an agent with access to a noisy quantum channel typically does not have access to its environment (otherwise, the description of the device would simply reduce to a larger unitary channel). Second, even in the case of unitary channels, the universal control of unitary channels is impossible, since their relative phase is not uniquely determined (cf. §2.5.2) [12, 184, 188]. This further prevents the possibility of a scenario where agents have access to a given set of communication devices and can freely combine them in superpositions.

In the next sections, we propose a framework for constructing a superposition of quantum channels that depends only on an input-output description of the available information-processing devices. We circumvent the above problems by using extensions of the original channels as the initial descriptions of the devices.

3.3 The vacuum extension of a quantum channel

In information processing, each use of a given device is counted as a resource. For example, a transmission line that can be used only once is modelled as a single quantum channel, whilst a transmission line that can be used twice independently is modelled as two quantum channels in parallel (cf. §2.1) [2, 146]. However, the physical apparatus performing the channel exists also in the absence of processing or transmitting information. For example, consider an optical fibre. When the optical fibre transmits a single photon with variable polarisation, we model it as a quantum channel $\mathcal{A} \in \mathsf{Chan}(A)$, where A is a two-dimensional system consisting of the horizontal and vertical polarisation degrees of freedom of a single photon. When the fibre is not used, we can model it as a quantum channel acting on a system containing the vacuum of the electromagnetic field [202]. In general, we assume that the system used for information processing or communicating is a sector A of a larger system S, which also contains a *vacuum sector* Vac, which is orthogonal to A.

Thus, we model a communication device as a quantum channel on the direct sum system $\widetilde{A} := A \oplus \mathrm{Vac}$. For simplicity, we shall assume that the vacuum sector is one-dimensional with a unique vacuum state $|\mathrm{vac}\rangle$. We call this extended description a *vacuum extension* of the original channel:

Definition 5. *Channel* $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{A})$ *is a* vacuum extension *of channel* $\mathcal{A} \in \mathsf{Chan}(A)$ *if*

$$\begin{aligned} \widetilde{\mathcal{A}}(\rho) &= \mathcal{A}(\rho) \qquad &\forall \rho \in \mathsf{St}(A)\,, \\ \widetilde{\mathcal{A}}(\rho) &= \mathcal{I}_{\mathrm{Vac}}(\rho) \qquad &\forall \rho \in \mathsf{St}(\mathrm{Vac})\,, \end{aligned} \tag{3.8}$$

where $\mathcal{I}_{\mathrm{Vac}}$ *is the identity channel on* Vac.

Theorem 4 implies that the Kraus operators of the vacuum extension $\widetilde{\mathcal{A}}$ have the form

$$\widetilde{A}_i = A_i \oplus \alpha_i \left|\text{vac}\right\rangle\left\langle\text{vac}\right| , \tag{3.9}$$

where $\{A_i\}_{i=1}^r$ is a Kraus representation of $\mathcal{A}$ and $\{\alpha_i\}_{i=1}^r$ are complex numbers, which we call *vacuum amplitudes*, satisfying the normalisation condition

$$\sum_i |\alpha_i|^2 = 1 . \tag{3.10}$$

The vacuum extension is an example of a superposition of channels as per Definition 3, and is therefore non-unique in general. However, the choice of vacuum extension is uniquely determined by the physics (i.e. Hamiltonian) of the device. The vacuum extension is the complete description of the computation or communication resource available to an agent. Crucially, it can be determined experimentally by input-output tomography of the device, without access to the environment.

Considering the case of quantum optics, a unitary $\widetilde{V}$ can be realised by a Hamiltonian acting on the two polarisation modes associated to systems A and B. For example, the unitary $Z \oplus e^{i\varphi} \left|\text{vac}\right\rangle\left\langle\text{vac}\right|$ can be generated by the Hamiltonian $H = \hbar[(\xi + \theta/2)a_H^\dagger a_H + (\xi - \theta/2)a_V^\dagger a_V]$ for $\theta = \pi,\ \varphi = \xi + \theta/2$ and time $t = 1$ in suitable units, where a_H (a_V) are the annihilation operators for the appropriate modes with horizontal (vertical) polarisation.

The action of channel $\widetilde{\mathcal{A}}$ on a generic quantum state $\rho \in \mathsf{St}(\widetilde{A})$ is

$$\widetilde{\mathcal{A}}(\rho) = \mathcal{A}(P_A \rho P_A) + \left\langle\text{vac}\right|\rho\left|\text{vac}\right\rangle \left|\text{vac}\right\rangle\left\langle\text{vac}\right| + \mathrm{F}\,\rho \left|\text{vac}\right\rangle\left\langle\text{vac}\right| + \left|\text{vac}\right\rangle\left\langle\text{vac}\right| \rho\, F^\dagger , \tag{3.11}$$

where the operator

$$F := \sum_i \overline{\alpha}_i A_i , \tag{3.12}$$

is called the *vacuum interference operator*. Note that the operator F depends only on the channel $\widetilde{\mathcal{A}}$, and not on the choice of Kraus operators, as can be seen by comparing the two sides of Eq. (3.11).

If the vacuum interference operator is zero, then the output state (3.11) is an incoherent mixture of a state of system A and the vacuum.

Definition 6. *For $F = 0$, we say that the vacuum extension $\widetilde{\mathcal{A}}$* has no coherence with the vacuum*, and we call it the* incoherent vacuum extension *of channel $\mathcal{A}$. For $F \neq 0$, we say that the vacuum extension $\widetilde{\mathcal{A}}$* has coherence with the vacuum.

3.4 The superposition of two independent channels

Having defined the notion of a vacuum extension of a quantum channel, we can finally proceed to build a superposition of two channels that depends only on a description of the devices themselves. We start with two systems A and B, with $\mathcal{H}_A \simeq \mathcal{H}_B$, and construct the corresponding vacuum-extended systems $\widetilde{A} := A \oplus \mathrm{Vac}$ and $\widetilde{B} := B \oplus \mathrm{Vac}$. Consider now the composite system formed by taking the tensor product of the two vacuum-extended systems:

$$\begin{aligned} \widetilde{A} \otimes \widetilde{B} = &\ \mathrm{Vac} \otimes \mathrm{Vac} \\ &\oplus (A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes B) \\ &\oplus A \otimes B\,. \end{aligned} \tag{3.13}$$

This contains (1) a no-particle sector $\mathrm{Vac} \otimes \mathrm{Vac}$, (2) a one-particle sector $(A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes B)$, and (3) a two-particle sector $A \otimes B$. The one-particle sector is isomorphic to the direct sum $A \oplus B$, since Vac is one-dimensional. When $A \simeq B$, the direct sum $A \oplus B$ can be physically realised by a particle with an internal message degree of freedom $M \simeq A \simeq B$, and an external path degree of freedom P. If P is a qubit of alternative paths $|0\rangle_P$ and $|1\rangle_P$, then $M \otimes P \simeq A \oplus B$.

Consider two channels $\mathcal{A}$ and $\mathcal{B}$, with vacuum extensions $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$. We construct the product channel $\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{B}}$, which represents the independent action of $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$. We can now define a superposition of channels $\mathcal{A}$ and $\mathcal{B}$, specified by the vacuum extensions $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$, as the restriction of the product channel $\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{B}}$ to the one-particle sector $(A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes B)$. We restrict our attention to the case where $A \simeq B$. Formally:

Definition 7. *The superposition of channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$ specified by the vacuum extensions $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{A})$ and $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{B})$ is the channel*

$$\begin{aligned} &\mathcal{S} : \mathsf{Chan}(\widetilde{A}) \times \mathsf{Chan}(\widetilde{B}) \to \mathsf{Chan}(M \otimes P) \\ &\mathcal{S}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})(\rho \otimes \omega) = \mathcal{U}^\dagger \circ (\widetilde{\mathcal{A}} \otimes \widetilde{\mathcal{B}}) \circ \mathcal{U}\,(\rho \otimes \omega)\,, \end{aligned} \tag{3.14}$$

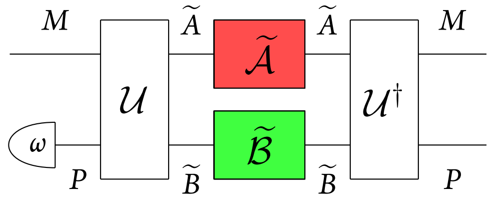


Figure 3.2: *The superposition of two independent channels $\mathcal{A}$ and $\mathcal{B}$ specified by the vacuum extensions $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$.* In communication scenarios, a single particle with internal message degree of freedom M is transmitted through the resulting superposition channel $\mathcal{S}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})$, with the path degree of freedom fixed in the state ω.

where the isomorphism

$$\begin{aligned} &\mathcal{U}(\cdot) = U(\cdot)U^\dagger \\ &U : \mathcal{H}_M \otimes \mathcal{H}_P \to (\mathcal{H}_A \otimes \mathcal{H}_{\text{Vac}}) \oplus (\mathcal{H}_{\text{Vac}} \otimes \mathcal{H}_B) \end{aligned} \tag{3.15}$$

is defined by

$$\begin{aligned} U(|\psi\rangle_M \otimes |0\rangle_P) &:= |\psi\rangle_{\widetilde{A}} \otimes |\text{vac}\rangle_{\widetilde{B}} \\ U(|\psi\rangle_M \otimes |1\rangle_P) &:= |\text{vac}\rangle_{\widetilde{A}} \otimes |\psi\rangle_{\widetilde{B}} \,. \end{aligned} \tag{3.16}$$

The superposition channel $\mathcal{S}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})$ is constructed from only the vacuum extensions $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$, which provide the full relevant descriptions of the two available devices. The map $\mathcal{S}$ is a quantum supermap (cf. 2.3.2) from a pair of vacuum-extended channels to a single bipartite quantum channel. The superposition of two channels specified by vacuum extensions is illustrated in Figure 3.2.

Physically, a superposition of two channels describes a single particle travelling in a superposition of two alternative trajectories, in each of which it undergoes one of the two alternative processes whilst the other process acts on the vacuum.

The following properties of the superpositions of two channels will be useful later. Explicitly, the action of the superposition of two channels specified by the vacuum extensions $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ is given by

$$\begin{aligned}\mathcal{S}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})(\rho\otimes\omega) = &\langle 0|\omega|0\rangle\,\mathcal{A}(\rho)\otimes|0\rangle\langle 0| + \langle 1|\omega|1\rangle\,\mathcal{B}(\rho)\otimes|1\rangle\langle 1| \\ &+\langle 0|\omega|1\rangle\,F_A\rho F_B^\dagger\otimes|0\rangle\langle 1| + \langle 1|\omega|0\rangle\,F_B\rho F_A^\dagger\otimes|1\rangle\langle 0|\,,\end{aligned} \tag{3.17}$$

where F_A (F_B) is the vacuum interference operator associated with the vacuum extension $\widetilde{\mathcal{A}}$ ($\widetilde{\mathcal{B}}$). The Kraus operators of the superposition channel $\mathcal{S}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})$ are

$$S_{ij} = A_{i\,M}\,\beta_j\otimes|0\rangle\langle 0|_P + \alpha_i\,B_{j\,M}\otimes|1\rangle\langle 1|_P\,, \tag{3.18}$$

where $A_{i\,M}, B_{j\,M}$ are Kraus operators corresponding to channels $\mathcal{A}$ and $\mathcal{B}$, respectively, and α_i and β_j are the associated vacuum amplitudes. When $\widetilde{\mathcal{A}} = \widetilde{\mathcal{B}}$, and the path is initialised in the $|+\rangle$ state, the above equation simplifies to

$$\mathcal{S}(\widetilde{\mathcal{A}},\widetilde{\mathcal{A}})(\rho\otimes|+\rangle\langle+|) = \frac{\mathcal{A}(\rho)+F\rho F^\dagger}{2}\otimes|+\rangle\langle+| + \frac{\mathcal{A}(\rho)-F\rho F^\dagger}{2}\otimes|-\rangle\langle-|\,, \tag{3.19}$$

where $F := \sum_i \overline{\alpha}_i\,A_i$.

3.4.1 Preview: physical implementation

A simple physical implementation of a superposition of two channels can be performed in the following way using single-photon quantum optics. The message is realised by the polarisation degree of freedom of a single photon, and the path is realised by the spatial degree of freedom of the same photon, in a Mach-Zehnder interferometer (see Figure 9.4 in §9.3). The horizontal and vertical polarisation modes are represented by the logical states $|0\rangle_M$ and $|1\rangle_M$ of the message, while localisation in the upper or lower arms of the interferometer are represented by the logical states $|0\rangle_P$ and $|1\rangle_P$ of the path.

The two devices corresponding to the two vacuum-extended channels are placed one on each arm of the Mach-Zehnder interferometer. A single photon is sent into one input port of the beamsplitter (with the vacuum in the other port). The state of the message $\rho\in\mathsf{St}(M)$ is controlled by half/quarter wave plates between the single-photon source and the first beamsplitter (i.e. prior to entering the interferometer). The state of the path $\omega\in\mathsf{St}(P)$ is determined by the parameters of the first beamsplitter.

After the two paths are recombined in the second beamsplitter (which is a 50:50 beamsplitter), finding a particle in one of the two outgoing modes corresponds to a measurement on the path in the $|\pm\rangle$ basis. A polarising beamsplitter (that splits a beam into its horizontal and vertical components) is placed at each of the two output ports of the interferometer. Then, a photon detector is placed at each of the two output ports of each of the two polarising beamsplitters, resulting in four photon detectors in total. A particle number measurement (of at most one particle) over these four photon detectors corresponds to a joint measurement of the message and path in the basis $\{|0\rangle_M \otimes |\pm\rangle_P, |1\rangle_M \otimes |\pm\rangle_P\}$. If additional half/quarter wave plates are placed between the polarising beamsplitters and photon detectors, then the output polarisation state can be rotated, corresponding to a measurement of the message in any chosen basis.

Further detail on the experimental implementations performed as part of this work is provided in Chapter 9.

3.4.2 Preview: communication using superposition of channels

Before continuing to formulate our framework for superpositions of channels, let us take a brief preview of its applications to communication, which are elaborated in Chapters 5,6 and 8. In a communication setting, the vacuum-extended channels describe transmission lines, with the input and outputs corresponding to the (spatially separate) locations of the sender and receiver, respectively.

From Eq. (3.19), we can already start to see how the superposition of independent channels is useful for reducing errors in communication. Consider the extreme case where $\mathcal{A}$ is channel with (classical or quantum) capacity equal to zero, such that no communication is possible through the channel on its own. Then, as long as $F \neq 0$ (i.e. $\widetilde{\mathcal{A}}$ has coherence with the vacuum), the output of the overall superposition channel (3.19) will have a non-trivial dependence on the input state ρ, enabling the possibility of (at least classical) communication at a non-zero rate.

Eq. (3.19) also suggests a natural decoding strategy for the receiver. By measuring in the Fourier basis $\{|+\rangle, |-\rangle\}$, the receiver can separate the two quantum operations $\mathcal{Q}_\pm := (\mathcal{A} \pm F \cdot F^\dagger)/2$. The $+$ outcome heralds constructive interference among the noisy processes along the two paths, while the $-$ outcome heralds destructive interference. This observation is the working principle of the error filtration technique of Gisin et al. [91], where selecting one of the two operations $\mathcal{Q}_\pm$ that are less noisy than the original channel enables a probabilistic reduction in noise. This advantage is quantified in a Shannon-theoretic manner in Chapter 5.

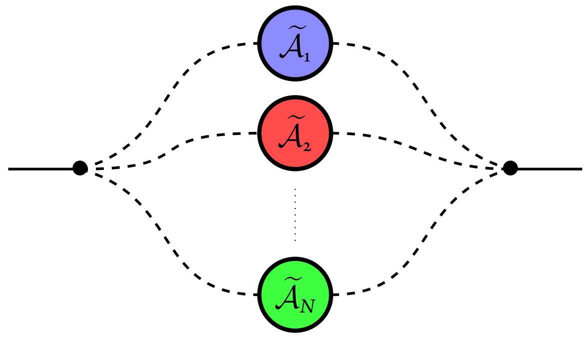


Figure 3.3: *Superposition of N independent channels specified by their vacuum extensions.* An input system $A = A_1 \oplus A_2 \oplus \cdots \oplus A_N$ branches out according to its sectors, with the j-th branch sent through the vacuum extension $\widetilde{\mathcal{A}}_j$. The outputs are finally recombined to form the overall output of the superposition channel.

3.5 Superposition of multiple independent channels

The superposition of multiple independent channel is defined by a direct generalisation of Definition 7, and is illustrated in Figure 3.3. Specifically, the superposition of N channels $\mathcal{A}^{(1)}, \ldots, \mathcal{A}^{(N)}$ specified by the vacuum extensions $\widetilde{\mathcal{A}}^{(1)}, \ldots, \widetilde{\mathcal{A}}^{(N)}$ is the channel with Kraus operators

$$S_{i_1 \cdots i_N} = \bigoplus_{j=1}^{N} \alpha_{i_1}^{(1)} \cdots \alpha_{i_{j-1}}^{(j-1)} A_{i_j}^{(j)} \alpha_{i_{j+1}}^{(j+1)} \cdots \alpha_{i_N}^{(N)} , \tag{3.20}$$

where $\{A_{i_j}^{(j)}\}$ is a Kraus representation of channel $\mathcal{A}^{(N)}$, and $\{\alpha_{i_j}^{(j)}\}$ is the corresponding set of vacuum amplitudes.

3.6 Superposition of multi-partite channels

The idea of combining multiple independent quantum channels in a superposition can be extended to correlated quantum channels. In general, correlations can occur between spatially separated regions acting at the same moment in time, or between different moments of time in same the region. In the following, we illustrate the main ideas of correlated channels using correlations in time, but the same ideas can be applied to any types of correlations.

Consider a 2-step correlated channel, for example the random unitary channel (2.51), described

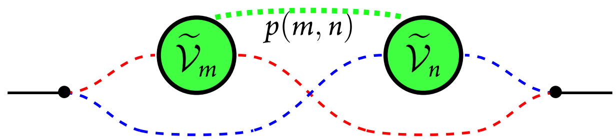


Figure 3.4: *The superposition of times of a time-correlated random unitary channel.* A single particle is sent at a superposition of two times (red and blue dashed lines), through the same transmission line (green ovals). The green dotted line represents the correlations between random unitary processes $\mathcal{V}_m$ and $\mathcal{V}_n$ taking place with probability $p(m,n)$ at the two subsequent uses of the transmission line, respectively.

in §2.7.1, and given again here for convenience:

$$\mathcal{R}(\rho_{12}) = \sum_{m,n} p(m,n)\,(V_m \otimes V_n)\,\rho_{12}\,(V_m \otimes V_n)^\dagger\,.$$

Consider the situation where the input to the channel is a single particle, carrying information in its internal degrees of freedom. Classically, the particle must be sent either at time t_1, or at time t_2, or at some random mixture of t_1 and t_2. When the particle is sent at time t_1, its evolution is given by the reduced channel $\mathcal{R}_1(\rho) := \sum_m p_1(m)\,V_m\rho V_m^\dagger$, where $p_1(m) := \sum_n p(m,n)$ is the marginal probability distribution of the unitaries at time t_1. Similarly, if the particle is sent at time t_2, its evolution is given by the channel $\mathcal{R}_2(\rho) := \sum_m p_2(n)\,V_n\rho V_n^\dagger$, with $p_2(n) := \sum_m p(m,n)$. A random choice of transmission times then results into a random mixture of the evolutions corresponding to channels $\mathcal{R}_1$ and $\mathcal{R}_2$. Crucially, the evolution of the particle is independent of any correlation that may be present in the probability distributions $p(m,n)$, that is, of any correlation between the first and the second use of the device.

More generally, we can consider a device described by an arbitrary 2-step correlated channel

$$\mathcal{C}_{12}(\rho_{12}) = \sum_m C_m \rho_{12} C_m^\dagger \in \mathsf{Chan}(S^{(1)}, S^{(2)})\,, \tag{3.21}$$

not necessarily of the random unitary form, where $\rho_{12} \in \mathsf{St}(S^{(1)} \otimes S^{(2)})$. The use of this device at a single time step t_1 is given by

$$\mathcal{C}_1(\rho_1) = \mathrm{Tr}_2\left[\mathcal{C}(\rho_{12})\right]\,, \tag{3.22}$$

where $\rho_1 = \mathrm{Tr}_2\,\rho_{12}$ is the reduced state on system $S^{(1)}$ (this follows the definition of a quantum comb, cf. §2.3.2).

In contrast, quantum physics allows for the possibility of transmitting a single particle in a way that is sensitive to the correlations between noisy processes at different times. The key idea is that the time when the particle is transmitted can be indefinite, as the particle could be sent through the transmission line at a coherent superposition of times t_1 and t_2 (see illustration in Figure 3.4). The superposition of transmission times could be achieved by adding an interferometric setup before the transmission line, letting the particle travel on a coherent superposition of two paths, one of which includes a delay [88]. This results in a time-bin qubit, described by a superposition of amplitudes corresponding to localisation at two different points in time, separated by a time difference much greater than a photon's coherence time [138].

Before developing the full formalism for describing single particle transmission through time-correlated channels, let us illustrate the key ideas with an example from quantum optics. Consider a single photon with degrees of freedom corresponding to the horizontal and vertical polarisation modes H_1, V_1 in the first time bin, and H_2, V_2 in the second time bin. First, the single photon is initialised in a polarisation state, which carries the message to be transmitted. It is then sent through an interferometer, with a time delay on one arm but which does not alter the polarisation state, so that the only role of the interferometric setup is to coherently control the moment of transmission. The result is a linear combination of (particle number) states of the form $(\alpha|1\rangle_{H1}|0\rangle_{V1} + \beta|0\rangle_{H1}|1\rangle_{V1}) \otimes |0\rangle_{H2}|0\rangle_{V2}$ and states of the form $|0\rangle_{H1}|0\rangle_{V1} \otimes (\alpha|1\rangle_{H2}|0\rangle_{V2} + \beta|0\rangle_{H2}|1\rangle_{V2})$, where $|\alpha|^2 + |\beta|^2 = 1$.

The composite system of the two modes in the first (second) time bin can be regarded as an extension of the system $S^{(1)}$ ($S^{(2)}$) in the time-correlated channel (3.21), constructed by appending a one-dimensional vacuum sector, i.e. $\widetilde{S}^{(1)} := S^{(1)} \oplus \mathrm{Vac}^{(1)}$ ($\widetilde{S}^{(2)} := S^{(2)} \oplus \mathrm{Vac}^{(2)}$). The states produced by the interferometric setup can then be written as a linear combination of states of the form $|\psi\rangle_1 \otimes |\mathrm{vac}\rangle_2$ and states of the form $|\mathrm{vac}\rangle_1 \otimes |\psi\rangle_2$, where $|\mathrm{vac}\rangle_i := |0\rangle_{Hi}|0\rangle_{Vi}$ is the vacuum state of the modes in system $\widetilde{S}^{(i)}$, and $|\psi\rangle_i := \alpha_i\,|1\rangle_{Hi}|0\rangle_{Vi} + \beta_i\,|0\rangle_{H_i}|1\rangle_{Vi}$ is a single-photon polarisation state, for $i \in \{1, 2\}$ (where $|\alpha_i|^2 + |\beta_i|^2 = 1$). For example, if the interferometer consists of 50:50 beamsplitters with a phase of zero between the two arms, then the state of the particle upon preparation of the time-bin qubit is $(|\psi\rangle_1 \otimes |\mathrm{vac}\rangle_2 + |\mathrm{vac}\rangle_1 \otimes |\psi\rangle_2)/\sqrt{2}$. The action of the time-correlated channel on the particle is then given by applying the channel to this state.

Generalising the above example, we model the transmission of a single particle through a bipartite channel by considering *abstract modes* described by the extended systems $\widetilde{S}^{(1)} := S^{(1)} \oplus \mathrm{Vac}^{(1)}$ and $\widetilde{S}^{(2)} := S^{(2)} \oplus \mathrm{Vac}^{(2)}$, each of which can contain a single particle equipped with an internal degree of freedom, such as a photon's polarisation, or no particle. This allows us to use the formalism of vacuum extensions to describe using the ports $S^{(1)}$ and $S^{(2)}$ in a superposition.

Mathematically, we proceed as follows. Consider a mulitport device described by a bipartite channel $\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes S^{(2)})$. A vacuum extension of the channel $\mathcal{B}$ is another bipartite channel $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)})$. In general, the systems $S^{(1)}, S^{(2)}$ can represent the systems accessible at different locations at the same time, or the systems accessible at the same location at two consecutive moments of time (in which case we have a 2-step correlated quantum channel), or more generally, they can represent any pair of independently addressable systems, representing the input/output ports of a multiport device.

In order to be able to send the same quantum particle to either of the ports of the device, we require the isomorphism $S^{(1)} \simeq S^{(2)} \simeq M$, where M is the message-carrying degree of freedom of the particle. The choice of port is determined by a control qubit C. In this case, we can define the situation in which a single particle is sent at a superposition of two different ports:

Definition 8. *The superposition of ports of a bipartite channel* $\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes S^{(2)})$ *specified by the vacuum extension* $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)})$ *is the channel*

$$\begin{aligned} &\mathcal{S} : \mathsf{Chan}(\widetilde{S}^{(1)}) \times \mathsf{Chan}(\widetilde{S}^{(2)}) \to \mathsf{Chan}(M \otimes C) \\ &\mathcal{S}(\widetilde{\mathcal{B}})(\rho \otimes \omega) = \mathcal{U}^\dagger \circ \widetilde{\mathcal{B}} \circ \mathcal{U}\,(\rho \otimes \omega)\,, \end{aligned} \tag{3.23}$$

where the isomorphism

$$\begin{aligned} &\mathcal{U}(\cdot) = U(\cdot)U^\dagger \\ &U : \mathcal{H}_M \otimes \mathcal{H}_C \to (\mathcal{H}_{S^{(1)}} \otimes \mathcal{H}_{\mathrm{Vac}}) \oplus (\mathcal{H}_{\mathrm{Vac}} \otimes \mathcal{H}_{S^{(2)}}) \end{aligned} \tag{3.24}$$

is defined by

$$\begin{aligned} U(|\psi\rangle_M \otimes |0\rangle_C) &:= |\psi\rangle_{\widetilde{S}^{(1)}} \otimes |\mathrm{vac}\rangle_{\widetilde{S}^{(2)}} \\ U(|\psi\rangle_M \otimes |1\rangle_C) &:= |\mathrm{vac}\rangle_{\widetilde{S}^{(1)}} \otimes |\psi\rangle_{\widetilde{S}^{(2)}}\,. \end{aligned} \tag{3.25}$$

Note that Definition 8 can be applied in particular to k-step quantum channels, which are a special case of k-partite channels. In this case we shall sometimes refer to the superposition of ports as a *superposition of times*. The illustration of the supermap $\mathcal{S}$ in the general case and in the special case of a 2-step channels is provided in Figure 3.5.

Note also, that the original Definition 7 is in fact a special case of Definition 8. In the case where the two ports represent spatially separated regions, either correlated or independent (as in Definition 7), we shall sometimes refer to the superposition of ports as a *superposition of paths*.

Let us briefly return to the example of the bipartite random unitary channel (2.51) in order to illustrate an application of the above definitions. In §6.1, by applying Definition 8 to this channel, we find that the use of the bipartite random unitary channel at a superposition of times is described by

$$\mathcal{S}(\widetilde{\mathcal{R}})(\rho \otimes \omega) = \sum_{m,n} p(m,n)\, W_{mn}\, (\rho \otimes \omega)\, W_{mn}^{\dagger}\,, \tag{3.26}$$

where

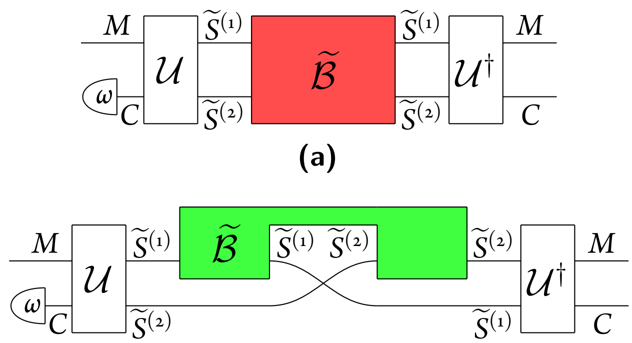
$$W_{mn} := V_m\, e^{i\varphi_n} \otimes |0\rangle\langle 0| + e^{i\varphi_m}\, V_n \otimes |1\rangle\langle 1|\,, \tag{3.27}$$

and $e^{i\varphi_m}$ is the vacuum amplitude associated with the unitary V_m (assuming that the vacuum extension each unitary V_m is another unitary). From these equations, it is clear that, in contrast to the possibilities allowed by classical physics, the overall evolution in general depends on the correlations $p(m,n)$ between the two uses of the device, even though only a single particle is sent through the channel.

Finally, Definition 8 can be extended to the transmission of a single particle through an N-partite multiport device. In this case, the device is represented by an N-partite quantum channel $\mathcal{B} \in \mathsf{Chan}(S^{(1)} \otimes \cdots \otimes S^{(N)})$, with $S^{(1)} \simeq S^{(2)} \simeq \cdots \simeq S^{(N)}$, and with vacuum extension $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{S}^{(1)} \otimes \cdots \otimes \widetilde{S}^{(N)})$. The superposition channel is then defined as the restriction of $\widetilde{\mathcal{B}}$ to the one-particle sector

$$\bigoplus_{j=1}^{N} \mathrm{Vac}^{(1)} \otimes \cdots \otimes \mathrm{Vac}^{(j-1)} \otimes S^{(j)} \otimes \mathrm{Vac}^{(j+1)} \otimes \cdots \otimes \mathrm{Vac}^{(N)} \simeq M \otimes C\,, \tag{3.28}$$

where C is now an N-dimensional control system.



(a)

(b)

Figure 3.5: (a) *Transmission of a single particle through a bipartite quantum channel* $\widetilde{\mathcal{B}}$ *(red)*. (b) *Transmission of a single particle through a 2-step quantum channel* $\widetilde{\mathcal{B}}$ *(green)*. In both cases, the particle is represented by a composite system $M \otimes C$, where M represents the degrees of freedom used as the message, and C represents the degrees of freedom used as the control. The isomorphism $\mathcal{U}$ converts the composite system $M \otimes C$ into the one-particle sector $(S^{(1)} \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes S^{(2)})$ of $\widetilde{S}^{(1)} \otimes \widetilde{S}^{(2)}$. The inverse map $\mathcal{U}^\dagger$ converts the output state back into $M \otimes C$. For applications to communication, we take the input of the control system C to be fixed in the state ω whilst the message system M is accessible to the sender.

4

Resource theories of communication

In this chapter, we formulate communication with superpositions of trajectories and communication with superpositions of causal orders as resource theories. To do this, we construct a general framework for resource theories of communication. Our general framework includes a minimal requirement that all communication paradigms must satisfy in order to form a meaningful resource theory. We show that standard quantum Shannon theory, quantum Shannon theory with superpositions of trajectories, and quantum Shannon theory with superpositions of causal orders all satisfy this requirement. Additionally, we consider the communication paradigm proposed in Ref. [96], which we call quantum Shannon theory with superpositions of encoding and decoding operations. In contrast, we show that this paradigm does not satisfy our minimal requirement, and therefore, we argue, does not constitute a meaningful paradigm of communication.

4.1 Standard quantum Shannon theory as a resource theory of communication

We begin by formulating standard quantum Shannon theory as a resource theory, setting the scene for its extension to more general resource theories of communication.

4.1.1 Quantum Shannon theory as a theory of resources

A central task in information theory is to quantify the amount of information that a given communication device can transmit. In general, the amount of information can be classical or quantum, or of other types. In this work, we shall focus on classical and quantum information. To make the quantification unambiguous, it is essential to specify how the given device can be used. The device represents a resource, and the rules on the possible uses of this resource can be formulated as a *resource theory* [60, 67].

A resource-theoretic approach to standard quantum Shannon theory was initiated by Devetak, Harrow, and Winter [74]. Further resource-theoretic formalisations have been put forward in Refs. [80, 131, 132, 186] in a variety of communication scenarios. Related resource theories of quantum devices have been recently formulated in Refs. [26, 169, 187] for purposes other than the theory of communication.

In this work we shall adopt the general framework for resource theories proposed by Coecke, Fritz, and Spekkens [67]. In this framework, the set of all possible resources is described by a set of objects, equipped with a set of operations acting on them. The set of operations is closed under sequential and parallel composition. For example, the set of operations, hereafter denoted by M, could be the set of all quantum channels acting on finite-dimensional quantum systems (the objects). The central idea of the resource-theoretic framework is to define a subset of operations $\mathsf{M}_{\text{free}} \subseteq \mathsf{M}$, which are regarded as *free*. The notion of resource is then defined relative to the set of free operations: a state or an operation is a non-trivial resource if and only if it is not free, and a resource is more valuable than another if the former can be converted into the latter by means of free operations.

Different choices of free operations generally define different resources. Intuitively, the set of free operations is meant to capture some operational restriction, which makes some operations 'easy to implement'. In principle, however, M_{free} could be any subset of operations, as long as it is closed under sequential and parallel composition. In this respect, the resource-theoretic approach is a conceptual tool to understand the power of the set M_{free}, irrespectively of whether implementing the operations in it is easy or not.

In quantum Shannon theory, the input resources are communication channels, or, more precisely, *uses* of communication channels. For example, the ability to transfer a single qubit from a sender to a receiver is modelled as a single use of a single-qubit identity channel.

To cast quantum Shannon theory in the resource-theoretic framework of Ref. [67], one has to regard the various types of quantum channels as objects, and to define the allowed operations that transform input channels into output channels, that is *quantum supermaps* (cf. §2.3.2) [52, 54, 57]. In the following, we shall define the sets of free supermaps M_{free} for some of the basic scenarios in quantum Shannon theory, setting the scene for the generalisations studied in the rest of this chapter.

4.1.2 Direct communication from a sender to a receiver through a single channel

Consider the basic communication scenario where a sender (Alice) communicates directly to a receiver (Bob). At the fundamental level, the possibility of communication consists of two ingredients: the availability of a piece of hardware that serves as a communication device, and the placement of that piece of hardware between the sender and the receiver. For example, the piece of hardware could be an optical fibre, and the placement could be provided by a communication company that laid the fibre between the sender's and the receiver's locations. In some situations, the placement is implicit: for example, the sender and receiver could be communicating through a medium, such as the air between them, which has been placed there, as it were, by Nature itself.

Mathematically, the communication device is described by a quantum channel $\mathcal{N} \in \mathsf{Chan}(X \to Y)$, which transforms systems of type X into systems of type Y. For example, the systems could be single qubits, encoded in the polarisation of single photons. At this level, the systems are not assigned a specific location in spacetime. Accordingly, we shall call the systems X and Y *unplaced systems*, and the channel $\mathcal{N} \in \mathsf{Chan}(X \to Y)$ an *unplaced channel*.

The placement of the device can be described by introducing a *placement operation*, which corresponds to putting the input (output) system at the sender's (receiver's) location. Mathematically, a placement operation is a supermap that transforms channels in $\mathsf{Chan}(X \to Y)$ into channels in $\mathsf{Chan}(A \to B)$, where system A (B) is of the same type (i.e. corresponds to a Hilbert space with the same dimension) as system X (Y), denoted as $A \simeq X$ ($B \simeq Y$), and is placed at the sender's

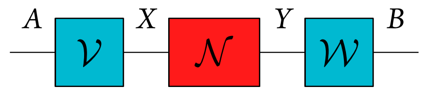


Figure 4.1: *Basic placement supermap* $\mathcal{S}^{A,B}_{\text{place}}(\mathcal{N}) := \mathcal{W}^B \circ \mathcal{N} \circ \mathcal{V}^A$. In this chapter, the unplaced communication channels $\mathcal{N}$ are drawn in red, while the placement supermaps (i.e. supermaps from unplaced channels to placed channels) are drawn in blue.

(receiver's) end, as illustrated in Figure 4.1. Explicitly, we define the *basic placement supermap* as:

$$\mathcal{S}^{A,B}_{\text{place}}(\mathcal{N}) := \mathcal{W}^B \circ \mathcal{N} \circ \mathcal{V}^A , \tag{4.1}$$

where $\mathcal{V}^A \in \mathsf{Chan}(A \to X)$ and $\mathcal{W}^B \in \mathsf{Chan}(Y \to B)$ are unitary channels implementing the isomorphisms $A \simeq X$ and $Y \simeq B$, respectively.

We shall call the systems A and B *placed systems*, and the channel $\mathcal{C} := \mathcal{S}^{A,B}_{\text{place}}(\mathcal{N})$ a *placed channel*. In the following, we shall use the letters $\mathcal{N}$ and $\mathcal{C}$ for unplaced and placed channels, respectively. Similarly, we shall use letters at the end of the alphabet (e.g. X, Y) to represent unplaced systems, and letters from the beginning or middle of the alphabet (e.g. A, B, R, S) to represent placed systems. In the figures in this Chapter (as well as in Chapter 7, where the distinction between placed and unplaced channels is important), we shall represent unplaced channels as red boxes, and placed channels as green boxes. This choice of colours reflects the fact that the placed channels are ready to be used by the communicating parties, while the unplaced channels have yet to be made available to them.

Once a device is in place, the sender and receiver can use it to communicate to one another. Typically, communication is achieved by connecting the communication device with other devices present at the sender's and receiver's locations. For example, one end of an optical fibre could be connected to a computer, used by the sender to type an email, and the other end of the fibre could be connected to another computer, used by the receiver to read the email. The operations performed by the sender and receiver can be described by a supermap transforming placed channels in $\mathsf{Chan}(A \to B)$ into placed channels in $\mathsf{Chan}(A' \to B')$, where A' and B' are two new input and output systems, also placed in the sender's and receiver's locations, respectively.

As reviewed in §2.3.2, Ref. [52] showed that the most general supermap $\mathcal{S}$ transforming a generic

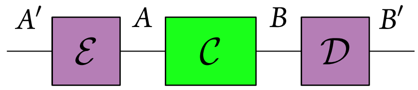


Figure 4.2: *Encoding-decoding supermap* $\mathcal{S}_{\mathcal{E},\mathcal{D}}(\mathcal{C}) := \mathcal{D} \circ \mathcal{C} \circ \mathcal{E}$. In this chapter, the placed quantum channels are drawn in green, while the encoding-decoding supermaps are drawn in violet.

input channel $\mathcal{C} \in \mathsf{Chan}(A \to B)$ into an output channel $\mathcal{S}(\mathcal{C}) \in \mathsf{Chan}(A' \to B')$ has the form

$$\mathcal{S}(\mathcal{C}) = \mathcal{D} \circ (\mathcal{C} \otimes \mathcal{I}_{\text{Aux}}) \circ \mathcal{E}\,, \tag{4.2}$$

where Aux is an auxiliary quantum system, and $\mathcal{E} \in \mathsf{Chan}(A' \to A \otimes \text{Aux})$ and $\mathcal{D} \in \mathsf{Chan}(B \otimes \text{Aux} \to B')$ are quantum channels. These supermaps define the set of all possible operations on input channels, and play the role of the set M in the general resource-theoretic framework described in §4.1.1.

To specify the set of free operations, one has to specify a *subset* of the set of all supermaps. The standard choice (in the absence of additional resources such as shared entanglement or shared randomness) is to require the free operations to have the form

$$\mathcal{S}_{\mathcal{E},\mathcal{D}}(\mathcal{C}) := \mathcal{D} \circ \mathcal{C} \circ \mathcal{E} \tag{4.3}$$

(see Figure 4.2 for an illustration). Operationally, this choice of M_{free} is justified by the fact that the supermaps (4.3) can be achieved by performing a local encoding operation $\mathcal{E}$ at the sender's side and a local decoding operation $\mathcal{D}$ at the receiver's side, without requiring the transmission of any system other than the system sent through the channel $\mathcal{C}$.

Note that while the supermaps (4.3) are the standard choice, other choices could be made. For example, one could consider quantum communication with the assistance of classical communication [30], or classical communication with the assistance of shared entanglement [32]. In these scenarios, the set of free supermaps is larger than the set of supermaps of the form (4.3), and contains supermaps that can be achieved with the additional resources under consideration. For completeness, the characterisation of such supermaps is provided in §4.1.6. In the following, however, we shall predominantly stick to the simplest choice of free supermaps, namely the choice in (4.3).

In general, we shall refer to supermaps from unplaced channels to placed channels as *placement* *supermaps*, and we shall interpret them as being performed either by a communication provider, or by Nature itself. We shall refer to supermaps from placed channels to placed channels as *party supermaps*, and shall interpret them as being performed by the communicating parties.

4.1.3 Direct communication from a sender to a receiver through multiple channels

So far, we have considered operations on a single quantum channel. We now extend the resource-theoretic formulation to scenarios where multiple communication channels (or multiple uses of the same communication channel) are available.

Consider a communication protocol that uses k communication devices, described by k unplaced channels $\mathcal{N}_1, \ldots, \mathcal{N}_k$, with $\mathcal{N}_i \in \mathsf{Chan}(X_i \to Y_i)$ for $i \in \{1, \ldots, k\}$. We denote by $(\mathcal{N}_1, \ldots, \mathcal{N}_k)$ the resource corresponding to a single use of each device. Again, the list $(\mathcal{N}_1, \ldots, \mathcal{N}_k)$ is interpreted as a description of the hardware before it is placed between the sender and receiver. For example, the hardware could be a list of optical fibres with some given specifications, such as attenuation coefficient, bandwidth, birefringence and length.

In Appendix B.1 we show that the list $(\mathcal{N}_1, \ldots, \mathcal{N}_k)$ can be interpreted as an equivalent notation for the product channel $\mathcal{N}_1 \otimes \cdots \otimes \mathcal{N}_k$, viewed as an element of a suitable set of channels (namely, k-partite no-signalling channels). In the following, we shall use the list notation $(\mathcal{N}_1, \ldots, \mathcal{N}_k)$ as a visual reminder that the channels $\mathcal{N}_1, \ldots, \mathcal{N}_k$ are unplaced.

In the direct communication scenario, it is understood that all the input systems are placed in Alice's laboratory, and all the output systems are placed in Bob's laboratory. Equivalently, this means that the communication devices are placed in parallel between the sender and the receiver. The operation of placing the devices in parallel is described by the *parallel placement supermap* $\mathcal{S}_{\text{par}}^{\mathbf{A},\mathbf{B}}$ defined by

$$\mathcal{S}_{\text{par}}^{\mathbf{A},\mathbf{B}}(\mathcal{N}_1, \ldots, \mathcal{N}_k) := \mathcal{S}_{\text{place}}^{A_1,B_1}(\mathcal{N}_1) \otimes \cdots \otimes \mathcal{S}_{\text{place}}^{A_k,B_k}(\mathcal{N}_k) \,. \tag{4.4}$$

where $\mathbf{A} := (A_1, \ldots, A_k)$ $[\mathbf{B} := (B_1, \ldots, B_k)]$ is a list of quantum systems placed in Alice's (Bob's) laboratory, with $A_i \simeq X_i$ and $B_i \simeq Y_i$ for every $i \in \{1, \ldots, k\}$. The result of the supermap is a placed quantum channel in $\mathsf{Chan}(A_1 \otimes \cdots \otimes A_k \to B_1 \otimes \cdots \otimes B_k)$.

A large body of results in standard quantum Shannon theory refers to channels combined in parallel as in Equation (4.4). For example, as discussed in §2.2.3, Smith and Yard [182] showed that, surprisingly, the parallel composition of two channels with zero quantum capacity can give rise to a channel with non-zero quantum capacity. This phenomenon became known as *activation of the quantum capacity*.

4.1.4 Network communication from a sender to a receiver

Let us now consider a communication scenario where the sender (Alice) and receiver (Bob) communicate through a network of communication devices. To begin with, we focus on the simple case where Alice and Bob communicate through two devices, which are connected by an intermediate party (Ray), who serves as a 'repeater' passing to Bob the information received from Alice.

The initial resource is described by a pair of unplaced channels $(\mathcal{N}_1, \mathcal{N}_2) \in \mathsf{Chan}(X_1 \to Y_1) \times \mathsf{Chan}(X_2 \to Y_2)$. The operation of placing channel $\mathcal{N}_1$ between Alice and Ray, and channel $\mathcal{N}_2$ between Ray and Bob is described by the *sequential placement supermap* $\mathcal{S}_{\text{seq}}^{A,R,R',B}$ defined by

$$\mathcal{S}_{\text{seq}}^{A,R,R',B}(\mathcal{N}_1, \mathcal{N}_2) := \mathcal{S}_{\text{place}}^{A,R}(\mathcal{N}_1) \otimes \mathcal{S}_{\text{place}}^{R',B}(\mathcal{N}_2) \,, \tag{4.5}$$

where system $A \simeq X_1$ is placed in Alice's laboratory, systems $R \simeq Y_1$ and $R' \simeq X_2$ are placed in Ray's laboratory, and system $B \simeq Y_2$ is placed in Bob's laboratory.

Note that the sequential placement (4.5) is formally identical to the parallel placement (4.4): in both cases, the placement of multiple channels is the tensor product of the placement of individual channels. The difference between parallel and sequential placement arises from the different spacetime locations in which the inputs and outputs of the channels are placed. In the parallel placement, all the input systems $\mathbf{A}$ are at the sender's location, and all the output systems $\mathbf{B}$ are at the receiver's location. In the sequential placement, the systems A, R, R', B appear in a strict sequential order: A before R, R before R', R' before B. This difference is crucial when it comes to specifying how the output of the placement supermap is to be used: in the case of parallel placement, the output of the supermap can be connected with local operations at the sender's and receiver's ends. In the case of sequential placement, intermediate operations are possible.

The difference between sequential and parallel placements is reflected by the different type of channels they generate. The output of the sequential placement supermap (4.5) is a 2-step quantum channel (see §2.7.1), where the first step represents the transfer of information from A to R, and the second step corresponds to the transfer of information from R' to B. Recall, that the difference between a two-step channel and a generic bipartite channel is that the two-step channel has to satisfy the additional condition (2.32) of quantum combs, which ensures compatibility with the

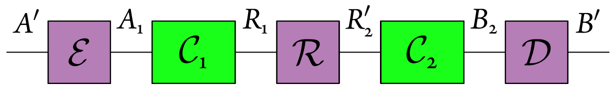


Figure 4.3: *Encoding-repeater-decoding supermap* $\mathcal{S}_{\mathcal{E},\mathcal{R},\mathcal{D}}(\mathcal{C}_1 \otimes \mathcal{C}_2) := \mathcal{D} \circ \mathcal{C}_2 \circ \mathcal{R} \circ \mathcal{C}_1 \circ \mathcal{E}$. In this chapter, party supermaps (i.e. supermaps from placed channels to placed channels) are drawn in violet.

causal ordering of the systems S_1, S_1', S_2, and S_2'.

The sequential placement supermap (4.5) transforms a pair of unplaced channels $(\mathcal{N}_1, \mathcal{N}_2)$ into a placed 2-step quantum channel $\mathcal{C}_1 \otimes \mathcal{C}_2$, with $\mathcal{C}_1 := \mathcal{S}^{A,R}_{\text{place}}(\mathcal{N}_1)$ and $\mathcal{C}_2 := \mathcal{S}^{R',B}_{\text{place}}(\mathcal{N}_2)$. Note that, in general, the set of 2-step quantum channels also contains maps that are not of the product form $\mathcal{C}_1 \otimes \mathcal{C}_2$. These maps correspond to two-step processes where a memory is passed from the first step to the second.

Once the devices have been placed, the sender, repeater, and receiver can connect them with their local devices, thus establishing a single channel that transfers information directly from the sender to the receiver. The most general supermaps from multi-step quantum channels to quantum channels have been characterised in Ref. [54]. Their action on product multi-step quantum channels $\mathcal{C}_1 \otimes \mathcal{C}_2$ is given by

$$\mathcal{S}(\mathcal{C}_1 \otimes \mathcal{C}_2) = \mathcal{D} \circ (\mathcal{C}_2 \otimes \mathcal{I}_{\text{Aux}_2}) \circ \mathcal{R} \circ (\mathcal{C}_1 \otimes \mathcal{I}_{\text{Aux}_1}) \circ \mathcal{E} \,, \tag{4.6}$$

where Aux_1 and Aux_2 are auxiliary systems, and $\mathcal{E}, \mathcal{R}$, and $\mathcal{D}$ are arbitrary channels in $\mathsf{Chan}(A' \to A \otimes \text{Aux}_1)$, $\mathsf{Chan}(R \otimes \text{Aux}_1 \to R' \otimes \text{Aux}_2)$, and $\mathsf{Chan}(B \otimes \text{Aux}_2 \to B')$, respectively.

The standard choice of free supermaps is the supermaps that are achievable without the auxiliary systems Aux_1 and Aux_2, that is, the supermaps of the form

$$\mathcal{S}_{\mathcal{E},\mathcal{R},\mathcal{D}}(\mathcal{C}_1 \otimes \mathcal{C}_2) := \mathcal{D} \circ \mathcal{C}_2 \circ \mathcal{R} \circ \mathcal{C}_1 \circ \mathcal{E} \,, \tag{4.7}$$

illustrated in Figure 4.3.

Communication through a network of $k \geq 2$ devices is described by a direct generalisation of the above example. Consider the situation where a sender communicates to a receiver with the assistance of $k-1$ intermediate repeaters. The communication devices are described by a list of unplaced channels $(\mathcal{N}_1, \dots, \mathcal{N}_k) \in \mathsf{Chan}(X_1 \to Y_1) \times \mathsf{Chan}(X_2 \to Y_2) \times \cdots \times \mathsf{Chan}(X_k \to$

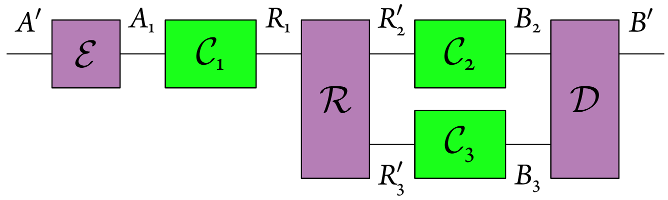


Figure 4.4: *A composite supermap.* The figure shows the placement of $k = 3$ channels between a sender, a single repeater ($r = 1$), and a receiver. The placement is then followed by an encoding-repeater-decoding supermap (in violet), representing the local operations performed by the sender, repeater, and receiver.

Y_k). The placement of the devices between the sender, repeaters, and receiver is described by the supermap

$$\begin{aligned} &\mathcal{S}_{\text{seq}}^{A,R_1,R'_1,\ldots,R_{k-1},R'_{k-1},B}(\mathcal{N}_1,\ldots,\mathcal{N}_k) \\ &:= \mathcal{S}_{\text{place}}^{A,R_1}(\mathcal{N}_1) \otimes \mathcal{S}_{\text{place}}^{R'_1,R_2}(\mathcal{N}_2) \otimes \cdots \otimes \mathcal{S}_{\text{place}}^{R'_{k-1},B}(\mathcal{N}_k)\,, \end{aligned} \tag{4.8}$$

where system $A \simeq X_1$ is placed in the sender's laboratory, system $B \simeq Y_k$ is placed in the receiver's laboratory, and systems $R_i \simeq Y_i$ and $R'_i \simeq X_{i+1}$ are placed in the laboratory of the i-th repeater, for $i \in \{1,\ldots,k-1\}$. The output of the supermap $\mathcal{S}_{\text{seq}}^{A,R_1,R'_1,\ldots,R_{k-1},R'_{k-1},B}$ is a k-step quantum channel.

Once the available devices have been placed, the communicating parties can connect their local devices to the placed communication channels. The corresponding supermap has the form

$$\begin{aligned} &\mathcal{S}_{\mathcal{E},\mathcal{R}_1,\ldots,\mathcal{R}_{k-1},\mathcal{D}}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_k) \\ &:= \mathcal{D} \circ \mathcal{C}_k \circ \mathcal{R}_{k-1} \circ \mathcal{C}_{k-1} \circ \cdots \circ \mathcal{R}_1 \circ \mathcal{C}_1 \circ \mathcal{E}\,, \end{aligned} \tag{4.9}$$

where $\mathcal{E} \in \mathsf{Chan}(A' \to A)$ is the encoding operation performed by the sender, $\mathcal{R}_i \in \mathsf{Chan}(R_i \to R'_i)$ is the repeater operation performed by the i-th intermediate party, and $\mathcal{D} \in \mathsf{Chan}(B \to B')$ is the decoding operation performed by the receiver.

More generally, one can consider any placement of $k \geq 2$ devices with $r \leq k-1$ intermediate repeaters. This includes placing some channels in parallel between two subsequent parties, in which case the placed channel is a $(r+1)$-step quantum channel. An example of this situation is illustrated in Figure 4.4. The most general placement supermaps corresponding to a definite causal structure of communicating parties are described in the following section.

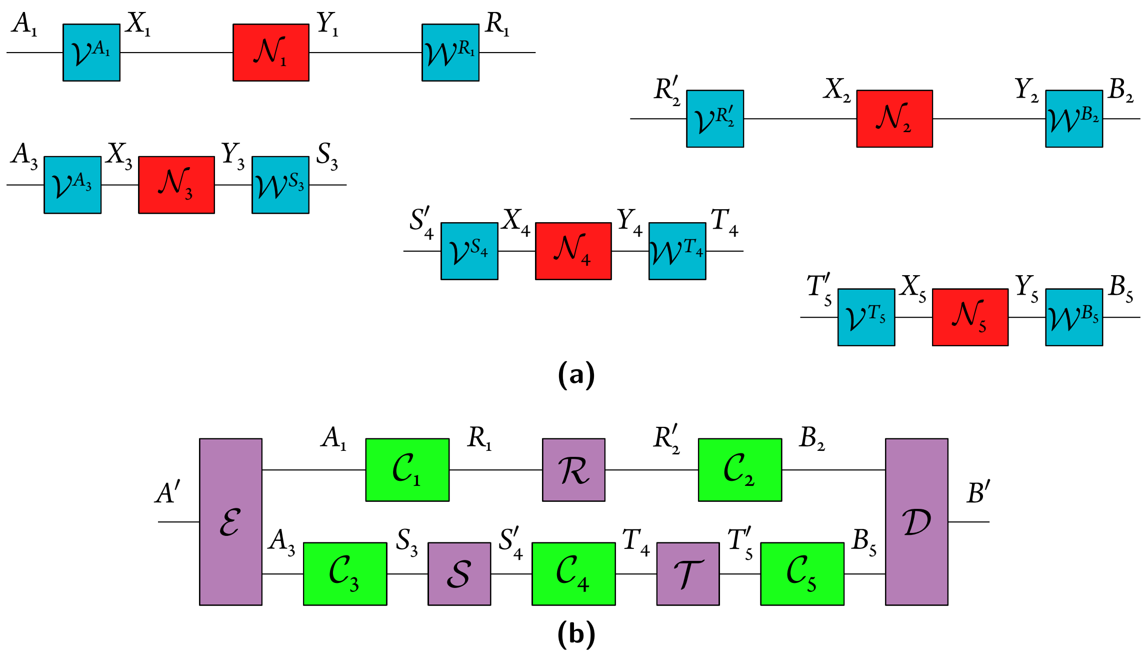


Figure 4.5: (a) *An illustration of the placement supermap* $\mathcal{S}_{\text{network}}^{A_1R_1R_2'B_2A_3S_3S_4'T_4T_5'B_5}$, *described by Eq.* (4.10), *acting on a list of five unplaced channels* $(\mathcal{N}_1,\dots,\mathcal{N}_5)$. (b) *An illustration of the encoding-repeater-decoding supermap* $\mathcal{S}_{\mathcal{E},\mathcal{R},\mathcal{S},\mathcal{T},\mathcal{D}}$, *described by Eq.* (4.12), *acting on the resulting placed channel of (a).*

4.1.5 Placement of channels in an arbitrary (definite) causal structure

Communication through a network of $k \geq 2$ devices, connected via $r \leq k-1$ intermediate parties, is described by specifying the causal structure of the communicating parties, and by considering supermaps that are compatible with that causal structure. In the case of a sender **A**, receiver **B** and a single repeater **R** (where a boldface letter **R** is identified with the list of input/output systems $(R_1,\dots,R_k,R_1',\dots,R_l')$ accessible to a given communicating party), the causal structure is implicitly given by the totally ordered set $\{\mathbf{A}\preceq\mathbf{R},\mathbf{R}\preceq\mathbf{B}\}$, where $\mathbf{A}\preceq\mathbf{B}$ denotes that **B** is in the future light cone of **A**. In this case, it is clear that only placements between **A** and **R**, **A** and **B**, or **R** and **B** are allowed.

In the case of $r \geq 2$ repeaters, a general causal structure is described by a partially ordered set (poset), with a choice of possible relations between the intermediate parties $\{\mathbf{R},\mathbf{S},\dots,\mathbf{T}\}$. Formally, a poset is a set endowed with a binary relation, which is reflexive, antisymmetric and transitive. The latter two properties ensure that loops in the causal structure are not allowed, i.e. if **A** precedes **R** and **R** precedes **S**, then **A** precedes **S**, and therefore **S** cannot also precede **A** (unless **A** = **R** = **S**). Physically, the description of causal structure as a poset is motivated by the structure of spacetime as described by special relativity [79].

Overall, the placement of communication devices between the communicating parties is described by a tensor product of basic placement supermaps, with the constraint that a placement from $\mathsf{Chan}(X_i \to Y_i)$ to $\mathsf{Chan}(S_i' \to T_i)$ is only possible if $\mathbf{S} \preceq \mathbf{T}$ in the causal structure.

We illustrate the scheme for a network of multiple repeaters with an example. Consider the communication scenario with a sender, a receiver, and $r = 3$ intermediate parties $\{\mathbf{R},\mathbf{S},\mathbf{T}\}$, arranged in a causal structure described by the poset $\{\mathbf{A} \preceq \mathbf{R}, \mathbf{R} \preceq \mathbf{B}, \mathbf{A} \preceq \mathbf{S}, \mathbf{S} \preceq \mathbf{T}, \mathbf{T} \preceq \mathbf{B}\}$. Suppose that the communicating parties have access to $k = 5$ devices, described by the list of unplaced channels $(\mathcal{N}_1, \ldots, \mathcal{N}_5) \in \mathsf{Chan}(X_1 \to Y_1) \times \cdots \times \mathsf{Chan}(X_5 \to Y_5)$. The use of the devices is specified by placing them in a particular configuration between the sender, receiver, and repeaters. One possible placement is given by

$$\begin{aligned}
&\mathcal{S}_{\text{network}}^{A_1 R_1 R_2' B_2 A_3 S_3 S_4' T_4 T_5' B_5}(\mathcal{N}_1, \ldots, \mathcal{N}_5) \\
&= \mathcal{S}_{\text{place}}^{A_1 R_1}(\mathcal{N}_1) \otimes \mathcal{S}_{\text{place}}^{R_2' B_2}(\mathcal{N}_2) \otimes \mathcal{S}_{\text{place}}^{A_3 S_3}(\mathcal{N}_3) \otimes \mathcal{S}_{\text{place}}^{S_4' T_4}(\mathcal{N}_4) \otimes \mathcal{S}_{\text{place}}^{T_5' B_5}(\mathcal{N}_5) \\
&= \mathcal{W}^{R_1} \circ \mathcal{N}_1 \circ \mathcal{V}^{A_1} \otimes \mathcal{W}^{B_2} \circ \mathcal{N}_2 \circ \mathcal{V}^{R_2'} \otimes \mathcal{W}^{S_3} \circ \mathcal{N}_3 \circ \mathcal{V}^{A_3} \otimes \mathcal{W}^{T_4} \circ \mathcal{N}_4 \circ \mathcal{V}^{S_4'} \otimes \mathcal{W}^{B_5} \circ \mathcal{N}_5 \circ \mathcal{V}^{T_5'} ,
\end{aligned} \tag{4.10}$$

essentially consisting of two sequences through repeaters, **R** and (**S**,**T**), placed in parallel between the sender and receiver, as illustrated in Figure 4.5a.

Note, that here the different intermediate parties are labelled **R**, **S**, ..., **T**. The subscript i (j) of the placed system R_i (R_j') at the communicating party **R** labels which input system X_i (output system Y_j) it corresponds to. In contrast, in the previous sections where each party only had access to a single system, R_i (R_i') denoted the single input (output) system of the i-th repeater party.

With the devices placed within the network of communicating parties, we once again consider the free operations on the placed channels. Consider a subset of $l \leq k$ placed channels $\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_l \in \mathsf{Chan}(\cdot \to R_1) \times \cdots \times \mathsf{Chan}(\cdot \to R_m) \times \mathsf{Chan}(R_{m+1}' \to \cdot) \times \cdots \times \mathsf{Chan}(R_l' \to \cdot)$, where the first $m \leq l$ channels have output systems at **R** (and any arbitrary placed input systems), and the remaining $l - m$ channels have input systems at **R** (and any arbitrary placed output systems). The final $k - l$ channels have neither input nor output systems at **R**. The free operations that can be performed at **R** are taken to be those of the form

$$\mathcal{S}_{\mathcal{R}}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_k) := [(\mathcal{C}_{m+1} \otimes \cdots \otimes \mathcal{C}_l) \circ \mathcal{R} \circ (\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_m)] \otimes (\mathcal{C}_{l+1} \otimes \cdots \otimes \mathcal{C}_k) , \tag{4.11}$$

where $\mathcal{R} \in \mathsf{Chan}(R_1 \otimes \cdots \otimes R_m \to R'_{m+1} \otimes \cdots \otimes R'_l)$. This includes as a special case the free operations $\mathcal{S}_{\mathcal{E}}$ and $\mathcal{S}_{\mathcal{D}}$ that can be performed by the sender and receiver, respectively, in which case $m = 0$ or $m = l$. Overall, the choice of free operations on placed channels is taken to be any sequential or parallel composition of (local) party supermaps of the form of Eq. (4.11). When two supermaps $\mathcal{S}_{\mathcal{R}}$ and $\mathcal{S}_{\mathcal{T}}$ commute, we use the shorthand $\mathcal{S}_{\mathcal{R},\mathcal{T}} := \mathcal{S}_{\mathcal{T}} \circ \mathcal{S}_{\mathcal{R}} = \mathcal{S}_{\mathcal{R}} \circ \mathcal{S}_{\mathcal{T}}$.

As an example, consider the placed channels given in Eq. (4.10) and let $\mathcal{C}_i = \mathcal{W} \circ \mathcal{N}_i \circ \mathcal{V}$. Then the action of the most general free supermap on the placed channels $\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_5$ is given by

$$\mathcal{S}_{\mathcal{E},\mathcal{R},\mathcal{S},\mathcal{T},\mathcal{D}}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_5) = \mathcal{D} \circ [(\mathcal{C}_2 \circ \mathcal{R} \circ \mathcal{C}_1) \otimes (\mathcal{C}_5 \circ \mathcal{T} \circ \mathcal{C}_4 \circ \mathcal{S} \circ \mathcal{C}_3)] \circ \mathcal{E} , \tag{4.12}$$

and is illustrated in Figurde 4.5b.

4.1.6 Free supermaps in assisted communication scenarios

For completeness, in this section we provide examples of supermaps that arise in communication scenarios involving the assistance of classical communication or entanglement.

Let us consider first the assistance of free classical communication [30], as illustrated in Figure 4.6. In this case, the free supermaps on placed channels have the form

$$\mathcal{S}_{\mathcal{E},\mathcal{D},\text{clas}}(\mathcal{C}) := \mathcal{D} \circ (\mathcal{C} \otimes \mathcal{I}^{\text{clas}}_{\text{Aux}}) \circ \mathcal{E} , \tag{4.13}$$

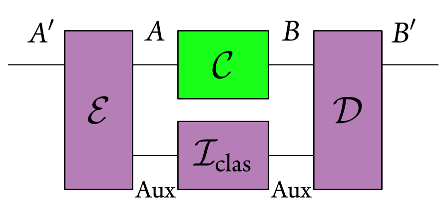


Figure 4.6: *Supermap describing encoding and decoding operations assisted by free classical communication.*

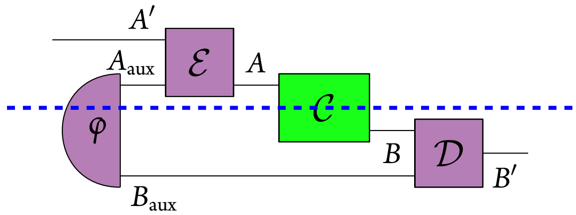


Figure 4.7: *Supermap describing encoding and decoding operations assisted by shared entanglement.* The blue dashed line denotes the partition between Alice (top) and Bob (bottom).

where $\mathcal{I}^{\text{clas}}_{\text{Aux}}$ is the classical identity channel, defined as $\mathcal{I}^{\text{clas}}_{\text{Aux}}(\rho) = \sum_j |j\rangle\langle j| \langle j|\rho|j\rangle$ for some orthonormal basis $\{|j\rangle\}$, and $\mathcal{E} \in \mathsf{Chan}(A' \to A \otimes \text{Aux})$ and $\mathcal{D} \in \mathsf{Chan}(B \otimes \text{Aux} \to B')$ are quantum channels.

Let us consider now classical communication with the assistance of shared entanglement [32]. In this case, the free operations on placed channels are those that can be achieved by performing encoding and decoding operations that act on a shared entangled state, as shown in Figure 4.7. Mathematically, these operations correspond to free supermaps of the form

$$\mathcal{S}_{\mathcal{E},\mathcal{D},\text{ent}}(\mathcal{C}) := \mathcal{D}_{BB_{\text{aux}}} \circ (\mathcal{C}_A \circ \mathcal{E}_{A'A_{\text{aux}}} \otimes \mathcal{I}_{B_{\text{aux}}}) \circ (\mathcal{I}_{A'} \otimes \varphi_{A_{\text{aux}}B_{\text{aux}}}), \tag{4.14}$$

where $\varphi_{A_{\text{aux}}B_{\text{aux}}}$ is an entangled state on system $A_{\text{aux}} \otimes B_{\text{aux}}$, $\mathcal{E} \in \mathsf{Chan}(A' \otimes A_{\text{aux}} \to A)$ and $\mathcal{D} \in \mathsf{Chan}(B \otimes B_{\text{aux}} \to B')$ are encoding and decoding channels, respectively, and the subscripts indicate the input systems of all channels.

4.1.7 Communication with correlated channels

So far, we have described each use of each communication device as a single unplaced quantum channel. However, as discussed in §2.7, we know that in general, multiple uses of the same communication device can be correlated. Similarly, simultaneous uses of two separate devices could exhibit correlations through space. Therefore, our resource-theoretic framework should include a description of the use of correlated quantum channels.

In fact, we already mentioned that a list of k unplaced channels can be viewed as a single k-partite no-signalling channel (cf. §4.1.3 and Appendix B.1). This means that the parallel and sequential placements described above can be directly applied to any no-signalling channel.

Specifically, the parallel placement places the device $\mathcal{N}_{1,\dots,k} \in \mathsf{NSChan}(A_1 \to B_1, \dots, A_k \to$

$B_k)$ as:

$$\mathcal{S}_{\text{par}}^{\mathbf{A},\mathbf{B}}(\mathcal{N}_{1,\ldots,k}) := (\mathcal{S}_{\text{place}}^{A_1,B_1} \otimes \cdots \otimes \mathcal{S}_{\text{place}}^{A_k,B_k})(\mathcal{N}_{1,\ldots,k}) \,, \tag{4.15}$$

where $\mathbf{A} := (A_1, \ldots, A_k)$ $[\mathbf{B} := (B_1, \ldots, B_k)]$ is a list of quantum systems placed in Alice's (Bob's) laboratory, with $A_i \simeq X_i$ and $B_i \simeq Y_i$ for every $i \in \{1, \ldots, k\}$, and NSChan is defined in §2.7.2. As before, the result of the supermap is a placed quantum channel in $\mathsf{Chan}(A_1 \otimes \cdots \otimes A_k \to B_1 \otimes \cdots \otimes B_k)$, representing the use of the correlated device with all ports used simultaneously.

Similarly, the sequential placement places the device $\mathcal{N}_{1,\ldots,k} \in \mathsf{NSChan}(A_1 \to B_1, \ldots, A_k \to B_k)$ as:

$$\mathcal{S}_{\text{seq}}^{\mathbf{A},\mathbf{B}}(\mathcal{N}_{1,\ldots,k}) := (\mathcal{S}_{\text{place}}^{A_1,B_1} \otimes \cdots \otimes \mathcal{S}_{\text{place}}^{A_k,B_k})(\mathcal{N}_{1,\ldots,k}) \,, \tag{4.16}$$

where $\mathbf{A} := (A_1, \ldots, A_k)$ $[\mathbf{B} := (B_1, \ldots, B_k)]$ is a list of quantum systems placed in Alice's (Bob's) laboratory, with $A_i \simeq X_i$ and $B_i \simeq Y_i$ for every $i \in \{1, \ldots, k\}$, where in particular, the system A_1 is sent before system A_2, which is sent before system A_3, etc. As before, the result of the supermap is a placed k-step quantum channel in $\mathsf{Chan}(A_1 \to B_1, \cdots, A_k \to B_k)$, representing the use of the correlated device with the ports used in a sequential order.

4.1.8 Terminology

In the rest of this work, the study of communication protocols involving only parallel placement between a sender and a receiver shall be called *standard quantum Shannon theory for direct communication*. The study of communication protocols involving both parallel and sequential placements between a sender, a receiver, and intermediate parties will be called *standard quantum Shannon theory for network communication*, or simply, *standard quantum Shannon theory*. We shall not consider assisted scenarios, such as entanglement-assisted communication, further in this work.

4.2 General resource theories of communication

Here we extend the framework of standard quantum Shannon theory to general resource theories of communication, arguing that any such theory must not include operations that enable communication independently of the communication devices initially available to the communicating parties.

4.2.1 Basic structure

The resource-theoretic formulation of standard quantum Shannon theory, discussed in the previous section, suggests a general scheme for constructing new resource theories of communication. The basic scheme is as follows:

1. One use of each communication device is described by an *unplaced quantum channel*, specifying how a given system type is transformed into another system type, but without assigning these system types to specific locations in spacetime. These system types are called *unplaced systems*.

2. The (uses of the) available communication devices are described by a *list of unplaced quantum channels*.

3. The sender, receiver, and possibly a set of intermediate parties are assigned spacetime regions, whose causal structure specifies who can send messages to whom. The physical systems accessed by the communicating parties are *placed systems*, that is, systems assigned specific locations in spacetime.

4. The placement of the communication devices in between the communicating parties is described by a *placement supermap*, that is, a supermap transforming lists of unplaced quantum channels into *placed quantum channels*. A placed quantum channel has placed systems as inputs and outputs, and can in general be a multistep process, represented by a multi-step quantum channel.

5. The operations performed by the sender, receiver, and intermediate parties are described by a *party supermap*, that is, a supermap on the set of placed quantum channels.

In the above scheme, a resource theory of communication is formulated by specifying which operations are considered as 'free' in points 4 and 5 above.

Free operations on placed channels (party supermaps) are interpreted as being implemented by the sender, the receiver, or intermediate parties. Free operations from unplaced to placed channels (placement supermaps) are interpreted as being performed by an external agent, e.g. a communication provider, or Nature itself. This is consistent with the intuitive idea that a communication infrastructure has to be set up before communication takes place. Overall, a resource theory of communication describes the actions performed by the communicating parties and by an external agent that places the communication devices between them.

In principle, one could also consider a third type of operations, from unplaced channels to unplaced channels. These operations would be performed by the third party *before* the channels are placed between the sender and receiver. For completeness, we shall include the possibility of these 'pre-placement operations' in our general scheme.

For example, the third party could decide to discard one of the devices in the list $(\mathcal{N}_1, \ldots, \mathcal{N}_k)$, and use only the remaining devices. This operation is given by the *discarding supermap*

$$\mathcal{S}^m_{\text{discard}}(\mathcal{N}_1, \ldots, \mathcal{N}_k) := (\mathcal{N}_1, \ldots, \mathcal{N}_{m-1}, \mathcal{N}_{m+1}, \ldots, \mathcal{N}_k)\,, \tag{4.17}$$

which discards the m-th channel from a list of k channels. Intuitively, the discarding supermap should always be included in the set of free supermaps, as the communication provider can always decide to discard a communication device in the construction of a communication network. We say that the set of free supermaps from unplaced channels to unplaced channels is *trivial* if it consists only of discarding supermaps and identity supermaps.

4.2.2 Resource theories of communication

For a resource theory of communication, the broader set of operations M from which the free operations M_{free} are chosen consists of (1) supermaps from unplaced channels to unplaced channels, (2) supermaps from unplaced channels to placed channels, and (3) supermaps from placed channels to placed channels. A detailed mathematical characterisation of all valid quantum supermaps of these three types is given in Appendix B of Ref. [119].

A resource theory of communication is then specified by fixing the set of free operations:

Definition 9. (Resource theory of communication.) *A resource theory of communication is specified by a set of free supermaps* $\mathsf{M}_{\text{free}} \subset \mathsf{M}$*, closed under sequential and parallel composition, containing (1) free supermaps from unplaced channels to unplaced channels, called* pre-placement supermaps*, (2) free supermaps from unplaced channels to placed channels, called* placement supermaps*, and (3) free supermaps from placed channels to placed channels, called* party supermaps*.*

In diagrams, we represent the placement supermaps by blue boxes, and the party supermaps by violet boxes.

Mathematically, the different channel types are objects in a symmetric monoidal category, and the free operations M_{free} correspond to the morphisms between them. This scheme matches the general framework of Coecke, Fritz and Spekkens [67].

The set M_{free} can be specified by a generating set of operations [67]. For example, standard quantum Shannon theory is the resource theory of communication where the free operations $\mathsf{M}_{\text{standard}}$ are generated from the following types of free operations:

(i) *Basic placement:* For a single channel $\mathcal{N} \in \mathsf{Chan}(X \to Y)$, the map

$$\mathcal{S}_{\text{place}}^{A,B}(\mathcal{N}) := \mathcal{W}^B \circ \mathcal{N} \circ \mathcal{V}^A,$$

where $\mathcal{V} \in \mathsf{Chan}(A \to X)$ $[\mathcal{W} \in \mathsf{Chan}(Y \to B)]$ is the unitary channel implementing the isomorphism between the unplaced system X (Y) and the placed system A (B).

(ii) *Insertion of local devices:* For l placed channels $\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_l \in \mathsf{Chan}[(A \to R_1), (R_1' \to R_2), \ldots, (R_{l-1}' \to B)]$, the *encoding map*

$$\mathcal{S}_{\mathcal{E}}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_l) := (\mathcal{C}_1 \circ \mathcal{E}) \otimes \mathcal{C}_2 \otimes \cdots \otimes \mathcal{C}_l\,, \tag{4.18}$$

the *repeater map*

$$\mathcal{S}_{\mathcal{R}_m}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_l) := \mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_{m-1} \otimes (\mathcal{C}_{m+1} \circ \mathcal{R}_m \circ \mathcal{C}_m) \otimes \mathcal{C}_{m+2} \otimes \cdots \otimes \mathcal{C}_l\,, \tag{4.19}$$

and the *decoding map*

$$\mathcal{S}_{\mathcal{D}}(\mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_l) := \mathcal{C}_1 \otimes \cdots \otimes \mathcal{C}_{l-1} \otimes (\mathcal{D} \circ \mathcal{C}_l)\,, \tag{4.20}$$

where $\mathcal{E} \in \mathsf{Chan}(A' \to A)$, $\mathcal{R}_m \in \mathsf{Chan}(R_m \to R_m')$, and $\mathcal{D} \in \mathsf{Chan}(B \to B')$ are quantum channels representing local devices at the sender's, m-th repeater's, and receiver's end, respectively.

Note that we omitted pre-placement supermaps, because the set of such supermaps is trivial in standard quantum Shannon theory.

The other supermaps shown earlier in Section 4.1 can be decomposed into the basic supermaps (i)–(ii). For example, the parallel placement (4.4) and sequential placement (4.8) are just the product of basic placement supermaps (i), which place individual channels in the appropriate configuration. Similarly, the encoding-decoding supermap (4.3) and the encoding-repeater-decoding-supermap (4.7) are just the result of multiple insertions of local devices (ii).

4.2.3 Generalised channel capacities

In standard quantum Shannon theory, the classical (quantum) capacity of a quantum channel $\mathcal{N}$ is defined as the maximum number of bits (qubits) that can be transmitted over k parallel uses of $\mathcal{N}$, per channel use and with vanishing error in the asymptotic limit $k \to \infty$. This is equivalent to the maximum number of classical (quantum) identity channels $\mathcal{I}_{\text{clas}}$ ($\mathcal{I}$) that the k parallel uses of $\mathcal{N}$ can simulate, per channel use and with vanishing error in the asymptotic limit $k \to \infty$, using arbitrary encoding/decoding channels [196]. (The classical identity channel $\mathcal{I}_{\text{clas}}$ is defined as the perfectly dephasing channel (cf. §2.2.3) with respect to a given orthonormal basis).

The standard definition of classical (quantum) capacity is appropriate for placed channels, which have already been arranged in between the sender and receiver, and therefore can only be used in parallel. However, unplaced channels could be arranged in more general configurations, generating a broader class of communication protocols.

In a general resource theory of communication, we define the *generalised classical (quantum) capacity* of a quantum channel $\mathcal{N}$ as the maximum number of classical (quantum) identity channels $\mathcal{I}_{\text{clas}}$ ($\mathcal{I}$) that can be generated by performing free operations of M_{free}, per channel use and with vanishing error in the asymptotic limit of $k \to \infty$ channel uses. Other types of generalised capacities can be defined similarly, with respect to some given ideal reference channel.

The generalised capacity is (trivially) a resource monotone [60, 67], meaning that it cannot be increased by applying free operations. Moreover, the generalised capacity increases (or stays the same) whenever the set of free operations is enlarged. Examples of this situation are the capacity enhancements observed in the presence of quantum control over the causal orders [58, 80, 94, 161, 162, 168]: in these protocols, the set of placements of standard quantum Shannon theory is enlarged to include placements in a superposition of alternative orders, and consequently various channel capacities have been shown to increase.

4.2.4 A minimal requirement for any resource theory of communication

Formally, every set of free supermaps defines a resource theory of communication. However, such a resource theory may not be a meaningful one. We argue that every meaningful resource theory of communication should at least satisfy a minimal requirement: the free operations should not allow the sender and receiver to communicate independently of the communication devices from which their communication protocol is built.

To illustrate this idea, consider the situation where two parties, Alice and Bob, communicate through a noisy telephone line. In the standard theory of communication, the key question is how to use this communication resource to transmit information reliably. Now, if Alice were to walk into Bob's lab, he would clearly be able to hear her through the air, but this would not be a new way to use the telephone line. Rather, it would be a way to bypass it. The air would act as a side-channel, allowing Alice and Bob to communicate to each other independently of how good or how bad their telephone line is.

The telephone line example has the following structure. Initially, Alice and Bob have access to a noisy communication channel $\mathcal{N} \in \mathsf{Chan}(A \to B)$. The operation of Alice moving into Bob's lab can be modelled as a *side-channel supermap*

$$\begin{aligned} \mathcal{S}_{\text{side}}^{(E,E')}: \quad & \mathsf{Chan}(A \to B) \to \mathsf{Chan}(A \otimes E \to B \otimes E') \\ & \mathcal{S}(\mathcal{N}) : \mathcal{N} \mapsto \mathcal{N} \otimes \mathcal{I}_{E,E'} \,, \end{aligned} \tag{4.21}$$

which juxtaposes the noisy channel $\mathcal{N}$ with a side-channel $\mathcal{I}_{E,E'} \in \mathsf{Chan}(E \to E')$ acting on some additional systems E and E' (the air in the proximity of Alice and Bob, respectively). If the channel $\mathcal{I}_{E,E'}$ is ideal, then the supermap $\mathcal{S}_{\text{side}}^{(E,E')}$ would let Alice communicate perfectly to Bob. This communication 'enhancement', however, is independent of the original channel $\mathcal{N}$. Every operation of the form (4.21) trivialises the notion of communication enhancement, and therefore should not be allowed in a resource theory of communication.

Building on the above example, we now propose a general notion of a side-channel generating operation:

Definition 10. (Side-channel generating operations.) *A supermap $\mathcal{S} \in \mathsf{M}$* generates a classical (quantum) side-channel *if there exist two free supermaps $\mathcal{S}_1 \in \mathsf{M}_{\text{free}}$ and $\mathcal{S}_2 \in \mathsf{M}_{\text{free}}$ and a placed quantum channel $\mathcal{C}$ with non-zero classical (quantum) capacity, such that, for all choices of input channels*

$(\mathcal{N}_1, \ldots, \mathcal{N}_k)$ *for supermap* $\mathcal{S}_1$, *we have that*

$$(\mathcal{S}_2 \circ \mathcal{S} \circ \mathcal{S}_1)(\mathcal{N}_1, \ldots, \mathcal{N}_k) = \mathcal{C}\,. \tag{4.22}$$

The above definition captures the idea that the supermap $\mathcal{S}$ can be used to construct a communication protocol that works independently of the communication devices originally available to the communicating parties. In the telephone line example, the channel $\mathcal{C}$ is the ideal channel $\mathcal{I}_{E,E'}$ describing the transmission of a message through the air between Alice and Bob.

We demand that any sensible resource theory of communication should forbid side-channel generating operations:

Condition 11. (No Side-Channel Generation.) *In a resource theory of classical (quantum) communication, no free operation* $\mathcal{S} \in \mathsf{M}_{\text{free}}$ *should generate a classical (quantum) side-channel.*

We stress that Condition 11 is a *minimal* requirement, and that, in particular cases, one may want to impose even stronger conditions on the allowed operations. In other words, we are not claiming that every resource theory of communication satisfying Condition 11 is an interesting one. Rather, Condition 11 is a bottom line that has to be satisfied when defining new resource theories of communication.

It is immediate to verify that standard quantum Shannon theory satisfies Condition 11. In the following, we shall show that

1. quantum Shannon theory with superpositions of causal orders satisfies Condition 11
2. quantum Shannon theory with superpositions of trajectories satisfies Condition 11
3. quantum Shannon theory with superpositions of encoding and decoding operations violates Condition 11.

In §4.5, we comment on the difference between our framework and the frameworks of Refs. [131, 132, 186], discussing an alternative to Condition 11, where the free supermaps are required to transform constant channels into constant channels.

4.3 Superposition of causal orders and superposition of trajectories

Here we formulate the resource theories of quantum Shannon theory with superpositions of causal orders and quantum Shannon theory with superpositions of trajectories, and we show that both theories satisfy the requirement of No Side-Channel Generation.

4.3.1 Quantum Shannon theory with superpositions of causal orders

As discussed in §2.4.1, indefinite causal order has been shown to provide various advantages in quantum information processing, such as in quantum query complexity [11, 47], quantum communication complexity [95], quantum metrology [201] and non-local games [149]. In all the above works, the combination of quantum devices in an indefinite causal order was shown to offer performances that cannot be matched by any quantum protocol that uses the input devices in a definite order.

A different category of advantages arises in the context of quantum communication [58, 59, 80, 94, 133, 161, 162, 168]. Here, protocols that combine communication channels through the quantum SWITCH have been shown to offer advantages with respect to the protocols allowed in standard quantum Shannon theory, as defined earlier in this chapter. These advantages are *not* advantages with respect to all possible protocols with definite causal order. They cannot be so, because the set of all protocols with definite causal order includes also trivial protocols where the original communication channels are juxtaposed with noiseless channels, as in the telephone line example of Equation (4.21).

The proper way to interpret the communication advantages shown in Refs. [58, 59, 80, 94, 133, 161, 162, 168] is to regard them as a comparison between two different resource theories of communication: standard quantum Shannon theory, and an extended resource theory that includes the quantum SWITCH among its placements.

Here we explicitly define such a resource theory, which we call quantum Shannon theory with superpositions of causal orders (SCO). The corresponding set of free operations shall be denoted by $\mathsf{M}_{\mathrm{SCO}}$. The generating free operations are operations (i)–(ii) of standard quantum Shannon theory, plus an additional placement supermap, based on the quantum SWITCH:

(iii) The *quantum* SWITCH *placement* $\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}$ maps a pair of unplaced quantum channels $(\mathcal{N}_1, \mathcal{N}_2) \in \mathsf{Chan}(X) \times \mathsf{Chan}(X)$ into a placed quantum channel $\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2) \in \mathsf{Chan}(A \to B \otimes O)$, where $A \simeq X$ ($B \simeq X$) is a quantum system placed at the sender's (receiver's) end, and O is a qubit system, called the *order qubit*, placed at the

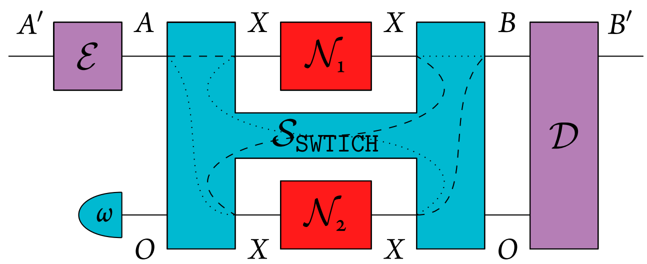


Figure 4.8: *Communication through the quantum* SWITCH. The quantum SWITCH placement $\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}$ (in blue) places two quantum channels ($\mathcal{N}_1$, $\mathcal{N}_2$) in a superposition of causal orders, determined by the fixed state $\omega \in \mathsf{St}(O)$, between a sender and receiver, and is followed by the encoding-decoding supermap $\mathcal{S}_{\mathcal{E},\mathcal{D}}$ (in violet). The dashed and dotted lines illustrate the two alternative orders of applying $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively.

receiver's end. Explicitly, the quantum channel $\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)$ is defined as

$$\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)(\rho) = \sum_{i,j} S_{ij}(\rho \otimes \omega) S_{ij}^{\dagger}\,, \tag{4.23}$$

where $\omega \in \mathsf{St}(O)$ is a state of the order qubit, and

$$S_{ij} := \mathrm{N}_i^{(2)} \mathrm{N}_j^{(1)} \otimes |0\rangle\langle 0| + \mathrm{N}_j^{(1)} \mathrm{N}_i^{(2)} \otimes |1\rangle\langle 1|\,, \tag{4.24}$$

$\{|0\rangle, |1\rangle\}$ being an orthonormal basis for the order qubit.

The quantum channel $\mathcal{S}^{A,B,\omega}_{\mathtt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)$ is independent of the Kraus decomposition of the channels $\mathcal{N}_1$ and $\mathcal{N}_2$.

A communication protocol using the quantum SWITCH placement is given in Figure 4.8. Note that the initial state of the order qubit is fixed as part of the placement, and is thus inaccessible to the sender [80, 168].

We stress that the quantum SWITCH placement should be understood here as an abstract supermap from two quantum channels to a new quantum channel. Whether this supermap can be physically realised, and how it can be realised, is another matter, which we do not consider in this work. Various ways to reproduce the action of the quantum SWITCH have been proposed, using conventional physics [93, 94, 97, 160, 165, 185], closed timelike curves [57], or quantum gravity scenarios [152, 203]. However, the resource theory $\mathsf{M}_{\mathrm{SCO}}$ should be considered as the abstract resource theory associated with the quantum SWITCH transformation, without reference to a specific physical implementation.

The motivation for including the quantum SWITCH among the free operations is to understand how the world *could* be, if quantum devices could be combined in a superposition of alternative orders. The study of quantum Shannon theory with the addition of the quantum SWITCH is similar in spirit the study of information tasks assisted by the Popescu-Rohrlich box [158], a fictional device that generates stronger than quantum correlations. Like the Popescu-Rohrlich box, the quantum SWITCH serves as a conceptual device, used to better understand standard quantum theory by comparing it to possible alternatives.

4.3.2 Quantum Shannon theory with superpositions of trajectories

The superposition of alternative evolutions was defined in Refs. [3, 4, 9, 146], and applied to quantum communication in Refs. [91, 123], where the ability to send quantum particles along a superposition of different trajectories provided the working principle for a new technique called error filtration. Shannon-theoretic advantages of the superposition of trajectories are demonstrated in Chapters 5, 6 and 8, as well as independently in Refs. [2, 133].

Here, we formulate the resource theory of quantum Shannon theory with superpositions of trajectories (ST), grounding the advantages presented in Chapters 5, 6 and 8 on a formal footing. The set of free operations in this resource theory, denoted by M_{ST}, is generated by the standard free operations (i)–(ii), with the addition of a *superposition placement* (iii*), based on the superposition of channels specified by vacuum extensions (cf. Definition 7), which creates a superposition of two alternative communication channels.

In order to define the superposition placement, we model the communication hardware by vacuum-extended quantum channels (cf. §3.3). Vacuum-extended channels represent communication devices that can act on the information carrier, or on the vacuum, or on any coherent superposition of the two. Using this feature, it is possible to coherently control the choice of channel through which the information carrier is sent. The result can be interpreted as a placement of the given different channels in a superposition of being on the path of the information carrier:

(iii*) The *superposition placement* $\mathcal{S}_{\mathrm{sup}}^{A,B,\omega}$ maps a pair of unplaced vacuum-extended channels $(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2) \in \mathsf{Chan}(\widetilde{X}) \times \mathsf{Chan}(\widetilde{X})$ into a placed quantum channel $\mathcal{S}_{\mathrm{sup}}^{A,B,\omega}(\mathcal{N}_1, \mathcal{N}_2) \in \mathsf{Chan}(A \to B \otimes P)$, where $A \simeq X$ $(B \simeq X)$ is a quantum system placed at the sender's (receiver's) end, and P is a qubit system, called the *path qubit*, placed at the receiver's end. Explicitly, the quantum channel $\mathcal{S}_{\text{sup}}^{A,B,\omega}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)$ is defined as

$$\mathcal{S}_{\text{sup}}^{A,B,\omega}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)(\rho) = \mathcal{U}^\dagger \circ (\mathcal{N}_1 \otimes \mathcal{N}_2) \circ \mathcal{U}(\rho \otimes \omega) \,, \tag{4.25}$$

where $\mathcal{U}$ is the isomorphism defined in Eq. (3.16) and $\omega \in \mathsf{St}(P)$ is a state of the path qubit.

Note, that the definition of the supermap describing the superposition placement (4.25) differs from the definition of the supermap describing the superposition of channels specified by vacuum extensions (3.14), in that the output of the former is a channel in $\mathsf{Chan}(A \to B \otimes P)$, whilst the output of the latter is a channel in $\mathsf{Chan}(A \otimes P \to B \otimes P)$. In other words, the initial state of the path qubit is considered part of the superposition placement. This is because in a physically motivated resource theory, the choice of path qubit should be considered part of the placement of the communication devices. If, instead, the path qubit were under the control of the sender, then the distinction between the path and the message disappears, and quantum Shannon theory with superpositions of trajectories would reduce simply to standard quantum Shannon theory with a larger message Hilbert space. Thus, the paradigm of quantum Shannon theory with superpositions of trajectories should be seen as exploring the Shannon-theoretic benefits of using a control system purely to combine alternative channels in a superposition, but where this control system is inaccessible to the sender.

A simple experimental setup implementing the superposition placement is that of a single-photon generator and a Mach-Zehnder interferometer, as described in §3.4.1. The two vacuum-extended channels are placed one on each arm of the interferometer. The sender controls a polarisation-shifter, which is placed between the single-photon generator and the input to the interferometer, determining the initial state of the message ρ. The parameters of the first beamsplitter in the interferometer determine the value of the path qubit ω. This results in a clear distinction between the placement of the path qubit, and the encoding of the message by the sender. The output state of the interferometer is then given by the composite message-path output state of the superposition channel (4.25).

An example of a communication protocol using the superposition placement is shown in Figure 4.9. Since the superposition placement is physically implementable, for example, in photonic systems, the resource theory M_{ST} is interesting both from a purely information-theoretic point of view as well as from a practical point of view.

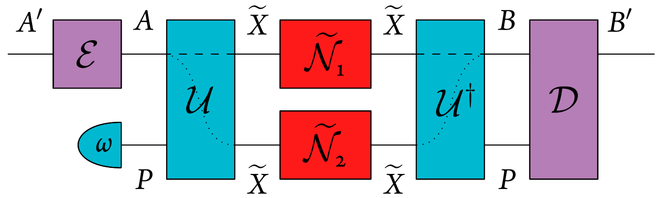


Figure 4.9: *Communication through a superposition of quantum channels.* The supermap $\mathcal{S}_{\text{sup}}^{A,B,\omega}$ (in blue) places two vacuum-extended channels $(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)$ on two alternative paths, and lets the transmitted system travel along both paths (dashed line and dotted line, respectively) in a quantum superposition, determined by the state $\omega \in \mathsf{St}(P)$. The resulting channel then undergoes the encoding-decoding supermap $\mathcal{S}_{\mathcal{E},\mathcal{D}}$ (in violet), describing the local operations performed at the sender's and receiver's ends.

The superposition placement can, just as the parallel and sequential placements, also be applied to correlated devices, described by no-signalling quantum channels. In particular, we can describe the transmission of a single particle at a superposition of different times, travelling through a time-correlated channel (cf. §3.6). Explicitly, the action of the superposition placement on a vacuum-extended unplaced no-signalling channel $\widetilde{\mathcal{N}}_{1,2} \in \mathsf{NSChan}(\widetilde{X}_1 \to \widetilde{Y}_1, \widetilde{X}_2 \to \widetilde{Y}_2)$ is given by

$$\mathcal{S}_{\text{sup}}^{A,B,\omega}(\widetilde{\mathcal{N}}_{1,2})(\rho) = \mathcal{U}^\dagger \circ \mathcal{N}_{1,2} \circ \mathcal{U}(\rho \otimes \omega)\,, \tag{4.26}$$

where $\mathcal{U}$ is the isomorphism defined in Eq. (3.16), resulting in a placed quantum channel in $\mathsf{Chan}(A \to B \otimes C)$, where $A \simeq X$ ($B \simeq X$) is a quantum system placed at the sender's (receiver's) end, and C is a control qubit, initialised in state $\omega \in \mathsf{St}(C)$, placed at the receiver's end.

Formally, Eq. (4.26) is based on the superposition of ports of a bipartite channel (cf. Definition 8), and could thus in principle be extended to take as input any 2-step channel, or even any bipartite channel. However, for the purposes of this work, we shall stick with considering unplaced channels as no-signalling channels only; in the applications presented in later chapters, our examples of correlated channels are all no-signalling channels.

For simplicity of presentation, here we considered only superpositions of two channels, both for the superposition of trajectories and for the superposition of orders. Both the superposition placement (iii*) and the quantum `SWITCH` placement (iii) can be straightforwardly generalised to N channels. The corresponding definitions can be found in Chapter 3 and Ref. [68], respectively.

4.3.3 Superpositions of causal orders and superpositions of trajectories do not generate side-channels

We now show that the supermaps (iii) or (iii*), combined with (i)–(ii), do not generate side-channels.

Proposition 12. *No supermap composed from the quantum* SWITCH *placement* (iii)*, basic placement* (i)*, and insertion of local devices* (ii) *generates side-channels.*

Proof. The supermaps (i)–(ii) of standard quantum Shannon theory do not generate side channels. Hence, it is sufficient to prove that the quantum SWITCH does not generate side-channels.

This is done by finding a choice of adversarial channels $\mathcal{N}_1$ and $\mathcal{N}_2$ such that $\mathcal{S}^{A,B,\omega}_{\texttt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)$ is a channel with zero classical capacity. One such choice is to pick $\mathcal{N}_1$ to be the identity channel $\mathcal{I}$ (with a single Kraus operator $N_i^{(1)} = I, i = 1$) and $\mathcal{N}_2$ to be the constant channel $\mathcal{N}_2(\rho) = |\psi_0\rangle\langle\psi_0|$ (with Kraus operators $N_j^{(2)} = |\psi_0\rangle\langle j|$, for some orthonormal basis $\{|j\rangle\}$). With this choice, the Kraus operators (4.24) in the definition of the channel $\mathcal{S}^{A,B,\omega}_{\texttt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)$ are

$$S_{ij} = |\psi_0\rangle\langle j| \otimes I \,, \tag{4.27}$$

and therefore one has

$$\mathcal{S}^{A,B,\omega}_{\texttt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)(\rho) := \sum_{i,j} S_{ij}(\rho \otimes \omega) S_{ij}^\dagger = |\psi_0\rangle\langle\psi_0| \otimes \omega \quad \forall \rho \in \mathsf{St}(A) \,. \tag{4.28}$$

Since the output of the channel $\mathcal{S}^{A,B,\omega}_{\texttt{SWITCH}}(\mathcal{N}_1, \mathcal{N}_2)$ is independent of its input, the channel has zero capacity (both classical and quantum), and no combination of it with the other supermaps (i)–(ii) can generate a channel with non-zero capacity. □

Proposition 13. *No supermap composed from the superposition placement* (iii*)*, basic placement* (i)*, and insertion of local devices* (ii) *generates side-channels.*

Proof. As in the proof of Proposition 12, it is sufficient to prove that the superposition placement (iii*) does not generate side-channels. This is done by finding a choice of adversarial vacuum-extended channels $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$ such that $\mathcal{S}^{A,B,\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)$ is a channel with zero classical capacity. One such choice is to pick the vacuum-extended channels $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$ defined by

$$\widetilde{\mathcal{N}}_1(\rho) = \widetilde{\mathcal{N}}_2(\rho) = \rho_0 \operatorname{Tr}[\rho\,(I - |\text{vac}\rangle\langle\text{vac}|)] + |\text{vac}\rangle\langle\text{vac}|\rho|\text{vac}\rangle\langle\text{vac}| \quad \forall \rho \in \mathsf{St}(\widetilde{X}) \,. \tag{4.29}$$

In other words, $\widetilde{\mathcal{N}}_1 = \widetilde{\mathcal{N}}_2$ is the incoherent vacuum-extension (cf. Definition 6) of the constant channel that maps every state into the fixed state ρ_0. For the vacuum-extended channels $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$, the vacuum interference operators are $F_1 = F_2 = 0$, and the superposition placement then yields the channel

$$\mathcal{S}_{\text{sup}}^{A,B,\omega}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)(\rho) = \rho_0 \otimes \text{diag}(\omega) , \tag{4.30}$$

with $\text{diag}(\omega) := \langle 0|\omega|0\rangle\, |0\rangle\langle 0| + \langle 1|\omega|1\rangle\, |1\rangle\langle 1|$, as one can verify from Equation (3.17). Since the channel $\mathcal{S}_{\text{sup}}^{A,B,\omega}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)$ is constant, it has zero (classical and quantum) capacity. □

Propositions 12 and 13 show that both quantum Shannon theory with superpositions of causal orders and quantum Shannon theory with superpositions of trajectories satisfy the requirement of No Side-Channel Generation, as stated in Condition 11.

4.4 Superpositions of encoding and decoding operations

A recent paper by Guérin, Rubino, and Brukner [96] argues that in order to claim meaningful communication advantages, the quantum `SWITCH` should be compared to a general class of operations termed 'superpositions of direct pure processes'. In this section, we analyse their arguments and examples, concluding that they rest on a communication model that violates the basic resource-theoretic framework.

4.4.1 The framework of SDPPs

The authors of Ref. [96] argue that quantum Shannon theory with superpositions of causal orders should be considered within a general framework of 'superpositions of direct pure processes' (SDPPs), which includes the quantum `SWITCH`: '*It seems that any reasonable resource theory that contains the quantum switch—a superposition of direct pure processes with different causal orders—should also allow superpositions of direct pure processes with the same causal order*' [96]. It is claimed, therefore, that the advantages of the quantum `SWITCH` should be compared to SDPPs with a definite causal order. In the following, we analyse the above claim, showing that, while the quantum `SWITCH` and the SDPPs considered in Ref. [96] share a similar mathematical structure, they have different operational features: in particular, the specific SDPPs compared with the quantum `SWITCH` in Ref. [96] generate side-channels, making the proposed advantages trivial from the resource-theoretic point of view.

In the language of this work, the SDPPs of Ref. [96] are supermaps that take two channels $(\mathcal{N}_1, \mathcal{N}_2)$, and return a superposition of k channels (cf. §3.2) which are individually of the form $\mathcal{D}_j \circ \mathcal{N}_2 \circ \mathcal{R}_j \circ \mathcal{N}_1 \circ \mathcal{E}_j$ or $\mathcal{D}'_j \circ \mathcal{N}_1 \circ \mathcal{R}'_j \circ \mathcal{N}_2 \circ \mathcal{E}'_j$, $j \in \{1, 2, \ldots, k\}$, for some encoding, repeater and decoding operations $\mathcal{E}_j, \mathcal{E}'_j, \mathcal{R}_j, \mathcal{R}'_j, \mathcal{D}_j$ and $\mathcal{D}'_j$. SDPPs with a definite causal order are defined as SDPPs which are superpositions of terms where the input channels $\mathcal{N}_1$ and $\mathcal{N}_2$ occur in the same, fixed order.

In the resource-theoretic scheme of this chapter, the SDPPs should be regarded as a set of free operations. One could, for example, consider them as a broader set of party supermaps, or alternatively, as a set of placement supermaps with internal encoding, decoding, and repeater operations, which are in a quantum superposition controlled by some quantum degree of freedom that is part of the placement.

The resource theory based on SDPPs is different from the resource theory of quantum Shannon theory with superpositions of trajectories. An important difference is that SDPPs are supermaps acting directly on the original channels, rather than their vacuum extensions. In this respect, SDPPs and the quantum `SWITCH` operate on the same type of input resource, and a comparison between them would indeed be even. However, in the following we shall show that building a resource theory of communication where all SDPPs are taken as free operations is problematic, because it violates the requirement of No Side-Channel Generation. This fact will be illustrated by analysing the specific examples of SDPPs proposed in Ref. [96].

4.4.2 Some SDPPs generate classical side-channels

One of the SDPPs proposed in Ref. [96] is the supermap depicted in Figure 4.10. This supermap corresponds to a protocol where two qubit channels $\mathcal{N}_1$ and $\mathcal{N}_2$ are applied after a `CNOT` gate, acting on the message qubit and on an additional control qubit C. The supermap, here denoted as $\mathcal{F} : (\mathcal{N}_1, \mathcal{N}_2) \mapsto \mathcal{F}(\mathcal{N}_1, \mathcal{N}_2)$, produces the output channel defined by

$$\mathcal{F}(\mathcal{N}_1, \mathcal{N}_2)(\rho) = [(\mathcal{N}_2 \circ \mathcal{N}_1) \otimes \mathcal{I}_C] \circ \mathcal{U}^{\texttt{CNOT}}(\rho \otimes |+\rangle\langle+|), \tag{4.31}$$

where $\mathcal{U}^{\texttt{CNOT}} := U^{\texttt{CNOT}\dagger}(\cdot)U^{\texttt{CNOT}}$ is the unitary channel corresponding to the `CNOT` gate

$$U^{\texttt{CNOT}} := I_M \otimes |0\rangle\langle 0|_C + X_M \otimes |1\rangle\langle 1|_C \ , \tag{4.32}$$

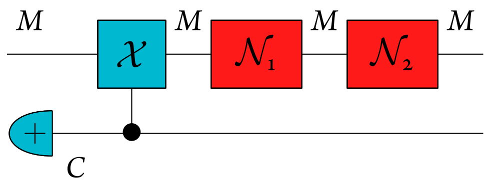


Figure 4.10: *An SDPP that transfers classical information through a side-channel, bypassing the original communication devices.* A control qubit C is prepared in the state $|+\rangle$ and is sent together with the message M through a CNOT gate. If the message is initialised in either of the states $|\pm\rangle$, then its interaction with the control through the CNOT will output the state $|\pm\rangle$ in C. The receiver is thus able to decode the original message by measuring the control qubit, irrespectively of the channels $\mathcal{N}_1$ and $\mathcal{N}_2$. Overall, this SDPP fails to satisfy the requirement of No Side-Channel Generation, which we regard as a minimal requirement for a sensible resource theory of communication.

X being the NOT gate. To avoid overloading the notation, here we have omitted the isomorphisms between unplaced and placed systems, and simply denoted the (placed and unplaced) message system by M.

Now, the map $\mathcal{F}$ enables perfect classical communication of one bit independently of the communication channels $\mathcal{N}_1$ and $\mathcal{N}_2$ [96]. Using the phase kickback mechanism of the CNOT gate [62], information encoded in the states $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$ is transferred from the message M to the control C *before* the noisy channels $\mathcal{N}_1$ and $\mathcal{N}_2$ are applied. Then, the information is safely carried by the control system to the receiver, completely bypassing the communication channels $\mathcal{N}_1$ and $\mathcal{N}_2$, and avoiding the resulting noise. In other words, this example of an SDPP is analogous to the example of the noisy telephone line discussed in §4.2.4: it achieves communication by completely bypassing the original channels.

More formally, one can see that the operation $\mathcal{F}$ generates a classical side-channel in the sense of Definition 10. Indeed, one can consider the party supermap corresponding to the encoding channel $\mathcal{E} = \mathcal{I}_M$ and the decoding channel $\mathcal{D} = \mathrm{Tr}_M$, which discards the message qubit. The result is the channel

$$\begin{aligned}\mathcal{C}(\cdot) &= \mathcal{D} \circ \mathcal{F}(\mathcal{N}_1, \mathcal{N}_2) \circ \mathcal{E}(\cdot) \\ &= |+\rangle\langle+| \cdot |+\rangle\langle+| + |-\rangle\langle-| \cdot |-\rangle\langle-| \ ,\end{aligned} \tag{4.33}$$

which is independent of $\mathcal{N}_1$ and $\mathcal{N}_2$ and provides a perfect transmission line for classical communication. In conclusion, the 'communication enhancement' of the SDPP (4.31) arises from a classical

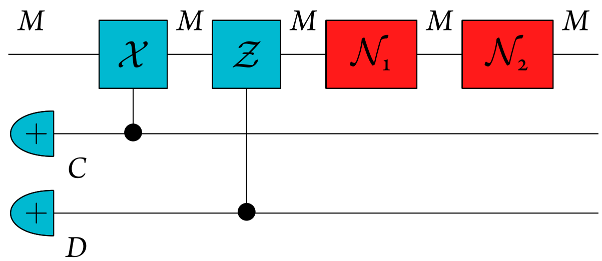


Figure 4.11: *An SDPP that transfers quantum information through a side-channel, for any noisy channels acting on the message.* Two control qubits C and D are both prepared in the state $|+\rangle$. A message M in state ρ is to be communicated. The composite system $M \otimes C$ is sent through a `CNOT` gate, followed by the composite system $M \otimes D$ going through a `CPHASE` gate. As shown in Equation (4.36), the receiver is able to recover the original input by measuring the control qubit D and performing a conditional correction on C, independently of the choice of noisy channels $\mathcal{N}_1$ and $\mathcal{N}_2$ that act on the message itself.

side-channel, which completely bypasses the original communication devices.

The authors of Ref. [96] also consider a specific interferometric implementation of the operation $\mathcal{F}$, which they claim avoids the criticism above. In Appendix G of Ref. [119] (the contents of which are omitted in this thesis for brevity), the arguments of Ref. [96] are analysed, concluding that, in fact, the above criticism still applies.

4.4.3 Some SDPPs generate quantum side-channels

In Appendix B of Ref. [96], the authors present an SDPP, stating that it *'allows us to perfectly transmit one qubit of quantum information, for all channels [...]'*. This statement is an explicit acknowledgement that the SDPP model permits the strongest possible kind of side-channels: perfect side-channels for quantum communication.

The example in Appendix B of Ref. [96] is presented as one that *'generalises, and improves upon, the observations made in the main text,'* the improvement being that only four control qubits are used instead of eight, which is the number of qubits used by the protocol in the main text. Here we review the example, showing that, in fact, one can improve it even further: the same perfect qubit side-channel can be generated by an SDPP that uses only two control qubits, instead of four.

Our improved version of the SDPP in Ref. [96] is depicted in Figure 4.11. It uses two control qubits C and D, in addition to the message qubit M. The corresponding supermap is given by

$$\begin{aligned}
&\mathcal{G}^{\omega,\xi} : (\mathcal{N}_1, \mathcal{N}_2) \mapsto \mathcal{G}^{\omega,\xi}(\mathcal{N}_1, \mathcal{N}_2) \\
&\mathcal{G}^{\omega,\xi}(\mathcal{N}_1, \mathcal{N}_2)(\rho_M) = [(\mathcal{N}_2 \circ \mathcal{N}_1) \otimes \mathcal{I}_C \otimes \mathcal{I}_D] \circ (\mathcal{U}^{\texttt{CPHASE}}_{MD} \otimes \mathcal{I}_C) \\
&\qquad \circ (\mathcal{U}^{\texttt{CNOT}}_{MC} \otimes \mathcal{I}_D)(\rho_M \otimes \omega_C \otimes \xi_D) ,
\end{aligned} \tag{4.34}$$

where ω and ξ are the initial quantum states of the control qubits C and D, respectively, and $\mathcal{U}^{\texttt{CNOT}}$ and $\mathcal{U}^{\texttt{CPHASE}}$ are the `CNOT` and `CPHASE` gates, respectively, defined by

$$\begin{aligned}
U^{\texttt{CNOT}}_{MC} &:= I_M \otimes |0\rangle\langle 0|_C + X_M \otimes |1\rangle\langle 1|_C \\
U^{\texttt{CPHASE}}_{MD} &:= I_M \otimes |0\rangle\langle 0|_D + Z_M \otimes |1\rangle\langle 1|_D \ ,
\end{aligned} \tag{4.35}$$

X and Z being Pauli gates.

Explicit calculation reveals that for both control qubits initialised in the $|+\rangle$ state, one obtains

$$\begin{aligned}
&\mathrm{Tr}_M \left[\mathcal{G}^{|+\rangle\langle+|,\,|+\rangle\langle+|}(\mathcal{N}_1, \mathcal{N}_2)(\rho_M)\right] \\
&\quad = \rho_C \otimes |+\rangle\langle+|_D + X\rho X_C \otimes |-\rangle\langle-|_D \ .
\end{aligned} \tag{4.36}$$

Therefore, the initial state ρ can be perfectly recovered independently of the noisy channels $\mathcal{N}_1$ and $\mathcal{N}_2$, by measuring D in the Fourier basis and then applying a `NOT` gate on C if the outcome is $|-\rangle$. The supermap defined by Equation (4.34) is an example of an SDPP that generates a perfect quantum side-channel, as it can perfectly transmit one qubit of quantum information for any choice of noisy channels.

The authors of Ref. [96] conclude with regard to their protocol: '*This example shows that SDPPs [...] can be used to perfectly send one qubit of information, essentially trivialising the problem of enhancing quantum and classical channel capacity if one were to take the set of all SDPPs as a resource*'. We agree, and argue that this is the reason why the set of all SDPPs does not define a sensible resource theory of communication.

A possible direction of future research would be to compare the quantum `SWITCH` with the subset of SDPPs that have definite causal order and do not generate side-channels. This may shed light on the mechanism that leads to enhancements in the quantum `SWITCH`, and on whether or not the characteristics of this mechanism can be reproduced by SDPPs with definite causal order.

More interestingly, it would be important to compare the side-channel non-generating SDPPs with definite causal order with *all* the side-channel non-generating SDPPs with indefinite causal order, rather than just restricting the comparison to the quantum SWITCH. For a given pair of channels, the maximum communication capacity achievable with indefinite causal order is—by definition—always larger than or equal to the maximum communication capacity achievable with definite causal order. The interesting question is whether there is a gap between the two, meaning that there exist communication advantages that can be achieved *only* with indefinite causal order.

4.5 Comparison with other frameworks

Our framework is based on the approach of Coecke, Fritz, and Spekkens [67], where the set of free operations is taken as the starting point from which the notion of resource is defined. An alternative approach is to start from a set of *'zero resources'* and to define the free operations as those that preserve this set. For resource theories of quantum channels, this approach was adopted in Refs. [131, 132, 186], where free channels were specified first, and free operations were defined as those supermaps that transform free channels into free channels.

In standard quantum Shannon theory, a natural choice for the set of free channels is the set of constant channels: no communication protocol in standard quantum Shannon theory can achieve communication using only constant channels. The set of supermaps that transform constant channels into constant channels was characterised in Ref. [186], where the authors showed that a supermap preserves the set of constant channels if and only if it is of the form $\mathcal{S}(\mathcal{N}) = \sum_i c_i \mathcal{D}_i \circ \mathcal{N} \circ \mathcal{E}_i$, where $\mathcal{E}_i$ and $\mathcal{D}_i$ are suitable channels, and (c_i) are real (possibly negative) coefficients, such that the map $\sum_i c_i \mathcal{E}_i \otimes \mathcal{D}_i$ is a quantum channel. Physically, these supermaps correspond to the transformations that can be achieved with the assistance of free no-signalling channels between the sender's and receiver's locations.

Going from standard quantum Shannon theory to its extensions, it is not clear whether constant channels should still be regarded as free. Clearly, a *placed* constant channel is useless for communication, because it does not transfer any information from the sender's laboratory to the receiver's laboratory. Hence, placed constant channels should still be considered as free. On the other hand, an *unplaced* constant channel may still be useful, depending on how it interacts with the placements allowed by the theory. This is indeed what happens when the allowed placements include the quantum SWITCH [80].

One might insist that operations that transform constant channels (placed or unplaced) into non-constant channels should not be allowed in a resource theory of communication. This requirement would amount to the following:

Condition 11'. (No Activation of Constant Channels.) *In a resource theory of communication, no free operation $\mathcal{S} \in \mathsf{M}_{\text{free}}$ should be able to transform a constant channel into a non-constant channel.*

Note that Condition 11' (No Activation of Constant Channels) is stronger than Condition 11 (No Side-Channel Generation). If a supermap violated Condition 11, by allowing the sender and receiver to communicate independently of the input channels, then in particular it would allow the sender and receiver to communicate with constant channels, thus violating Condition 11'. In fact, Condition 11' is *strictly stronger* than Condition 11. The quantum SWITCH placement transforms two completely depolarising channels into a non-constant channel [80], thereby violating Condition 11'. On the other hand, the quantum SWITCH placement does not permit the sender and receiver to communicate independently of the input channels: for example, if the input channels are the completely depolarising channel and the identity, the quantum SWITCH placement outputs the channel

$$\mathcal{S}^{A,B,\omega}_{\texttt{SWITCH}}(\mathcal{N}_{\text{dep}}, \mathcal{I}) = \mathcal{N}_{\text{dep}} \otimes \omega \,, \tag{4.37}$$

which is constant and does not permit any communication. Hence, the quantum SWITCH placement satisfies Condition 11, while it violates Condition 11'.

One motivation for assuming Condition 11' would be the idea that the communication provider could 'break' some of the available devices, by turning them into constant channels, *before placing them between the sender and receiver*. This pre-placement operation would be described by a *constant supermap*, of the form

$$\mathcal{S}^{\mathcal{N}_0} : \mathcal{N} \mapsto \mathcal{N}_0 \qquad \forall \mathcal{N} \in \mathsf{Chan}(X \to Y) \,, \tag{4.38}$$

where $\mathcal{N}_0$ is a constant channel. If such constant supermaps were allowed, then Conditions 11 and 11' would become equivalent: a placement supermap that transforms some constant channel into a non-constant channel could be preceded by a constant supermap, thus enabling communication independently of the input channels. However, it is not obvious why constant supermaps should be regarded as free. Ultimately, assuming constant supermaps to be free is equivalent to assuming *by fiat* that constant channels are zero-resource channels, and therefore can be generated for free.

In summary, it is important to distinguish between two requirements: (a) constant channels should not be transformed into non-constant channels, and (b) it should not be possible to communicate independently of the input devices. While requirement (b) may still be too weak to guarantee that a resource theory of communication is interesting, it appears that there are interesting resource theories of communication that violate the requirement (a) and still lead to non-trivial Shannon-theoretic structures.

4.6 Summary

We established a general framework of resource theories of communication. In our framework, the input resources are communication devices, which can be placed between the communicating parties, and combined with local operations performed by the communicating parties. A resource theory is specified by a choice of placement operations, describing how the communication devices are arranged, and by a choice of party operations, describing the action of the communicating parties.

We formulated a minimal requirement that every resource theory of communication should satisfy: no combination of the allowed operations should be able to bypass the communication devices initially available to the communicating parties. We showed that quantum Shannon theory with superpositions of causal order of communication channels [80] and quantum Shannon theory with superpositions of trajectories of information carriers satisfy this requirement, while quantum Shannon theory with superpositions of encoding and decoding operations [96] does not.

Overall, the resource-theoretic framework proposed in this chapter allows for rigorous comparisons between different resource theories of communication, and can be used for the exploration of new models of quantum communication, with both foundational and practical implications.

5
Communication through independent channels in a superposition of trajectories

In the previous two chapters, we defined the superposition of two quantum channels, and constructed a resource theory to use the placement of two channels in superposition for communication. Here we finally present concrete examples of the communication advantages over standard quantum Shannon theory from using independent channels in a superposition. We present examples of classical communication through a superposition of two pure erasure channels, quantum communication through a superposition of two pure dephasing channels, as well as perfect classical and quantum communication through a superposition of an asymptotically large number of such channels. First, we discuss the definitions of communication capacities in this setting.

5.1 Communication capacities assisted by superpositions of trajectories

In §4.2.3, we defined the generalised classical or quantum capacity of an unplaced quantum channel as the maximum number of classical or quantum identity channels that can be generated by performing any combination of the chosen set of free operations, per channel use in the asymptotic limit of $k \to \infty$ channel uses. In practice, however, we typically deal with placed channels, which have already been constructed using the free operations. Therefore, let us make a few comments on the relationship between these two types of capacities.

In the resource theory that includes the superposition placement of two channels with a fixed path state ω, the generalised classical (quantum) capacity of a vacuum-extended channel $\widetilde{\mathcal{N}}$ corresponds to the scenario where the k uses of $\widetilde{\mathcal{N}}$ are first combined in pairs and placed in a superposition, and the resulting $k/2$ placed channels are used in parallel. The resulting capacity is by definition equal to half of the standard classical (quantum) capacity of the corresponding placed channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}, \widetilde{\mathcal{N}})$. (In this and following chapters, we shall suppress the superscripts of the placed input and output systems A, B for brevity.)

More broadly, the generalised classical (quantum) capacity of a vacuum-extended channel $\widetilde{\mathcal{N}}$ in the resource theory that contains the superposition placement of N channels with a fixed path state ω is by definition equal to $1/N$ times the standard classical (quantum) capacity of the corresponding placed channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}^{\otimes N})$.

When the sender has access to multiple non-identical transmission lines, which can be combined together in a superposition, the relationship between the capacity of the resulting placed channel and the generalised capacities of the individual constituent unplaced channels becomes highly non-trivial. To enable these cases to be treated on an equal footing, in the remainder of this work, we shall mostly discuss only the capacities of placed channels.

A circuit diagram corresponding to the communication scenario with two unplaced channels $\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2$ placed in a superposition is shown in Figure 5.1. The placed channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_1, \widetilde{\mathcal{N}}_2)$ is used k times, with a global encoding operation $\mathcal{E}_k$ and decoding operation $\mathcal{D}_k$. Its capacity is determined by case when $k \to \infty$.

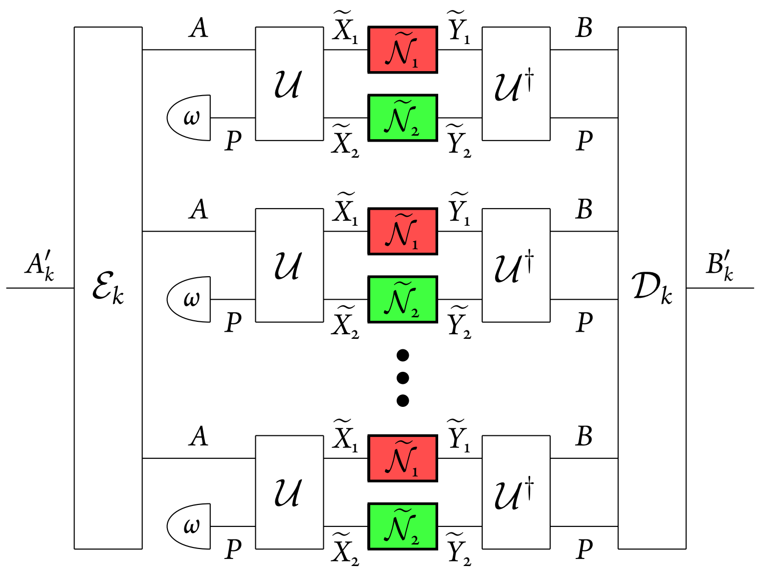


Figure 5.1: *Communication of a message encoded in k particles, with each particle travelling in a superposition of two paths.* The sender encodes the state of a quantum system A'_k onto the internal state of k particles, using a global encoding operation $\mathcal{E}_k$. Each particle is sent through one of the channels $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$, with the choice of the channel controlled by the state ω of the particle's external degree of freedom. Finally, the receiver applies a global decoding operation $\mathcal{D}_k$ on the output particles, returning a quantum state in system B'_k.

5.2 Classical communication through a superposition of pure erasure channels

Consider a direct communication scenario between a sender and a receiver that wish to communicate classical information. Suppose that they have access to two communication channels, each of which are described by the pure erasure channel

$$\mathcal{E}(\rho) = |\psi_0\rangle\langle\psi_0|\,, \tag{5.1}$$

where both $\rho, |\psi_0\rangle\langle\psi_0| \in \mathsf{St}(A)$, which has zero classical capacity. Clearly, standard quantum Shannon theory does not permit any information to be sent to the receiver using these two channels in any combination. Note that a convex combination (i.e. an incoherent mixture) of these two channels is still a zero-capacity channel. In the following, we show that, in contrast, a coherent superposition of the two channels can lead to an effective channel with non-zero classical capacity.

Consider now, that the communication devices used in the protocol can take the vacuum as input, and are described by a vacuum extension $\widetilde{\mathcal{E}}$ with Kraus operators $\{\widetilde{E}_i = |\psi_0\rangle\langle i| \oplus \alpha_i\, |\text{vac}\rangle\langle\text{vac}|\}$, and vacuum amplitudes $\{\alpha_i\}$, for $i \in \{1, \ldots, d\}$. In this case, the sender can transmit the message in a superposition of travelling through each of the two channels by initialising the path in the $|+\rangle$ state. The output state of the overall superposition channel can be computed using Eq. (4.25), which gives

$$\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}(\widetilde{\mathcal{E}}, \widetilde{\mathcal{E}})(\rho) = |\psi_0\rangle\langle\psi_0| \otimes \left(p|+\rangle\langle+| + (1-p)\frac{I}{2}\right) \tag{5.2}$$
$$p = \langle\alpha|\rho|\alpha\rangle, \quad |\alpha\rangle = \sum_i \alpha_i |i\rangle\,.$$

Since the output state depends on the input state ρ, the receiver will be able to decode some of the information in the original message. More specifically, the overall channel is a measure and re-prepare channel, equivalent to the measurement with orthogonal projectors $\{|\alpha\rangle\langle\alpha|, I - |\alpha\rangle\langle\alpha|\}$ followed by a re-preparation of the states $|+\rangle\langle+|$ or $I/2$, depending on the outcome. This is in turn equivalent to a classical binary asymmetric channel, with 0 mapped deterministically to 0, and 1 mapped to a uniform mixture of 0 and 1. The classical capacity of this channel is $\log_2(5/4) \approx 0.32$ bits [163] and can be achieved using polar codes [105]. In a quantum setting, the sender can encode 0 in the state $|\alpha\rangle$ and 1 in an orthogonal state $|\alpha_\perp\rangle$, and then use the optimal classical code.

The essential feature that enabled communication through vacuum-extended pure erasure channels is the fact that the channels preserved coherence between the message and the vacuum. If, on the other hand, we had chosen an incoherent vacuum extension, e.g. the vacuum extension with Kraus operators $\{\widetilde{E}'_i := |\psi\rangle\langle i| \oplus 0_{\mathrm{Vac}}\}$, $i \in \{1, \ldots, d\}$ and $\widetilde{E}'_{d+1} := 0_A \oplus |\mathrm{vac}\rangle\langle\mathrm{vac}|$, the overall channel would be equivalent to a measurement on the path followed by an erasure channel on the message. The output state (calculated from Eq. (4.25)) would have been

$$\mathcal{S}_{\mathrm{sup}}^{|+\rangle\langle+|}(\widetilde{\mathcal{E}}', \widetilde{\mathcal{E}}')(\rho) = |\psi_0\rangle\langle\psi_0| \otimes \frac{I}{2}\,, \tag{5.3}$$

which is independent of the input state ρ, and therefore has zero classical capacity.

Ref. [2] presents similar communication advantages for a superposition of two independent completely depolarising channels, and was completed independently from this work.

5.3 Quantum communication through a superposition of entanglement-breaking channels

Recall the output of a superposition of two identical channels, given in Eq. (3.19). This equation shows that it can be possible to obtain probabilistic noiseless transmission of a quantum state, if the destructive interference term $\mathcal{A} - F \cdot F^\dagger$ is proportional to a unitary gate.

For example, consider the completely dephasing channel

$$\mathcal{D}(\rho) = |0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1|\,, \tag{5.4}$$

which has zero quantum capacity. For a vacuum extension described by Kraus operators $\{|i\rangle\langle i| \oplus \frac{1}{\sqrt{2}}|\mathrm{vac}\rangle\langle\mathrm{vac}|\}_{i=0}^{1}$, the vacuum interference operator is $F = I/\sqrt{2}$. For a superposition of two such completely dephasing channels, this means that the destructive interference term is proportional to the unitary gate $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$, which can be undone by the receiver. The probability of obtaining this term is $1/4$, which means that the superposition of channels enables the perfect transmission of single qubit with 25% probability. In the remaining cases, we obtain constructive interference, with the conditional evolution of the message described by the channel

$$\frac{2}{3}(\mathcal{A}(\rho) + F\rho F^\dagger) = \frac{2}{3}(|0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1| + \rho/2) = \frac{2}{3}\rho + \frac{1}{3}Z\rho Z\,, \tag{5.5}$$

whose quantum capacity is $1 - h(1/3) \approx 0.08$ qubits, where $h(x) = -x\log_2 x - (1-x)\log(1-x)$ is the binary entropy [196]. Since the same encoding works in both cases, the overall quantum capacity of the superposition channel is at least 0.08 qubits.

5.4 Perfect communication through asymptotically many paths

Having seen that the superposition of paths can enable the transfer of both classical and quantum information at a non-zero rate through zero-capacity channels, one might wonder whether the superposition of independent channels can enable perfect noiseless communication through noisy channels. To answer this question, we look again at Equation (3.19). In order for the superposition channel to be completely noiseless (i.e. a unitary), both maps $\mathcal{A} \pm F\rho F^\dagger$ must be proportional to unitary channels. However, the map $\mathcal{A} + F \cdot F^\dagger$ is proportional to a unitary channel if and only if the original channel $\mathcal{A}$ is a unitary channel itself. From Eq. (3.17), we see that the same result holds for the superposition of two different channels $\mathcal{A}$ and $\mathcal{B}$, and, from Eq. (3.20), more generally for a superposition of N independent channels. That is, a superposition of a finite number of independent noisy channels can never lead to a perfectly noiseless channel [58].

Yet, such extreme destructive interference of the noise can in fact occur in the asymptotic limit of a superposition of N independent channels.

Consider the transmission of a single particle through a superposition of N independent and identical transmission lines, each described by the vacuum extension $\widetilde{\mathcal{A}}$ of some channel $\mathcal{A}$. With the path initialised in the maximally coherent state $|e_0\rangle = \sum_{j=0}^{N-1} |j\rangle/\sqrt{N}$, the output of the superposition channel is

$$\mathcal{S}_{\text{sup}}^{|e_0\rangle\langle e_0|}(\widetilde{\mathcal{A}}^{\otimes N})(\rho) = \frac{\mathcal{A}(\rho) + (N-1)\,F\rho F^\dagger}{N} \otimes |e_0\rangle\langle e_0| + \frac{\mathcal{A}(\rho) - F\rho F^\dagger}{N} \otimes \big(I - |e_0\rangle\langle e_0|\big)\,, \tag{5.6}$$

with $F := \sum_i \overline{a}_i\, A_i$. As before, the output state of the path is diagonal in the Fourier basis, heralding the possibility of constructive or destructive interference. In the large N limit, the superposition channel tends to a mixture of the two quantum operations $F \cdot F^\dagger$ and $\mathcal{A} - F \cdot F^\dagger$. This limiting behaviour leads to extreme possibilities for the correction of noise:

5.4.1 Asymptotically large number of pure erasure channels

Consider a communication scenario where the communicating parties have access to $N \to \infty$ independent copies of the pure erasure channel $\mathcal{E}(\rho) = |\psi_0\rangle\langle 0|\rho|0\rangle\langle\psi_0| + |\psi_0\rangle\langle 1|\rho|1\rangle\langle\psi_0|$, with vacuum extension described by Kraus operators $\widetilde{E}_0 = |\psi_0\rangle\langle 0| \oplus \alpha_0 |\text{vac}\rangle\langle\text{vac}|$ and $\widetilde{E}_1 = |\psi_0\rangle\langle 1| \oplus \alpha_1 |\text{vac}\rangle\langle\text{vac}|$. In this case, the vacuum interference operator $F = |\psi_0\rangle\langle\alpha|$ and Equation (5.6) gives

$$\mathcal{S}_{\text{sup}}^{|e_0\rangle\langle e_0|}(\widetilde{\mathcal{E}}^{\otimes N})(\rho) \to \langle\alpha|\rho|\alpha\rangle\, |\psi_0\rangle\langle\psi_0| \otimes |e_0\rangle\langle e_0| + \langle\alpha_\perp|\rho|\alpha_\perp\rangle\, |\psi_0\rangle\langle\psi_0| \otimes \omega_\perp\,, \tag{5.7}$$

with $\langle\alpha_\perp|\alpha\rangle = 0$ and $\omega_\perp := (I - |e_0\rangle\langle e_0|)/(N-1)$. This channel is equivalent to a measurement in the basis $\{|\alpha\rangle, |\alpha_\perp\rangle\}$, followed by a preparation of one of the orthogonal states $|e_0\rangle\langle e_0|$ and $\omega_\perp$, conditional on the outcome. Since these two states are orthogonal, the superposition channel acts as a perfect noiseless channel for the communication of classical bits.

5.4.2 Asymptotically large number of completely depolarising channels

Suppose, that the communicating parties have access to $N \to \infty$ independent copies of the completely depolarising channel $\mathcal{D}(\rho) = (\rho + X\rho X + Y\rho Y + Z\rho Z)/4$ with vacuum extension described by Kraus operators $\widetilde{D}_0 = (I \oplus 1)/2$, $\widetilde{D}_1 = (X \oplus i)/2$, $\widetilde{D}_2 = (Y \oplus i)/2$, and $\widetilde{D}_3 = (Z \oplus i)/2$. Then, the vacuum interference term is proportional to a unitary channel, since $F = (\cos\theta\, I - i \sin\theta S)/2$, with $\cos\theta = 1/2$ and $S = (X + Y + Z)/\sqrt{3}$. As a result, when the measurement on the path heralds the quantum operation $F \cdot F^\dagger$, which occurs with 25% probability, a qubit is noiselessly transmitted from the sender to the receiver. That is, noiseless quantum communication with 25% probability is possible in the $N \to \infty$ limit.

5.4.3 Asymptotically large number of completely dephasing channels

Finally, suppose that the communicating parties have access to $N \to \infty$ independent copies of the completely dephasing channel $\mathcal{A}(\rho) = |0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1|$, with the vacuum extension described by Kraus operators $|0\rangle\langle 0| \oplus |\text{vac}\rangle\langle\text{vac}|/\sqrt{2}$ and $|1\rangle\langle 1| \oplus |\text{vac}\rangle\langle\text{vac}|/\sqrt{2}$. This gives $F = I/\sqrt{2}$ and $\mathcal{A}(\rho) - F\rho F^\dagger = Z\rho Z/2$, such that both quantum operations $F \cdot F^\dagger$ and $\mathcal{A} - F \cdot F^\dagger$ are proportional to unitary channels. By measuring the path, the receiver can find out which unitary channel acted and correct it, enabling perfect noiseless quantum communication.

5.5 Optimal control state for maximizing the communication rate

In the previous subsections, we presented all of the communication advantages from a superposition of channels with the path initialised in the maximally coherent Fourier state, which for qubits is the state $|+\rangle$. This leads to the question whether such a state is always optimal. Here, we show that for a superposition of two identical channels (cf. Eq. (4.25)) the Holevo capacity (a useful lower bound to the classical capacity – cf. §2.2) is indeed maximised for the path initialised as $\omega = |+\rangle\langle+|$.

We prove this in the following lemma:

Lemma 14. *Let $\mathcal{C}_\omega$ be an arbitrary channel of the form*

$$\mathcal{C}_\omega(\rho) = \mathcal{L}_+(\rho) \otimes \omega + \mathcal{L}_-(\rho) \otimes Z\omega Z\,, \tag{5.8}$$

where $\mathcal{L}_\pm$ are arbitrary linear maps. Then, for every quantum state ω, the Holevo capacity satisfies the bound $\chi\left(\mathcal{C}_\omega\right) \leq \chi\left(\mathcal{C}_{|+\rangle\langle+|}\right)$.

Proof. The Holevo capacity is known to be monotonically decreasing under the addition of quantum channels, namely $\chi(\mathcal{E}) \geq \chi(\mathcal{F} \circ \mathcal{E})$ for every pair of channels $\mathcal{E}$ and $\mathcal{F}$. For every channel $\mathcal{C}_\omega$ of the form (5.8), we have the relation

$$\mathcal{C}_\omega = (\mathcal{I}_M \otimes \mathcal{P}_\omega) \circ \mathcal{C}_{|+\rangle\langle+|}\,, \tag{5.9}$$

where $\mathcal{P}_\omega$ is the quantum channel defined by

$$\mathcal{P}_\omega(\gamma) := \langle+|\gamma|+\rangle\, \omega + \langle-|\gamma|-\rangle\, Z\omega Z \tag{5.10}$$

for an arbitrary state γ. Hence, we have $\chi(\mathcal{C}_\omega) = \chi\left[(\mathcal{I}_M \otimes \mathcal{P}_\omega) \circ \mathcal{C}_{|+\rangle\langle+|}\right] \leq \chi(\mathcal{C}_{|+\rangle\langle+|})$. □

Since Eq. (4.25) is of the form (5.8), the Holevo information of a superposition of two identical and independent channels is, in particular, maximised when the path is initialised in the $|+\rangle$ state.

In the next section, we consider the superposition of correlated quantum channels. We compare the performance of a superposition of correlated channels with that of a superposition of independent channels, where we shall make use of Lemma 14.

6

Communication through correlated channels at a superposition of times

In this chapter, we explore the communication enhancements possible when a time-correlated quantum device is used at a superposition of different times. First, we discuss enhancements in classical communication, and show that a transmission line which is completely depolarising on the message at any definite moment in time can be transformed into a perfect communication channel for classical information, for specific choices of the physical parameters.

Then, we extend this result to networks of multiple time-correlated transmission lines. The action of the quantum `SWITCH` can be reproduced in this setting. We show that (1) time-correlations are essential in order to reproduce the communication advantages of the quantum `SWITCH`, and (2) more sophisticated patterns of time correlations in the same setting can surpass the communication capacity associated with the quantum `SWITCH`.

Finally, we discussed enhancements in quantum communication, and find that a transmission line which is completely dephasing on the message at any definite moment in time can, for particular parameters, be transformed into a perfect communication channel for qubits.

6.1 Sending a single particle through a time-correlated channel

In this section, we consider a single particle sent at a superposition of two moments of time through a time-correlated channel. For the transmission of the particle, we consider channels that conserve the number of particles, i.e. that map states of a given sector into states of the same sector. This is the case, for example, for linear optical elements, which preserve the photon number.

In the following, we specialise to the case of correlated channels of the random unitary form

$$\mathcal{R} = \sum_{m,n} p(m,n)\, \mathcal{V}_m \otimes \mathcal{V}_n \in \mathsf{Chan}(S^{(1)}, S^{(2)})\,, \tag{6.1}$$

where $\mathcal{V}_m(\cdot) := V_m(\cdot)V_m^\dagger$ is a unitary channel, $\{V_m\}$ is a set of unitary gates, and $p(m,n)$ is a joint probability distribution. (Recall, that $\mathsf{Chan}(S^{(1)}, S^{(2)})$ is shorthand for a 2-step channel $\mathsf{Chan}(S^{(1)} \to S^{(1)}, S^{(2)} \to S^{(2)})$ – see notation introduced in §2.7.1.) The vacuum extension of each unitary V_m is taken to be another unitary

$$\widetilde{V}_m := V_m \otimes e^{i\varphi_m} \left|\text{vac}\right\rangle\left\langle\text{vac}\right|\,, \tag{6.2}$$

where the vacuum amplitude is given by a complex phase, representing the coherent action of each possible noisy process on the one-particle and vacuum sectors. This leads to the vacuum extension

$$\widetilde{\mathcal{R}} = \sum_{m,n} p(m,n)\, \widetilde{\mathcal{V}}_m \otimes \widetilde{\mathcal{V}}_n \in \mathsf{Chan}(\widetilde{S}^{(1)}, \widetilde{S}^{(2)})\,, \tag{6.3}$$

with $\widetilde{\mathcal{V}}_m(\cdot) := \widetilde{V}_m(\cdot)\widetilde{V}_m^\dagger$.

The use of the channel $\mathcal{R}$, specified by the vacuum extension $\widetilde{\mathcal{R}}$, at a superposition of times (cf. the superposition placement (4.25)) is given by:

$$\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{R}}) = \sum_{m,n=0}^{r-1} p(m,n)\, \mathcal{U}^\dagger \circ \left(\widetilde{\mathcal{V}}_m \otimes \widetilde{\mathcal{V}}_n\right) \circ \mathcal{U}\,. \tag{6.4}$$

Explicitly, this is equivalent to the effective channel

$$C_\omega(\rho) := \mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{R}})(\rho) = \sum_{m,n} p(m,n)\, W_{mn}\,(\rho\otimes\omega)\, W^\dagger_{mn}\,, \tag{6.5}$$

where ρ (respectively, ω) is an arbitrary state of the message (respectively, control), and

$$W_{mn} := V_m\, e^{i\varphi_n} \otimes |0\rangle\langle 0| + e^{i\varphi_m}\, V_n \otimes |1\rangle\langle 1|\,, \tag{6.6}$$

where $e^{i\varphi_m}$ is the vacuum amplitude in Eq. (6.2).

It is useful to consider the case where the probability distribution $p(m,n)$ is symmetric, that is, $p(m,n) = p(n,m)$ for every m and n. In this case, the superposition channel has the simple expression

$$C_\omega(\rho) = \mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{R}})(\rho) = \frac{\mathcal{R}_1(\rho)+\mathcal{G}(\rho)}{2}\otimes\omega + \frac{\mathcal{R}_1(\rho)-\mathcal{G}(\rho)}{2}\otimes Z\omega Z\,, \tag{6.7}$$

where $\mathcal{R}_1$ is the reduced channel defined by

$$\mathcal{R}_1(\rho) := \sum_m p_1(m)\, V_m\rho V_m^\dagger \qquad p_1(m) := \sum_n p(m,n)\,, \tag{6.8}$$

and $\mathcal{G}$ is the linear map defined by

$$\mathcal{G}(\rho) := \sum_{m,n} p(m,n)\, e^{i(\varphi_n-\varphi_m)}\, V_m\rho V_n^\dagger\,. \tag{6.9}$$

Here, the channel $\mathcal{R}_1$ is the quantum channel representing the evolution of the message when it is sent at a definite time (either t_1 or t_2). It depends only on the marginal probability distribution $p_1(m) := \sum_n p(m,n)$ and is independent of the correlations. In contrast, the map $\mathcal{G}$ generally depends on the correlations between the evolution of the particle at two mutually exclusive moments of time. We call $\mathcal{G}$ the *interference term*.

6.2 Classical communication through correlated white noise

6.2.1 Correlated white noise

Consider the case where the evolution at any definite time step is completely depolarising on the message-carrying sector M, that is,

$$\mathcal{C}_{|j\rangle\langle j|}(\rho) = \mathrm{Tr}(\rho)\frac{I}{d} \otimes |j\rangle\langle j| \qquad \forall \rho \in \mathsf{St}(M)\,, \forall j \in \{0,1\}\,, \tag{6.10}$$

where $\mathcal{C}_{|j\rangle\langle j|}$ is the quantum channel obtained by plugging $\omega = |j\rangle\langle j|$ into Eq. (6.7). Eq. (6.10) implies that whenever the particle is sent at a definite moment of time, the message is replaced by white noise. Accordingly, the channel $\mathcal{R}_1$ in Eq. (6.8) is a completely depolarising channel (cf. 2.2.2).

When the probability distribution $p(m,n)$ is symmetric, Eq. (6.7) becomes

$$\mathcal{C}_\omega(\rho) = \frac{I/d + \mathcal{G}(\rho)}{2} \otimes \omega + \frac{I/d - \mathcal{G}(\rho)}{2} \otimes Z\omega Z\,. \tag{6.11}$$

In the realisation of the random unitary channel, we shall take the unitaries $\{V_m\}$ to be an orthogonal basis for the space of $d \times d$ matrices. Accordingly, the set $\{V_m\}$ will contain d^2 unitaries, labelled by integers from 0 to $d^2 - 1$. For qubits, we shall take $\{V_m\}$ to be the four Pauli matrices $\{I, X, Y, Z\}$, labelled as $V_0 = I$, $V_1 = X$, $V_2 = Y$, and $V_3 = Z$.

In terms of the probability distribution $p(m,n)$, the condition (6.10) amounts to requiring that the marginal probability distributions $p_1(m)$ and $p_2(n)$ be uniform, that is

$$p_1(m) = p_2(n) = \frac{1}{d^2} \qquad \forall m,n \in \{0,\ldots,d^2-1\}\,. \tag{6.12}$$

The probability distributions $p(m,n)$ satisfying Eq. (6.12) form a convex polytope whose extreme points are probability distributions of the form $p(m,n) = \delta_{m,\sigma(n)}/d^2$, where σ is a permutation of the set $\{0,\ldots,d^2-1\}$ [35].

For the identity permutation, satisfying $\sigma(m) = m$ for all values of m, the probability distribution $p(m,n)$ is symmetric, and the interference term (6.9) is the completely depolarising channel $\mathcal{G}(\rho) = \mathrm{Tr}(\rho)I/d\ \forall\rho$. Hence, the channel $\mathcal{C}_\omega$ in Eq. (6.11) is completely depolarising, and no information can be transmitted through it, no matter what state ω is used. In the following, we shall show that, instead, other types of permutations enable the perfect transmission of classical information.

6.2.2 Perfect communication through correlated completely depolarising channels

Here we focus on the case where the message is a qubit ($d = 2$). Let σ be a permutation that swaps two pairs of indices, for example mapping $(0, 1, 2, 3)$ into $(1, 0, 3, 2)$. In this case, the probability distribution $p(m, n) = \delta_{m,\sigma(n)}/4$ is symmetric, and the interference term is

$$\mathcal{G}(\rho) = \frac{\rho X e^{i(\varphi_1 - \varphi_0)} + Y\rho Z e^{i(\varphi_3 - \varphi_2)} + \text{h.c.}}{4}, \tag{6.13}$$

where h.c. denotes the Hermitian conjugate of the preceding matrices.

Note that $\mathcal{G}(\rho)$ depends only on the differences $\varphi_1 - \varphi_0$ and $\varphi_3 - \varphi_2$. We now show that, by suitably choosing the differences $\varphi_1 - \varphi_0$ and $\varphi_3 - \varphi_2$, and the state ω, it is possible to achieve a perfect transmission of classical information. When $\varphi_1 - \varphi_0 = 0$ and $\varphi_3 - \varphi_2 = \pi/2$, the interference term becomes

$$\mathcal{G}(\rho) = \frac{\{\rho, X\} - \{Z\rho Z, X\}}{4}, \tag{6.14}$$

where $\{A, B\} = AB + BA$ denotes the anticommutator of two generic operators A and B. In particular, choosing $\rho = |\pm\rangle\langle\pm|$, with $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$, we obtain

$$\mathcal{G}(|\pm\rangle\langle\pm|) = \pm\frac{I}{2}. \tag{6.15}$$

Combining this relation with the depolarising condition $\mathcal{C}(|\pm\rangle\langle\pm|) = I/2$, and inserting these two relations into into Eq. (6.7), we obtain

$$\mathcal{C}_\omega(|\pm\rangle\langle\pm|) = \frac{I}{2} \otimes \omega_\pm, \tag{6.16}$$

with $\omega_+ := \omega$ and $\omega_- := Z\omega Z$. In other words, the net effect of the superposition of correlated depolarising channels is to transfer information from the message to the output state of the control.

Putting the control in the state $\omega = |+\rangle\langle+|$, one obtains the orthogonal output states $\omega_\pm =$

$|\pm\rangle\langle\pm|$. Hence, a sender can encode a bit into the states $|\pm\rangle$, and a receiver will be able to decode the bit in principle without error, by measuring the control system in the basis $\{|+\rangle, |-\rangle\}$.

In summary, there exist time-correlated channels that look completely depolarising when the message is sent at any definite moment of time, and yet allow for a perfect transmission of classical information by sending messages at a coherent superposition of different times.

6.2.3 Maximum capacity in the lack of correlations

We now show that correlations in the probability distribution $p(m,n)$ are essential in order to achieve the perfect communication task discussed in the previous subsection. Specifically, we prove that no perfect communication is possible in the lack of correlations, that is, when the probability distribution factorises as $p(m,n) = p_1(m)\, p_2(n) = 1/d^4$ (cf. Eq. (6.12)). For qubit messages $(d = 2)$, we show that, in the lack of correlations,

1. the classical capacity of the channel $\mathcal{C}_\omega$ is upper bounded by 0.5 bits, meaning that it is impossible to transmit more than 0.5 bits per use of the channel,
2. the maximum classical capacity of the channel $\mathcal{C}_\omega$ over arbitrary states ω of the control system and over arbitrary (not necessarily random-unitary) realisations of the completely depolarising channel is equal to 0.16 bits.

The first result follows from an analytical upper bound on the classical capacity, while the second result follows from numerical optimisation.

We begin by deriving the form of the superposition of two uncorrelated depolarising channels. Note, the superposition of times of the correlated depolarising channel (6.11) is mathematically equivalent to the superposition of paths of two independent completely depolarising channels. Let $\mathcal{D} : \rho \mapsto \mathrm{Tr}(\rho) I/d$ be the completely depolarising channel, with vacuum extension $\widetilde{\mathcal{D}}$. Using the superposition placement (4.25), we obtain, for a fixed state ω of the control system,

$$\mathcal{C}_{\omega,F}(\rho) := \mathcal{S}^{\omega}_{\mathrm{sup}}(\widetilde{\mathcal{D}}, \widetilde{\mathcal{D}})(\rho) = \frac{I/d + F\rho F^\dagger}{2} \otimes \omega + \frac{I/d - F\rho F^\dagger}{2} \otimes Z\omega Z\,, \tag{6.17}$$

where F is the vacuum interference operator associated with $\widetilde{\mathcal{D}}$.

The derivation of the bound consists of three steps:

1. The first step is to prove that, in the lack of correlations and for message dimension $d = 2$, the channel $\mathcal{C}_{\omega,F}$ is entanglement-breaking [107], i.e. it transforms all entangled states into separable states. This is formally stated in Proposition 15 below and proven in Appendix A.2. For entanglement-breaking channels, it is known that the classical capacity coincides with the *Holevo capacity* [179].

2. The second step is to observe that state of the control that maximises the Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ is $\omega = |+\rangle\langle+|$. This follows from Lemma 14 and the result holds for arbitrary message dimension $d \geq 2$. In fact, it holds even in the presence of correlations, as long as the probability distribution $p(m, n)$ is symmetric.

3. Finally, the third step is to show that, in the lack of correlations and for arbitrary message dimension $d \geq 2$, the Holevo capacity of the channel $\mathcal{C}_{|+\rangle\langle+|,F}$ is upper bounded by $1/d$. This follows from the more general result of Proposition 16 below.

Proposition 15. *The channel $\mathcal{C}_{\omega,F}$ defined in Eq. (6.17) is entanglement-breaking for $d = 2$.*

The proof is given in Appendix A.2.

Proposition 16. *The Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ defined in Eq. (6.17) is upper bounded as*

$$\chi(\mathcal{C}_{\omega,F}) \leq \frac{\log(2d)}{d} + \frac{\frac{1}{d} + ||F||_\infty^2}{2} \log \frac{\frac{1}{d} + ||F||_\infty^2}{2} + \frac{\frac{1}{d} - ||F||_\infty^2}{2} \log \frac{\frac{1}{d} - ||F||_\infty^2}{2}, \qquad (6.18)$$

where F is the vacuum interference operator defined in Eq. (3.12).

The proof is provided in Appendix A.3.

Putting the three steps together, we obtain that, in the lack of correlations and for qubit messages, the classical capacity of the channel $\mathcal{C}_{\omega,F}$ is upper bounded by $1/2$ for every possible state ω:

Corollary 17. *The classical capacity of the channel $\mathcal{C}_{\omega,F}$ defined in Eq. (6.17) is upper bounded as $C(\mathcal{C}_{\omega,F}) \leq 1/d$. In particular, for $d = 2$, one has the bound $C(\mathcal{C}_{\omega,F}) \leq 0.5$.*

Proof. Immediate from the fact that the right-hand-side of Eq. (6.18) is monotonically decreasing with $\|F\|_\infty$, and that $\|F\|_\infty$ is upper bounded by $1/\sqrt{d}$ (Lemma 22), together with Lemma 14 and the fact that $\mathcal{C}_{\omega,F}$ is entanglement-breaking. □

Hence, the perfect transmission of 1 bit achieved in §6.2.2 is impossible in the lack of correlations.

Numerical evaluation of the capacity

The evaluation of the Holevo capacity involves an optimisation over all possible input ensembles. For quantum channels with d-dimensional input, the optimisation can be restricted to ensembles with up to d^2 linearly independent pure states [71]. In practice, however, the optimisation is often hard to carry out even in dimension $d = 2$. To make the optimisation feasible, we found that in our case the optimisation can be reduced to an optimisation over ensembles that depend only on three real parameters $q, p_0, p_1 \in [0, 1]$.

Theorem 18. *The classical capacity of the superposition of two independent qubit completely depolarising channels, $\mathcal{C}_{\omega,F}$ [Eq. (6.17)] is given by*

$$C(\mathcal{C}_{\omega,F}) \leq \max_{0\leq q,p_0,p_1\leq 1} H\Big[\mathcal{C}_{\omega,F}\Big(\rho_q\Big)\Big] - qH\left[\mathcal{C}_{\omega,F'}(|\psi_0\rangle\langle\psi_0|)\right] - (1-q)H\left[\mathcal{C}_{\omega,F'}(|\psi_1\rangle\langle\psi_1|)\right] \ ,$$

$$\begin{cases} |\psi_0\rangle &= \sqrt{p_0}\,|0\rangle + \sqrt{1-p_0}\,|1\rangle \ , \\ |\psi_1\rangle &= \sqrt{p_1}\,|1\rangle + \sqrt{1-p_1}\,|1\rangle \ , \\ \rho_q &= [qp_0 + (1-q)p_1]\,|0\rangle\langle 0| \\ & \quad +[q(1-p_0) + (1-q)(1-p_1)]\,|1\rangle\langle 1| \end{cases} \tag{6.19}$$

where $F' = a\,|0\rangle\langle 0| + b\,|1\rangle\langle 1|$, with a, b being the singular values of F.

Proof. Follows directly from Lemma 23, which is proven in Appendix A.4, and Lemma 24, which is given in Appendix A.4 and proven in Ref. [121]. □

Corollary 19. *For every vacuum extension of the completely depolarising channel and for every state of the control qubit, the classical capacity of any channel of the form Eq. (6.17) resulting from the superpo-*

sition of two independent qubit completely depolarising channels is upper bounded as

$$C(\mathcal{C}_{\omega,F}) \le \max_{\substack{a\ge 0, b\ge 0\\ a^2+b^2\le 1/2}} \max_{0\le q,p_0,p_1\le 1} H\Big[\mathcal{C}_{\omega,F}\Big(\rho_q\Big)\Big] - qH\left[\mathcal{C}_{\omega,F}(|\psi_0\rangle\langle\psi_0|)\right] - (1-q)H\left[\mathcal{C}_{\omega,F}(|\psi_1\rangle\langle\psi_1|)\right] ,$$

$$\begin{cases} |\psi_0\rangle &= \sqrt{p_0}\,|0\rangle + \sqrt{1-p_0}\,|1\rangle , \\ |\psi_1\rangle &= \sqrt{p_1}\,|1\rangle + \sqrt{1-p_1}\,|1\rangle , \\ \rho_q &= [qp_0 + (1-q)p_1]\,|0\rangle\langle 0| + [q(1-p_0) + (1-q)(1-p_1)]\,|1\rangle\langle 1| , \\ F &= a\,|0\rangle\langle 0| + b\,|1\rangle\langle 1| \end{cases} \tag{6.20}$$

Proof. Follows from Theorem 18, which shows that only channels with a vacuum interference operator that is diagonal in the computational basis are relevant for maximising the capacity, and the fact that any vacuum interference operator F must satisfy the condition $\operatorname{Tr} F^\dagger F \le 1/d$ [2], which implies the inequality $|a|^2 + |b|^2 \le 1/d$ for an operator of the form $F = a\,|0\rangle\langle 0| + b\,|1\rangle\langle 1|$.
□

Using Theorem 18, we numerically evaluate the largest value of the Holevo capacity, and therefore the classical capacity, for all possible qubit channels (i.e. $d = 2$) of the form (6.11) with $p(m,n) = 1/16$. We set the state of the control to $\omega = |+\rangle\langle +|$, which we know to guarantee the maximum Holevo capacity (cf. Lemma 14).

The resulting value of the Holevo capacity is a function of the phases $\{\varphi_m\}_{m\in\{0,1,2,3\}}$ in Eq. (6.9). One phase, say φ_0, can be set to 0 without loss of generality, as it represents a global phase. In Figure 6.1a, we provide a 3-dimensional plot showing the exact values of the Holevo capacity, and therefore by the arguments above, the classical capacity, for all possible values of the phases φ_1, φ_2, and φ_3. The maximum over all possible choices of phases is 0.16 bits.

More generally, using the results of Corollary 19, we also find that 0.16 bits is the maximum capacity achievable with arbitrary (not necessarily random unitary) channels that reduce to the depolarising channel in the one-particle sector. The value 0.16 was previously found to be a lower bound to the classical capacity [2], and our result shows that the lower bound is actually tight: 0.16 is the best classical capacity one can obtain by sending a single particle through a superposition of paths traversing two identical, independent channels that are completely depolarising in the one-particle sector.

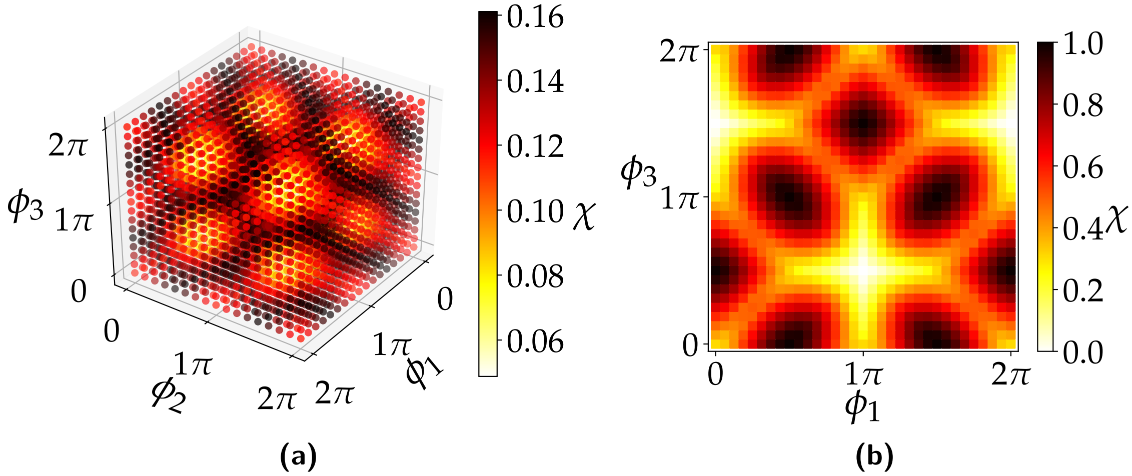


Figure 6.1: *Performance in the transmission of a single particle through a correlated depolarising channel.* (a) Classical capacity in the lack of correlations. Without loss of generality, $\varphi_0 = 0$. The maximum capacity is 0.016 bits. (b) Lower bound to the classical capacity achieved with the correlated probability distribution $p(m,n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. Without loss of generality, we set $\varphi_0 = \varphi_2 = 0$. The maximum lower bound is 1 bit.

6.2.4 Lower bound to the classical capacity in the presence of correlations

In the correlated case, we do not have a proof that the classical capacity coincides with the Holevo capacity. On top of that, the evaluation of the Holevo capacity generally requires an optimisation over all possible ensembles of d^2 linearly independent pure states, which is computationally challenging. Here, we circumvent this problem by computing a lower bound to the Holevo capacity, obtained by restricting the optimisation to the set of all *orthogonal* ensembles, that is, input ensembles consisting of two orthogonal qubit states. In general, this lower bound may not be tight [87, 99, 115], but it is nevertheless interesting as it quantifies the maximum performance of a natural set of encoding strategies. Since the Holevo capacity is always a lower bound to the classical capacity, the above lower bound is also a lower bound to the classical capacity.

Here, we evaluate the lower bound for the correlated channel with $p(m,n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. This particular choice is interesting because as we have seen in §6.2.2, it can reach the maximum capacity of 1 bit. We now inspect how the lower bound depends on the phases.

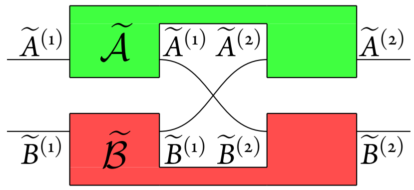


Figure 6.2: *The channel $\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})$, obtained by connecting two vacuum-extended 2-step channels $\widetilde{\mathcal{A}}$ (green) and $\widetilde{\mathcal{B}}$ (red) such that output of the first port of each channel is connected to the input of the second port of the other channel.*

Since the interference term (6.13) depends only on the differences $\varphi_1 - \varphi_0$ and $\varphi_3 - \varphi_2$, we set $\varphi_0 = \varphi_2 = 0$ and scan the possible values of φ_1 and φ_3. For the state of the control system, we again choose $\omega = |+\rangle\langle+|$, as it maximises the Holevo capacity (cf. Lemma 14). The lower bound to the Holevo capacity is shown in Figure 6.1b for all values of φ_1 and φ_3.

6.3 Communication through multiple time-correlated channels

Time-correlated channels can be used to mimic the use of ordinary quantum channels in a superposition of different causal orders [50, 147]. In this section we show that time correlations are a necessary resource for reproducing the benefits of the superposition of orders in quantum communication, and that, in fact, time correlations are an even more powerful resource than the ability to combine channels in a superposition of orders.

6.3.1 A network of time-correlated channels

Suppose that two time-correlated channels are arranged as in Figure 6.2, and that a single particle is sent through a superposition of two alternative paths visiting each of the two channels exactly once. We describe this mathematically as follows: Let $\mathcal{A}$ and $\mathcal{B}$ be two-step channels, with vacuum extensions $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{A}^{(1)}, \widetilde{A}^{(2)})$ and $\widetilde{\mathcal{B}} \in \mathsf{Chan}(\widetilde{B}^{(1)}, \widetilde{B}^{(2)})$. For simplicity, here we take all the systems $\widetilde{A}^{(1)}, \widetilde{A}^{(2)}, \widetilde{B}^{(1)}, \widetilde{B}^{(2)}$ to be isomorphic.

We now connect the 2-step channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ in such a way that the output of the first port of each channel is fed into the input of the second port of the other channel, as in Figure 6.2. This particular composition of two 2-step channels is described by a supermap $\mathcal{Z}$ that maps pairs of channels in $\mathsf{Chan}(\widetilde{A}^{(1)}, \widetilde{A}^{(2)}) \times \mathsf{Chan}(\widetilde{B}^{(1)}, \widetilde{B}^{(2)})$ into bipartite channels in $\mathsf{Chan}(\widetilde{A}^{(1)} \otimes \widetilde{B}^{(1)} \to$

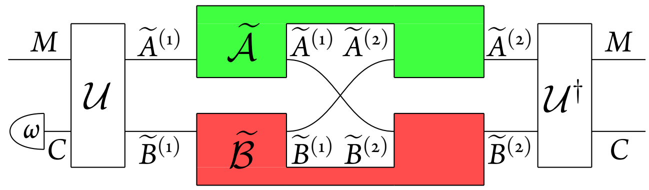


Figure 6.3: *The superposition channel $\mathcal{S}[\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})]$ of two 2-step channels $\mathcal{A}$ and $\mathcal{B}$, specified by the vacuum extensions $\widetilde{\mathcal{A}}$ (green) and $\widetilde{\mathcal{B}}$ (red), where the alternative paths traverse the two correlated channels in the opposite order.*

$\widetilde{B}^{(2)} \otimes \widetilde{A}^{(2)}$).

We can now consider the scenario in which a single particle is sent in a superposition of going through the A-port and the B-port of the channel $\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})$. Following the superposition placement (4.25), the evolution of the particle is described by the superposition channel

$$\mathcal{S}^{\omega}_{\rm sup}\left[\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})\right](\cdot) := \mathcal{U}^\dagger \circ \mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}}) \circ \mathcal{U}(\cdot \otimes \omega)\,, \tag{6.21}$$

with $\mathcal{U}$ defined as in Eq. (3.16). The superposition channel $\mathcal{S}^{\omega}_{\rm sup}\left[\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})\right]$ is illustrated in Figure 6.3.

Let us apply the above construction to the special case where the channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are of the random unitary form

$$\begin{aligned}\widetilde{\mathcal{A}} = \widetilde{\mathcal{R}}_A &:= \sum_{m,n} p_A(m,n)\, \widetilde{\mathcal{V}}^{(A)}_m \otimes \widetilde{\mathcal{V}}^{(A)}_n \\ \widetilde{\mathcal{B}} = \widetilde{\mathcal{R}}_B &:= \sum_{k,l} p_B(k,l)\, \widetilde{\mathcal{V}}^{(B)}_k \otimes \widetilde{\mathcal{V}}^{(B)}_l\,,\end{aligned} \tag{6.22}$$

where $\widetilde{\mathcal{V}}^{(A)}_m$ and $\widetilde{\mathcal{V}}^{(B)}_k$ are the unitary channels corresponding to the unitary operators

$$\begin{aligned}\widetilde{V}^{(A)}_m &:= V^{(A)}_m \oplus e^{i\varphi^{(A)}_m}\,|\mathrm{vac}\rangle\langle\mathrm{vac}| \\ \widetilde{V}^{(B)}_k &:= V^{(B)}_k \oplus e^{i\varphi^{(B)}_k}\,|\mathrm{vac}\rangle\langle\mathrm{vac}|\,,\end{aligned} \tag{6.23}$$

respectively. With this choice, we have

$$\mathcal{Z}(\widetilde{\mathcal{R}}_A,\widetilde{\mathcal{R}}_B)=\sum_{m,n,k,l}p_A(m,n)p_B(k,l)\,(\widetilde{\mathcal{V}}_l^{(B)}\circ\widetilde{\mathcal{V}}_m^{(A)})\otimes(\widetilde{\mathcal{V}}_n^{(A)}\circ\widetilde{\mathcal{V}}_k^{(B)})\,. \tag{6.24}$$

When the control system is initialised in the state ω, the overall evolution of the message and the control is then described by the effective channel $\mathcal{E}_\omega$ defined as

$$\mathcal{E}_\omega(\rho):=\mathcal{S}^\omega_{\text{sup}}\left[\mathcal{Z}\left(\widetilde{\mathcal{R}}_A,\widetilde{\mathcal{R}}_B\right)\right]\,(\rho)=\sum_{m,n,k,l}p_A(m,n)\,p_B(k,l)\,W_{mnkl}\,(\rho\otimes\omega)\,W^\dagger_{mnkl}\,, \tag{6.25}$$

with

$$W_{mnkl}:=V_l^{(B)}V_m^{(A)}\,e^{i(\varphi_k^{(B)}+\varphi_n^{(A)})}\otimes|0\rangle\langle 0|+V_n^{(A)}V_k^{(B)}\,e^{i(\varphi_m^{(A)}+\varphi_l^{(B)})}\otimes|1\rangle\langle 1|\,. \tag{6.26}$$

Expanding out, we obtain

$$\begin{aligned}
\mathcal{S}^\omega_{\text{sup}}\left[\mathcal{Z}\left(\widetilde{\mathcal{R}}_A,\widetilde{\mathcal{R}}_B\right)\right](\rho)&=\sum_{m,n,k,l}p_A(m,n)\,p_B(k,l)\,W_{mnkl}(\rho\otimes\omega)W^\dagger_{mnkl} \qquad (6.27)\\
&=\sum_{m,n,k,l}\Big\{p_A(m,n)\,p_B(k,l)\,V_l^{(B)}V_m^{(A)}\rho V_m^{(A)\dagger}V_l^{(B)\dagger}\otimes\omega_{00}\,|0\rangle\langle 0|\\
&\qquad+p_A(m,n)\,p_B(k,l)\,V_n^{(A)}V_k^{(B)}\rho V_k^{(B)\dagger}V_n^{(A)\,\dagger}\otimes\omega_{11}\,|1\rangle\langle 1|\\
&\qquad+p_A(m,n)\,p_B(k,l)\,V_l^{(B)}V_m^{(A)}\rho V_k^{(B)\,\dagger}V_n^{(A)\dagger}\,e^{i\left[\varphi_n^{(A)}+\varphi_k^{(B)}-\varphi_m^{(A)}-\varphi_l^{(B)}\right]}\otimes\omega_{01}\,|0\rangle\langle 1|+\text{h.c.}\Big\}\\
&=\mathcal{R}_B\mathcal{R}_A(\rho)\otimes\omega_{00}\,|0\rangle\langle 0|+\mathcal{R}_A\mathcal{R}_B(\rho)\otimes\omega_{11}\,|1\rangle\langle 1|+\mathcal{K}(\rho)\otimes\omega_{01}\,|0\rangle\langle 1|+[\mathcal{K}(\rho)]^\dagger\otimes\omega_{10}\,|1\rangle\langle 0|\,,
\end{aligned}$$

where $\omega_{mn}:=\langle m|\,\omega\,|n\rangle$ and $\mathcal{K}$ is the linear map defined by

$$\mathcal{K}(\rho):=\sum_{m,n,k,l}p_A(m,n)\,p_B(k,l)\,V_l^{(B)}V_m^{(A)}\rho V_k^{(B)\,\dagger}V_n^{(A)\dagger}\,e^{i\left[\varphi_n^{(A)}+\varphi_k^{(B)}-\varphi_m^{(A)}-\varphi_l^{(B)}\right]}\,. \tag{6.28}$$

We now restrict our attention to the case where

1. the two channels $\widetilde{\mathcal{R}}_A$ and $\widetilde{\mathcal{R}}_B$ are identical (this implies that one can choose without loss of generality $p_A(m,n)=p_B(m,n):=p(m,n)$ for every m and n, $V_m^{(A)}=V_m^{(B)}:=V_m$, and $\varphi_m^{(A)}=\varphi_m^{(B)}=:\varphi_m$ for every m),

2. the probability distribution $p(m,n)$ is symmetric, namely $p(m,n) = p(n,m)$ for every m, n.

Under these conditions, the operator $\mathcal{K}(\rho)$ is self-adjoint for every density matrix ρ, and the effective channel can be rewritten as

$$\mathcal{S}^{\omega}_{\text{sup}}\left[\mathcal{Z}\left(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B\right)\right](\rho) = \frac{(\mathcal{R}_1)^2(\rho) + \mathcal{K}(\rho)}{2} \otimes \omega + \frac{(\mathcal{R}_1)^2(\rho) - \mathcal{K}(\rho)}{2} \otimes Z\omega Z\,, \tag{6.29}$$

with

$$\mathcal{R}_1(\rho) := \sum_{m,n} p(m,n)\, V_m \rho V_m^\dagger\,. \tag{6.30}$$

Thus, we have found the output of the superposition of two identical 2-step random unitary channels with symmetric probability distributions, connected as in Figure 6.3. Notably, Equation (6.29) has the same form as Equation (6.7) describing a superposition of a single 2-step random unitary channel, with $\mathcal{R}_1$ replaced by $(\mathcal{R}_1)^2$ and $\mathcal{G}(\rho)$ replaced by $\mathcal{K} = \mathcal{G}^2$.

6.3.2 Reproducing the quantum SWITCH

An interesting special case occurs when the probability distributions $p_A(m,n)$ and $p_B(k,l)$ are perfectly correlated, that is

$$\begin{aligned} p_A(m,n) &= p_{1A}(m)\delta_{mn} \quad \forall m, n \\ p_B(k,l) &= p_{1B}(k)\,\delta_{kl} \quad \forall k, l\,, \end{aligned} \tag{6.31}$$

where $p_{1A}(m)$ and $p_{1B}(k)$ are the marginal probability distributions of $p_A(m,n)$ and $p_B(k,l)$, respectively. Under this condition, the network in Figure 6.3 reproduces the action of two random unitary channels in a superposition of two alternative orders [50].

In particular, here we are interested in the case where the channels $\mathcal{A}$ and $\mathcal{B}$ are random unitary, with Kraus operators $A_m := \sqrt{p_{1A}(m)}\, V_m^{(A)}$ and $B_k := \sqrt{p_{1B}}\, V_k^{(B)}$. With this choice, the output of the quantum SWITCH placement (Eq. (4.23), which we shall call $\mathcal{W}_\omega$), coincides with the channel $\mathcal{E}_\omega$ in Eq. (6.25) under the condition that the probability distributions $p_A(m,n)$ and $p_B(k,l)$ are perfectly correlated (cf. Eq. (6.31)).

When the channels $\mathcal{A}$ and $\mathcal{B}$ are completely depolarising, Ref. [80] showed that the channel $\mathcal{W}_\omega$ resulting from the quantum SWITCH can transmit 0.049 bits of classical information, provided that the control is initialised in the state $\omega = |+\rangle\langle+|$. Later, the value 0.049 was proven to be exactly equal to the classical capacity [59]. Since the channels $\mathcal{E}_\omega$ and $\mathcal{W}_\omega$ coincide, we conclude that the time-correlated network in Figure 6.3 can achieve a capacity of 0.049 bits.

In the following, we provide two new results:

1. We show that time correlations are strictly necessary in order to achieve the quantum SWITCH capacity of 0.049 bits. Specifically, we show numerically that the maximum classical capacity in the uncorrelated case is 0.018 bits for random-unitary realisations of the completely depolarising channel, and 0.024 bits for arbitrary realisations. This result shows that, when the quantum SWITCH is reproduced by the network in Figure 6.3, the origin of the communication enhancement is not just the interference of paths, but rather the combined effect of the interference of paths *and* of the time correlations. We discuss this point further in Chapter 7.

2. We show that there exist time correlations that achieve a classical capacity of at least 0.31 bits. This result shows that the access to time correlations is generally a stronger resource than the ability to combine ordinary channels in a superposition of orders.

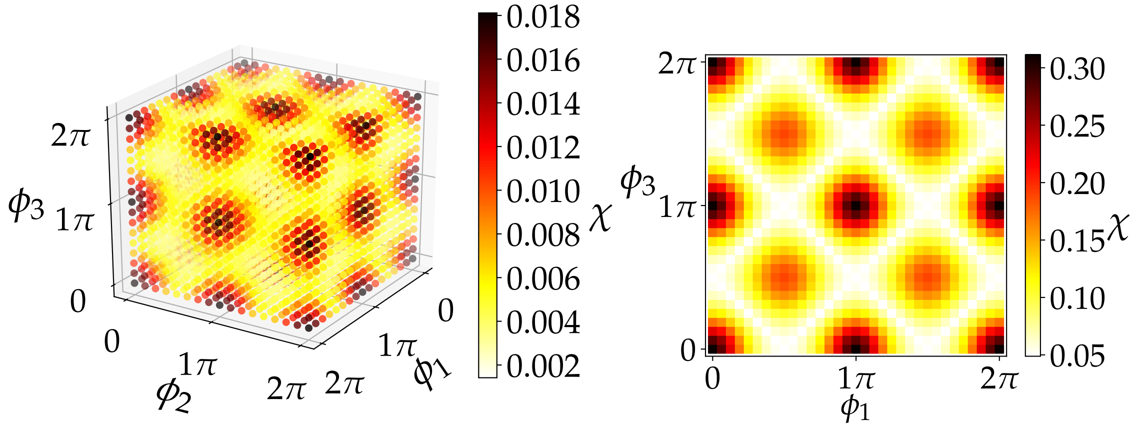
6.3.3 Maximum capacity in the lack of correlations

Here we evaluate the maximum amount of classical information that can be transmitted through the network in Figure 6.3 when the channels are completely depolarising and no correlation is present. That is, in the case of random unitary channels, the probabilities distribution factorises as $p_A(m,n) = p_{A,1}(m)p_{A,2}(n) = p_B(k,l) = p_{B,1}(k)p_{B,2}(l) = 1/4 \cdot 1/4 = 1/16 \ \forall m,n,k,l \in \{0,1,2,3\}$.

Consider the scenario of Figure 6.3, in the special case where the 2-step channels $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are of the product form $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 \otimes \widetilde{\mathcal{A}}_2$ and $\widetilde{\mathcal{B}} = \widetilde{\mathcal{B}}_1 \otimes \widetilde{\mathcal{B}}_2$, respectively. In this case, the combination of the channels in the network of Figure 6.2 gives the bipartite channel

$$\mathcal{Z}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}) = \widetilde{\mathcal{B}}_2\widetilde{\mathcal{A}}_1 \otimes \widetilde{\mathcal{A}}_2\widetilde{\mathcal{B}}_1 \,. \tag{6.32}$$

When a single particle is sent into one of the two ports of this channel, the resulting evolution is



(a) **(b)**

Figure 6.4: *Performance in the transmission of a single particle through a network of correlated depolarising channels, arranged as in Figure 6.3.* (a) Classical capacity in the lack of correlations. Without loss of generality, $\varphi_0 = 0$. The maximum capacity is 0.018 bits. (b) Lower bound to the classical capacity achieved with maximal correlations corresponding to the probability distributions $p_A(m,n) = p_B(m,n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3. Without loss of generality, $\varphi_0 = \varphi_2 = 0$. The maximum lower bound is 0.31 bits.

described by the superposition channel

$$\mathcal{S}^{\omega}_{\text{sup}}\left[\mathcal{Z}(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}})\right]=\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{B}}_2\widetilde{\mathcal{A}}_1\otimes\widetilde{\mathcal{A}}_2\widetilde{\mathcal{B}}_1)\,. \tag{6.33}$$

We now restrict our attention to the case where the channels $\widetilde{\mathcal{A}}_1$, $\widetilde{\mathcal{A}}_2$, $\widetilde{\mathcal{B}}_1$, and $\widetilde{\mathcal{B}}_2$ are all equal to each other, and are all equal to $\widetilde{\mathcal{D}}$, a vacuum extension of the completely depolarising channel. In this case, the action of the superposition channel is

$$\mathcal{E}_\omega=\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{D}}^2\otimes\widetilde{\mathcal{D}}^2)(\rho)=\frac{I/d+F^2\rho F^{2\dagger}}{2}\otimes\omega+\frac{I/d-F^2\rho F^{2\dagger}}{2}\otimes Z\omega Z\,, \tag{6.34}$$

where F is the vacuum interference operator associated to channel $\widetilde{\mathcal{D}}$. The above equation follows from Eq. (6.17) and from the observation that the vacuum interference operator of $\widetilde{\mathcal{D}}^2$ is F^2.

Note that one has the equality

$$\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{D}}^2\otimes\widetilde{\mathcal{D}}^2)(\rho)\equiv\mathcal{C}_{\omega,F^2}(\rho)\,, \tag{6.35}$$

using the notation of Eq. (6.17). That is, in the lack of correlations, the configuration of channels depicted in Figure 6.3 gives rise to the effective channel in Equation (6.17), with F replaced by F^2.

The evaluation of the maximum capacity follows the same steps as in §6.2.3. The main observations are:

1. in the lack of correlations, the channel $\mathcal{E}_\omega$ in Eq. (6.25) is entanglement-breaking, and therefore its classical capacity coincides with the Holevo capacity

2. the control state ω that maximises the Holevo capacity of the channel $\mathcal{E}_\omega$ is $\omega=|+\rangle\langle+|$

3. without loss of generality, the maximisation of the Holevo capacity can be reduced to ensembles that depend only on three real parameters q, p_0, and p_1 in $[0,1]$.

Observations 1.–2. follow from the same results used to prove the analogous claims in §6.2.3. Observations 3 follows from the following. The equality (6.35) means that Theorem 18 and Corollary 19 apply to this scenario as well, with F replaced by F^2. In particular, the classical capacity can be determined numerically using Corollary 19, with the maximisation constraint now being for the vacuum interference operator $F^2=g\,|0\rangle\langle 0|+h\,|1\rangle\langle 1|,\quad g+h\leq 1/d$, where $g,h\geq 0$. This proves Observations 3.

Having established observations 1.–3. above, we use Theorem 18 to evaluate the capacity of the channel $\mathcal{E}_\omega$ in Eq. (6.25) by scanning all possible values of the phases $\{\varphi_m\}_{m=0}^3$. The result is the plot shown in Figure 6.4a. The largest classical capacity over all random unitary realisations is 0.018 bits, which is strictly smaller than the value 0.049 bits achieved by the superposition of orders.

Finally, we extend the optimisation from random unitary realisations to arbitrary realisations of the completely depolarising channel, using Corollary 19. For this broader class of realisations, we numerically obtain that the maximum capacity is 0.024 bits.

In summary, the best classical capacity that can be obtained by sending a single particle through the network described by Eq. (6.34), in the lack of correlations between the two paths, is 0.024 bits, and when restricted to random unitary realisations of the completely depolarising channel is 0.018 bits.

6.3.4 Time correlations surpassing the quantum SWITCH capacity

We now show that the classical capacity of 0.049 bits, achieved by the quantum SWITCH, can be surpassed using more general time correlations. We prove this result explicitly, by exhibiting a pair of time-correlated channels that achieve a capacity at least 0.31 bits.

In particular, suppose that the unitaries $\{V_m\}_{m=0}^{d^2-1}$ form an orthogonal basis, and that the probability $p(m,n)$ has the form $p(m,n) = \delta_{n,\sigma(m)}/d^2$, for a permutation σ that makes $p(m,n)$ symmetric. In this case, the marginal probabilities $p(m) = p(n) = 1/d^2$, i.e. we recover the completely depolarising channel at each individual time step, and Eq. (6.29) becomes

$$\mathcal{E}_\omega(\rho) := \mathcal{S}^\omega_{\text{sup}}\left[\mathcal{Z}\left(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B\right)\right](\rho) = \frac{I/d + \mathcal{K}(\rho)}{2} \otimes \omega + \frac{I/d - \mathcal{K}(\rho)}{2} \otimes Z\omega Z\,, \tag{6.36}$$

with

$$\mathcal{K}(\rho) = \frac{1}{d^4}\sum_{m,k} V^{(B)}_{\sigma(k)} V^{(A)}_m \rho V^{(B)\dagger}_k V^{(A)\dagger}_{\sigma(m)}\, e^{i\left[\varphi^{(A)}_{\sigma(m)} + \varphi^{(B)}_k - \varphi^{(A)}_m - \varphi^{(B)}_{\sigma(k)}\right]}. \tag{6.37}$$

We now specialise further to the case where $d = 2$ and σ is the permutation that exchanges 0 with 1, and 2 with 3. This choice is motivated by the fact that the permutation σ guarantees the maximum communication capacity in the case where a single time-correlated channel is used (cf.

§6.2.2). With this choice, we obtain

$$\mathcal{K}(\rho) := \frac{1}{8}\left\{\left[\cos 2(\varphi_1 - \varphi_0) + \cos 2(\varphi_3 - \varphi_2)\right] \rho + 2X\rho X\right\} . \tag{6.38}$$

Note that the channel $\mathcal{E}_\omega$ depends only on the phase differences $\varphi_1 - \varphi_0$ and $\varphi_3 - \varphi_2$, via Eq. (6.38).

We now provide a lower bound to the classical capacity of the channel $\mathcal{E}_\omega$. As we did in §6.2.4, we lower bound the classical capacity by the Holevo capacity, and, in turn, we lower bound the Holevo capacity by restricting the maximisation to orthogonal input ensembles. For the state of the control qubit, we pick $\omega = |+\rangle\langle+|$, which is the choice that maximises the Holevo capacity (cf. Lemma 14).

The lower bound to the classical capacity is shown in Figure 6.4b for all possible values of the phases φ_1 and φ_3. The highest lower bound over all combinations of phases $\{\varphi_m\}_{m=0}^{3}$ is given by 0.31 bits. This value is larger than the classical capacity of 0.049 bits achieved by the quantum `SWITCH`, corresponding to perfect correlations $p_A(m,n) = p_B(m,n) = \delta_{m,n}/4$. This result implies that not only can time correlations reproduce the superposition of causal orders, but they can also surpass its advantages.

6.4 Quantum communication through correlated dephasing noise

In this section, we consider enhancements in the communication of *quantum* information through correlated superpositions of channels. As in the previous sections, we consider 2-step random unitary channels of the form (6.1) with vacuum extensions of the form (6.3). We begin by presenting a result very similar to the dephasing example of §5.3, this time formulated as a realisation in terms of random unitary channels. This shows that the superposition placement can transform two independent zero-quantum-capacity dephasing channels into a channel with non-zero quantum capacity. We then present an extension of this result to the case of correlated channels, showing that correlations can enable dephasing channels with zero quantum capacity when used at a definite time step to be transformed into a *perfect quantum* communication channel when used at an indefinite time.

6.4.1 Probabilistic correction of dephasing noise on independent paths

Consider the dephasing channel (cf. §2.2.3)

$$\mathcal{D}_1(\rho) = \frac{1}{2}\rho + \frac{1}{2}Z\rho Z, \tag{6.39}$$

which implements completely dephasing noise on the message degree of freedom. This has zero quantum capacity, i.e. cannot transmit any quantum information.

The full physical description of this process including its action on the vacuum is given by the vacuum-extended channel

$$\widetilde{\mathcal{D}}_1(\rho) = \frac{1}{2}(I \oplus e^{i\varphi_0} I_{\text{Vac}})\rho(I \oplus e^{-i\varphi_0} I_{\text{Vac}}) + \frac{1}{2}(Z \oplus e^{i\varphi_3} I_{\text{Vac}})\rho(Z \oplus e^{-i\varphi_3} I_{\text{Vac}}), \tag{6.40}$$

for some phases φ_0, φ_3, where the second term of each direct sum corresponds to the one-dimensional vacuum sector with $I_{\text{Vac}} = |\text{vac}\rangle\langle\text{vac}|$.

Consider now two such channels, where the noise probabilities on each channel are independent (i.e. $p(m,n) = p(m)p(n) = 1/2 \cdot 1/2 = 1/4$). When used in a superposition of paths with the path initialised in the $|+\rangle$ state, this results in the effective channel

$$\mathcal{C}_{|+\rangle\langle+|}(\rho) = \mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}(\widetilde{\mathcal{D}}_1, \widetilde{\mathcal{D}}_1)(\rho) = \frac{\mathcal{D}_1(\rho) + F\rho F^\dagger}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{D}_1(\rho) - F\rho F^\dagger}{2} \otimes |-\rangle\langle-| \,, \tag{6.41}$$

where

$$F = \frac{1}{2}\sum_{m=0,3} e^{-i\varphi_m} V_m = \frac{Ie^{-i\varphi_0} + Ze^{-i\varphi_3}}{2} \,, \tag{6.42}$$

with $V_0 = I,\ V_3 = Z$.

We quantify the transmission of quantum information through this channel by its quantum capacity. However, the quantum capacity is hard to evaluate and has only been computed in very few cases [126, 127]. A useful lower bound can be found by considering the coherent information of the channel (§2.2), which in turn has a useful lower bound given by the coherent information of the bipartite state resulting from sending half of the maximally entangled state $|\Phi^+\rangle := (|00\rangle + |11\rangle)/\sqrt{2}$ through the channel. We shall call this quantity the *coherent information with respect to the maximally entangled state*.

The coherent information with respect to the maximally entangled state is plotted against the

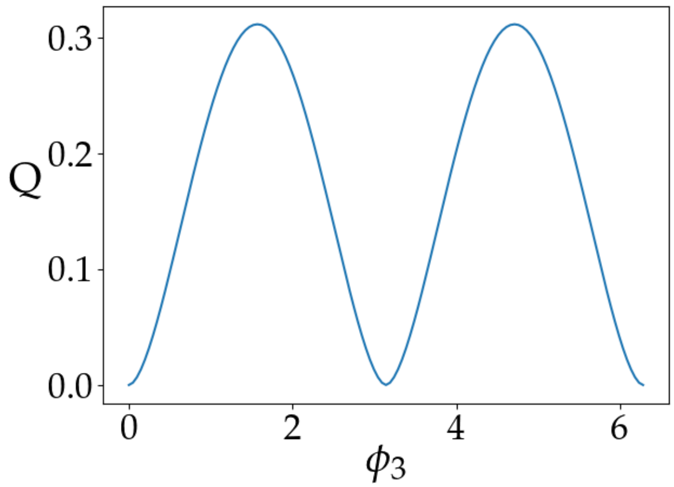


Figure 6.5: *Coherent information (with respect to the maximally entangled state) Q against the phase φ_3 for independent dephasing noise.* W.l.o.g., we set $\varphi_0 = 0$.

phase φ_3 in Figure 6.5. Since $F\rho F^\dagger$ depends only on the phase difference $\varphi_3 - \varphi_0$, we can set $\varphi_0 = 0$ without loss of generality. This shows that the effective channel $\mathcal{C}_{|+\rangle\langle+|}$ has non-zero quantum capacity for all but a vanishingly small region of the phase differences.

In particular if $\varphi_3 - \varphi_0 = 3\pi/2$, then

$$\mathcal{C}_{|+\rangle\langle+|}(\rho) = \left(\frac{\mathcal{D}_1(\rho)}{2} + \frac{Q_+ \rho Q_+^\dagger}{4} \right) \otimes |+\rangle\langle+| + \frac{Q_- \rho Q_-^\dagger}{4} \otimes |-\rangle\langle-| \,, \tag{6.43}$$

where

$$Q_\pm := \frac{I \pm iZ}{\sqrt{2}} \,. \tag{6.44}$$

$Q_\pm$ is proportional to a unitary, so upon measurement of the control in the $|-\rangle$ state, which occurs with 25% probability, the effective channel enables perfect correction of the noise.

A similar result holds for $\varphi_3 - \varphi_0 = \pi/2$, with $Q_+ \leftrightarrow Q_-$.

However, note that this superposition of independent channels can never lead to an effective channel with perfect quantum capacity, as per the observation in §5.4. We now show that, in contrast, this is possible for a superposition of correlated channels.

6.4.2 Perfect correction of dephasing noise on correlated paths

Consider now the 2-step random unitary channel

$$\mathcal{D}_{12}(\rho) = \frac{1}{2} I \otimes Z \rho I \otimes Z + \frac{1}{2} Z \otimes I \rho Z \otimes I \,, \tag{6.45}$$

which reduces to the dephasing channel (6.39) above when used at a definite time step. This is equivalent to a random unitary channel of the form (6.1), with maximal correlations described by the probability distribution $p(m,n) = \delta_{m,\sigma(n)}/2$, with $m, n \in \{0, 3\}$, where σ is a permutation that maps $0 \leftrightarrow 3$. Again, if a single photon is sent at a definite time through the channel, no quantum information can be transmitted. We again take the vacuum extension of each unitary to be another unitary, so that the full physical process is described by the vacuum extension

$$\begin{aligned} \widetilde{\mathcal{D}}_{12}(\rho) =& \frac{1}{2}(I \oplus e^{i\varphi_0} I_{\text{Vac}}) \otimes (Z \oplus e^{i\varphi_3} I_{\text{Vac}}) \rho (I \oplus e^{-i\varphi_0} I_{\text{Vac}}) \otimes (Z \oplus e^{-i\varphi_3} I_{\text{Vac}}) \\ &+ \frac{1}{2}(Z \oplus e^{i\varphi_3} I_{\text{Vac}}) \otimes (I \oplus e^{i\varphi_0} I_{\text{Vac}}) \rho (Z \oplus e^{-i\varphi_3} I_{\text{Vac}}) \otimes (I \oplus e^{-i\varphi_0} I_{\text{Vac}}) \,. \end{aligned} \tag{6.46}$$

When used at a superposition of times, this results in the effective channel

$$\mathcal{C}_{|+\rangle\langle+|}(\rho) = \mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}(\mathcal{D}_{12})(\rho) = \frac{\mathcal{D}_1(\rho) + \mathcal{G}(\rho)}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{D}_1(\rho) - \mathcal{G}(\rho)}{2} \otimes |-\rangle\langle-| \,, \tag{6.47}$$

where

$$\mathcal{G}(\rho) := \frac{1}{2} \sum_{m=0,3} e^{i[\varphi_{\sigma(m)} - \varphi_m]} V_m \rho V_{\sigma(m)}^\dagger . \tag{6.48}$$

For this particular pattern of maximal correlations, when a single particle is sent through a superposition of times through the 2-step channel, the effective channel has up to *perfect* quantum capacity of 1 qubit per channel use, depending on the phase difference $\varphi_3 - \varphi_0$. The coherent information plotted against the phase φ_3 is shown in Figure 6.6.

In particular if $\varphi_3 - \varphi_0 = 3\pi/2$, then

$$\mathcal{C}_{|+\rangle\langle+|}(\rho) = \frac{Q_+ \rho Q_+^\dagger}{2} \otimes |+\rangle\langle+| + \frac{Q_- \rho Q_-^\dagger}{2} \otimes |-\rangle\langle-| \,, \tag{6.49}$$

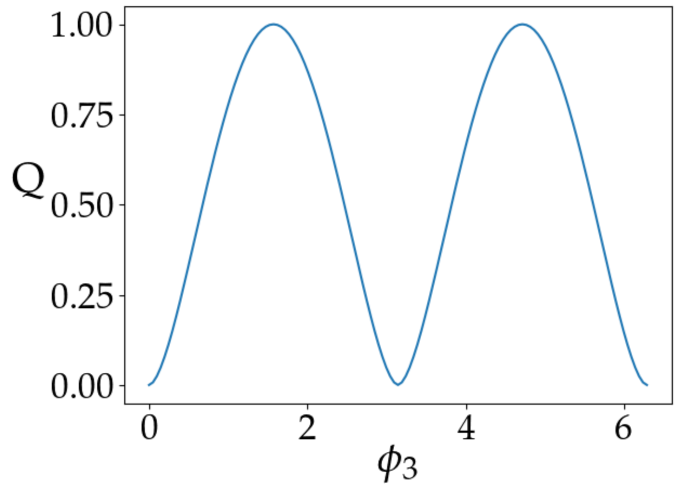


Figure 6.6: *Coherent information (with respect to the maximally entangled state) Q against the phase φ_3 for correlated dephasing noise with maximal correlations described by the probability distribution $p(m,n) = \delta_{m,\sigma(n)}/2$, for $m,n \in \{0,3\}$, where σ is the permutation that maps* $0 \leftrightarrow 3$. W.l.o.g., we set $\varphi_0 = 0$.

where

$$Q_{\pm} := \frac{I \pm iZ}{\sqrt{2}}\,. \tag{6.50}$$

A similar result holds for $\varphi_3 - \varphi_0 = \pi/2$, with $Q_+ \leftrightarrow Q_-$. Both Q_+ and Q_- are proportional to a unitary, so the effective channel (6.49) has perfect quantum capacity of 1 qubit per channel use.

Thus we have shown that a time-correlated transmission line, which has zero quantum capacity when used at any definite time, can be transformed into a perfect communication channel for qubits, by coherently controlling the time of transmission of a single particle.

7

Superpositions of causal orders vs. superpositions of trajectories

In this chapter, we return to the question of comparing the superpositions of causal orders and the superposition trajectories, first mentioned in §2.5.4. This comparison was first raised in Ref. [2], where Abbott et al. gave an example of a communication protocol where two completely depolarising channels are coherently superposed.

The authors quantified the transmission of information in terms of the Holevo capacity, and showed that the Holevo capacity achievable by combining the two channels in a superposition is greater than the Holevo capacity achievable by combining them using the quantum SWITCH. The value of the former Holevo capacity is 0.16 bit [2] (which we in fact showed is equal to the classical capacity in Chapter 6), whilst the value of the latter Holevo capacity is 0.049 bits [80] (which was later shown to be equal to the classical capacity in Ref. [59]). As a result, the authors argued that the communication advantages of the quantum SWITCH '*should therefore rather be understood as resulting from coherent control of quantum communication channels*,' as opposed to being specifically due to indefinite causal order.

In the following, we shall analyse this claim in detail, and argue that from a resource-theoretic point of view, the underlying assumptions of this claim do not hold water. Finally, we make a comparison between the superposition of independent channels and the corresponding superposition of correlated channels that can reproduce the action of the quantum SWITCH (as shown in Chapter 6). We argue that this comparison suggests that the communication advantages associated with the quantum SWITCH in fact arise from an interplay between coherent control and the correlations created by an indefinite causal order.

7.1 Resource-theoretic comparison

The above observation of Abbott et al. is presented as a comparison between two alternative ways to transform two completely depolarising channels into a new quantum channel with non-zero capacity. However, we shall argue that their conclusions do not follow from their observation.

First, it is not clear how a comparison between the values of the Holevo capacity (or classical capacity) for the quantum SWITCH and for the superposition of channels could be used to make a definite statement about the 'origin' of their respective advantages. At most, the comparison could show that the ability to control trajectories is more powerful than the ability to control causal orders.

Second, the comparison made in Ref. [2] is uneven, because

1. it does not compare supermaps acting on the same input channels, and

2. it does not compare superpositions of channels where the depolarising channels act the same number of times.

A detailed analysis of these two points is provided in the following.

1. Different input channels. In quantum Shannon theory with superpositions of trajectories, the input resources are vacuum-extended channels, while in quantum Shannon theory with superpositions of causal orders the input resources are ordinary (non-vacuum-extended) channels.

A vacuum-extended channel is a stronger resource than the corresponding channel, because it can have coherence with the vacuum, in the sense of Definition 6. We now argue that coherence with the vacuum is indeed the underlying resource implicit in the communication advantages of

Ref. [2]. Suppose that a particle is sent in a superposition of two paths, going through two communication devices, each of which acts as a completely depolarising channels on the internal degree of freedom of the particle. The two devices are described by vacuum extensions of the completely depolarising channel, and act as

$$\begin{aligned}\widetilde{\mathcal{N}}_{\text{dep}}(\rho) =& (1-\langle\text{vac}|\rho|\text{vac}\rangle)\,\frac{I}{d}+\langle\text{vac}|\rho|\text{vac}\rangle\,|\text{vac}\rangle\langle\text{vac}| \\ &+F\rho\,|\text{vac}\rangle\langle\text{vac}|+|\text{vac}\rangle\langle\text{vac}|\,\rho\,F^{\dagger}\,,\end{aligned} \tag{7.1}$$

where F is the vacuum interference operator defined in Equation (3.12). Now, if the channels have no coherence with the vacuum (that is, if $F = 0$), then their superposition yields the constant channel

$$\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}},\widetilde{\mathcal{N}}_{\text{dep}})(\rho)=\frac{I}{d}\otimes\text{diag}(\omega)\,, \tag{7.2}$$

following from Equation (4.30) with $\rho_0 = I/d$. Since the output is independent of the input, the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}},\widetilde{\mathcal{N}}_{\text{dep}})$ cannot be used to communicate.

The above analysis shows that the presence of coherence with the vacuum is necessary for the advantages observed by Abbott *et al.* [2]. In contrast, the presence of coherence with the vacuum is, in principle, unnecessary for the advantages of the quantum SWITCH. For example, the implementation of the quantum SWITCH via closed timelike curves [53, 57], illustrated in Figure 7.1, does not require any coherence with the vacuum.

In summary, the advantages of the superposition of causal orders and the superposition of trajectories arise from different input resources, with the resources used in the latter (vacuum-extended channels exhibiting coherence with the vacuum) being strictly stronger than the resources used in the former (ordinary, non-vacuum-extended channels).

2. *Different numbers of uses of the depolarising channel.* Superpositions of trajectories and superpositions of causal orders refer to two different communication scenarios:

(a) in superpositions of trajectories, the particle travels through *only one* depolarising channel (*either* $\mathcal{N}_1$ *or* $\mathcal{N}_2$),

(b) in superpositions of causal orders the particle travels through *two* depolarising channels (*both* $\mathcal{N}_1$ *and* $\mathcal{N}_2$).

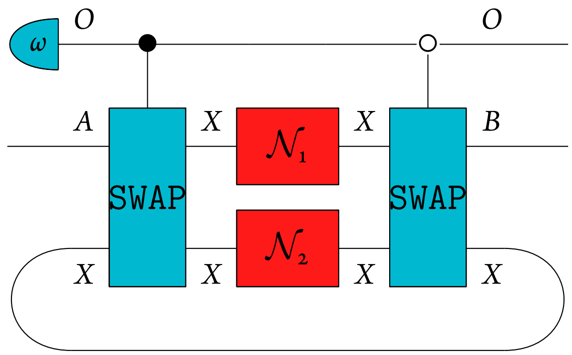


Figure 7.1: *An implementation of the quantum* SWITCH *placement (in blue) using closed timelike curves.* A quantum state $\rho \in \mathrm{St}(A)$ is routed through one of the two channels $\mathcal{N}_1$ or $\mathcal{N}_2$ by a SWAP gate controlled by the state of the order qubit $\omega \in \mathrm{St}(O)$. A second SWAP gate (controlled in the opposite way) routes the state to a closed timelike curve, which transfers the incoming system back through the first SWAP gate, and through one of the two channels $\mathcal{N}_2$ or $\mathcal{N}_1$.

From this point of view, there is little surprise that Scenario (a) allows more communication than Scenario (b), given that in Scenario (b) the particle is exposed twice to depolarising noise, as acknowledged also in Ref. [2].

One may argue that the difference between Scenarios (a) and (b) is irrelevant, because the completely depolarising channel $\mathcal{N}_{\rm dep}(\cdot) := \mathrm{Tr}(\cdot) I/d$ satisfies the equality

$$\mathcal{N}_{\rm dep} \circ \mathcal{N}_{\rm dep} = \mathcal{N}_{\rm dep} \,, \tag{7.3}$$

meaning that applying the channel twice in a row is the same as applying it once.

However, the input resource for the superposition of two channels is not two depolarising channels themselves, but rather their vacuum extensions. Crucially, algebraic identities like the one in Equation (7.3) do not carry over to the vacuum extensions: in general, the relation $\mathcal{N}_1 \circ \mathcal{N}_2 = \mathcal{N}_3$ *does not* imply the relation $\widetilde{\mathcal{N}}_1 \circ \widetilde{\mathcal{N}}_2 = \widetilde{\mathcal{N}}_3$. In the particular case of depolarising channels, we have the following result:

Proposition 20. *The condition $\widetilde{\mathcal{N}}_{\rm dep} \circ \widetilde{\mathcal{N}}_{\rm dep} = \widetilde{\mathcal{N}}_{dep}$ is satisfied if and only if the vacuum extension $\widetilde{\mathcal{N}}_{\rm dep}$ has no coherence with the vacuum.*

The proof is given in Appendix A.5.

In summary, the only case in which the equation $\widetilde{\mathcal{N}}_{\rm dep} \circ \widetilde{\mathcal{N}}_{\rm dep} = \widetilde{\mathcal{N}}_{\rm dep}$ would justify a comparison between the Holevo (or classical) capacity with a single depolarising channel and the Holevo (or classical) capacity with two depolarising channels is exactly the case in which the vacuum extension $\widetilde{\mathcal{N}}_{\text{dep}}$ has no coherence with the vacuum, and therefore the superposition of channels provides no advantage.

In order to make a more even comparison with the quantum `SWITCH`, one should analyse the scenario where information is sent along a superposition of two paths, each visiting *two* depolarising channels. Mathematically, this superposition is described by the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}})$, instead of the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}})$ considered in Ref. [2].

The Holevo capacity of the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}})$ was studied in Chapter 6, where we found that the maximum value of the Holevo capacity over all possible vacuum extensions is 0.024 bits, which we also proved was equal to the classical capacity. This is in fact less than the classical capacity of 0.049 bits associated with the quantum `SWITCH`.

To illustrate the difference between the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}})$ and the channel $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}})$, we calculate the Holevo capacity (which is equal to the classical capacity in this specific case) for both cases when the depolarising channel is implemented by a uniform randomisation of Pauli channels (cf. Eqs. (6.5) and (6.25), respectively), where the vacuum amplitude corresponding to each unitary V_m is simply a complex phase $e^{i\varphi_m}$. Figure 7.2 show a scatter plot with the capacities of both channels in the same graph against the norm of the corresponding vacuum interference operator, F or F^2, for the same combination of phases φ_1, φ_2, φ_3 as shown in Figures 6.1 and 6.4. This shows that the capacity generally increases with the norm of the vacuum interference operator, which is consistent with the idea that coherence with the vacuum is responsible for the capacity enhancement. Since $||F^2||_\infty < ||F||_\infty$ for all F constructed from random unitary channels with unitary vacuum extensions of the form (6.3), this provides an explanation of why the capacity of $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}})$ is always less than the capacity of $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}})$.

For the above reasons, we argue that the comparison between the quantum `SWITCH` and superposition of independent communication channels presented in Ref. [2] is uneven.

7.2 Further discussion

The initial question posed in Ref. [2], namely to what extent indefinite causal order *per se*, as opposed to coherent control on its own, is responsible for the communication advantages of the quan-

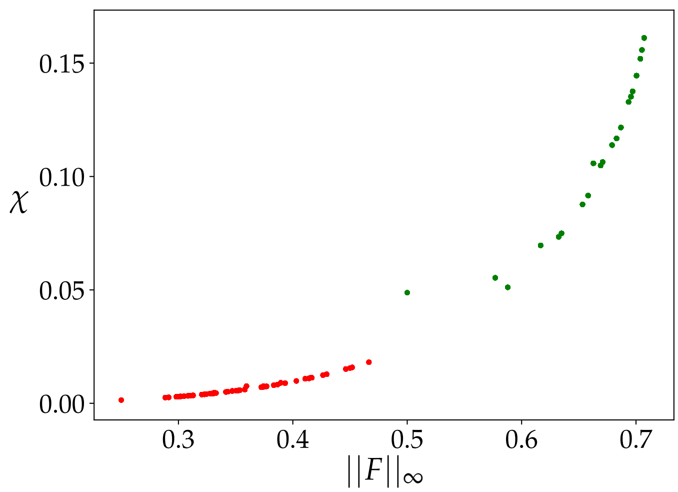


Figure 7.2: *Green: A plot of the classical capacity against the norm of the vacuum interference operator* $||F||_\infty$ *for the channel* $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}})$. *Red: A plot of the classical capacity against the norm of the vacuum interference operator* $||F^2||_\infty$ *for the channel* $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}})$. In both cases $F = \sum_{m=0}^{3} \frac{1}{4} e^{-i\varphi_m} V_m$ and is sampled over the phase parameters $\{\varphi_1, \varphi_2, \varphi_3\}$ with a numerical precision of $\pi/8$ for each parameter. We set $\varphi_0 = 0$ without loss of generality, as $F\rho F^\dagger$ is invariant under the phase group $U(1)$. The classical capacity is here equal to the Holevo capacity and the Holevo capacity was calculated using Theorem 18.

tum SWITCH is an important one. We address this question in the following way. As argued above, a direct resource-theoretic comparison between the superposition of causal orders and superposition of trajectories is difficult to make. However, as shown in §6.3, the quantum SWITCH of two channels can always be reproduced by a superposition of paths through two 2-step correlated channels, connected as in Figure 6.3. This means that we can instead make a direct comparison using the superposition of paths through two 2-step correlated channels, as in Figure 6.3, between (a) the case with correlations that reproduce the quantum SWITCH and (b) the case with trivial correlations, where the two uses of each channel factories into independent uses.

First, consider the case of classical communication through completely depolarising noise. Recall from §6.3, that $\mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}, \widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}}) = \mathcal{S}^{\omega}_{\text{sup}}\left[\mathcal{Z}\left(\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B\right)\right]$, as defined in Eq. (6.29) with the probability distributions in both $\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B$ given by $p(m,n) = p(m)p(n) = 1/4 \times 1/4 = 1/16$, for $m, n \in \{0, 1, 2, 3\}$. We showed that the maximum classical capacity achievable by a channel of this form is 0.024 bits.

However, for the same channel with the probability distributions in both $\widetilde{\mathcal{R}}_A, \widetilde{\mathcal{R}}_B$ given by $p(m,n) = \delta_{mn}/4$ for $m, n \in \{0, 1, 2, 3\}$, we recover the ability to reach the quantum SWITCH capacity of 0.049 bits.

Second, consider the case of quantum communication through completely dephasing noise. Refs. [58, 168] showed that the quantum SWITCH enables noiseless quantum communication through two dephasing channels of the form $(X\rho X + Y\rho Y)/2$. This can be reproduced by a pair of 2-step channels of the form (6.46), with $I \to X,\ Z \to Y$ and any choice of $\{\varphi_m\}$, placed in a superposition as in Eq. (6.29), corresponding to a correlated probability distribution $p(m,n) = \delta_{mn}/2$ for $m, n \in \{1, 2\}$.

However, as shown in §5.4, the superposition of two independent channels can never lead to a noiseless channel, unless the original channels are themselves unitary channels. This means that the corresponding superposition of independent channels, that is, with $p(m,n) = 1/2 \times 1/2 = 1/4$ for $m, n \in \{1, 2\}$, will never reach the advantages associated with the correlated case.

This shows that when correlations are present in the probability distribution, the superposition of channels can achieve noiseless transmission of both classical and quantum information, which is impossible in the analogous scenarios without correlations. Therefore, we conclude that the communication advantages in the correlated case are not merely due to the coherent control of the channels, but rather due to the non-trivial interplay between coherent control and the time-correlations in the noise.

The superposition of causal orders introduces correlations between the two branches of the superposition that are equivalent to the correlations induced by using two time-correlated channels in a superposition of paths. Therefore, although we may not be able to make a formal resource-theoretic comparison between the superposition of causal orders and the superposition of trajectories, we argue that it is reasonable to understand the communication advantages achievable by the quantum `SWITCH` as arising from a combination of the correlations due to indefinite causal order and coherent control.

8

Communication networks with superpositions of paths

Communicating classical or quantum information over large distances via single photons is a crucial feature of many quantum protocols, such as quantum key distribution [28, 81] or superdense coding [29]. Yet, the distance at which information can be transmitted via quantum states is limited by both photon loss and noise, with the probability of each of these errors scaling exponentially with distance [42]. This means that in a quantum communication network, information can travel over only a finite number of nodes before the detection probability of the photon becomes effectively zero.

In this chapter, we present a scheme involving quantum repeaters on intermediate nodes that can mitigate noisy processes on the internal degree of freedom of a quantum particle. The central feature of our scheme is the transmission of a single quantum particle in a superposition of alternative paths through the communication network. In the idealised case, the protocol enables classical communication at a finite rate through asymptotically long sequences of noisy quantum channels.

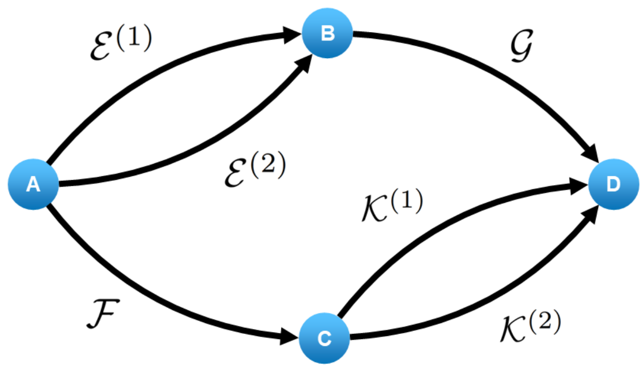


Figure 8.1: *Quantum network with coherent control of the communication paths.* The vertices represent the communicating parties, while the arrows represent the available communication channels. When multiple communication paths are available, the information carriers can propagate in a superposition of alternative paths.

In more realistic scenarios, our scheme provides a finite classical capacity advantage for various sequences of noisy channels.

8.1 Erasure error model

We begin with a simple, physical model of noise on a d-dimensional internal degree of freedom of a quantum particle, e.g., for $d = 2$, the polarisation of a single photon. Consider the case where the particle propagates from an input node via several intermediate nodes to an output node in a communication network (Figure 8.1). We specialise to the case where transmission from one node to another is given by a quantum erasure channel[1]

$$\mathcal{E}(\rho) = p\,|\eta\rangle\langle\eta| + (1-p)\rho, \tag{8.1}$$

which acts on the internal degree of freedom carrying the message. Here, $\rho \in \mathsf{St}(M)$ is the input quantum state on the d-dimensional system M of the message, $|\eta\rangle \in \mathcal{H}_M$ is some fixed error state,

[1]Here we take the erasure state to lie in the same Hilbert space as the message, which differs from some sources in the literature that define the erasure state in the quantum erasure channel to lie outside of the message Hilbert space [31].

and p is the erasure probability. After the particle propagates through n nodes, its state is

$$\mathcal{E}^n(\rho) = [1 - (1-p)^n] |\eta\rangle\langle\eta| + (1-p)^n \rho\,, \tag{8.2}$$

where the probability of detecting the original input state ρ exponentially decays to zero as the number of nodes n increases.

8.2 Superposition of sequences of identical channels

We now generalise the above scenario to the possibility of sending the particle through a superposition of alternative paths through the network. As such, we now model the transmission between any two nodes by a vacuum extension $\widetilde{\mathcal{E}}$ of the quantum erasure channel.

Consider the transmission of a single particle in a superposition of two paths, each through n copies of independent identical channels $\widetilde{\mathcal{E}}$, as depicted in Figure 8.2. If the control is initialised in the $|+\rangle$ state, then the output state is

$$\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}(\widetilde{\mathcal{E}}^n, \widetilde{\mathcal{E}}^n)(\rho) = \frac{\mathcal{E}^n(\rho) + F^n \rho F^{\dagger^n}}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{E}^n(\rho) - F^n \rho F^{\dagger^n}}{2} \otimes |-\rangle\langle-|\,. \tag{8.3}$$

(This can be found by applying Eq. (4.25) n times.) Consider this output state in the asymptotic limit $n \to \infty$. It is apparent from Eq. (8.2) that $\mathcal{E}^n \to |\eta\rangle\langle\eta|$ in this limit (assuming $p > 0$). If the largest eigenvalue of F has modulus strictly less than 1, then $F^n \rho F^{\dagger^n} \to 0$; in this case, the output state is $|\eta\rangle\langle\eta| \otimes I_C$, which has only a trivial dependence on the input state ρ. If, however, F has an eigenvector $|\varphi\rangle$ with eigenvalue 1, then $F^n \to |\varphi\rangle\langle\varphi|$, $F^n \rho F^{\dagger^n} \to \langle\varphi|\rho|\varphi\rangle |\varphi\rangle\langle\varphi|$, and the output state has a non-trivial dependence on the input state ρ. As such, one expects that the superposition channel can transmit information at a non-zero rate, even in the asymptotic limit of an infinite sequence of noisy channels.

8.3 Communication through a superposition of asymptotically long sequences of identical erasure channels

Let us consider what a physically motivated vacuum extension of the erasure channel could look like. An erasure process whereby any input state is replaced, with probability p, by some fixed error

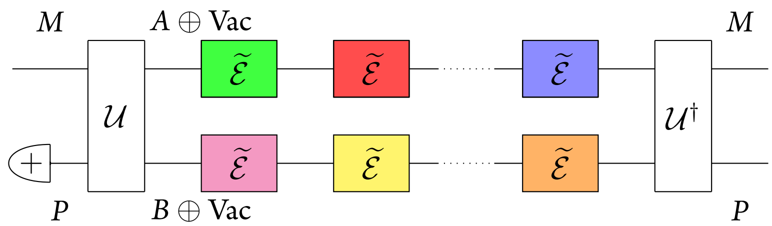


Figure 8.2: *A circuit diagram of a superposition of two sequences of n channels $\mathcal{E}$, specified by the vacuum extensions $\widetilde{\mathcal{E}}$.* The message is encoded in system $M \simeq A \simeq B$. Note that each channel is drawn in a different colour because they each represent an independent transmission line.

state $|\eta\rangle$ occurs, for example, upon a swap interaction between the message M and an environment E initially in state $|\eta\rangle$ (where $M \simeq E$). That is, the message-environment system is subject to the unitary interaction $W_{ME} = \texttt{SWAP}_{ME}$ with probability p. Since no interaction occurs if the vacuum state is input in place of the message, the overall vacuum-extended interaction is

$$\widetilde{W}_{\widetilde{M}E} = \texttt{SWAP}_{ME} \oplus I_{\mathrm{Vac}E} \,, \tag{8.4}$$

where $\widetilde{M} = M \oplus \mathrm{Vac}$. The vacuum extension $\widetilde{\mathcal{E}}$ of the erasure channel is then

$$\widetilde{\mathcal{E}}(\rho) = \mathrm{Tr}_E \left(p\, \widetilde{W}_{\widetilde{M}E} \rho \otimes |\eta\rangle\langle\eta|_E\, \widetilde{W}^\dagger_{\widetilde{M}E} + (1-p)\rho \otimes |\eta\rangle\langle\eta|_E \right) \,, \tag{8.5}$$

with Kraus operators

$$\begin{aligned} \widetilde{E}_i &= \sqrt{p}|\eta\rangle\langle i| \oplus \sqrt{p}\,\langle i|\eta\rangle\, I_{\mathrm{Vac}} \quad \text{for } i = 0, \ldots, r-1 \,, \\ \widetilde{E}_r &= \sqrt{1-p} I_{\widetilde{M}} \,, \end{aligned} \tag{8.6}$$

where we can identify $E_i = \sqrt{p}|\eta\rangle\langle i|$ and $\alpha_i = \sqrt{p}\,\langle i|\eta\rangle$ for $i = 0, \ldots, r-1$, and $E_r = \sqrt{1-p} I_M$, $\alpha_r = \sqrt{1-p}$.

From this physical realisation of an erasure channel and its vacuum extension, we obtain a vacuum interference operator $F = |\eta\rangle\langle\eta| + (1-p)(I_M - |\eta\rangle\langle\eta|)$, whose eigenvector $|\eta\rangle$ has eigenvalue 1. Assuming $p > 0$, $F^n \to |\eta\rangle\langle\eta|$ as $n \to \infty$. If the path is initialised in the $|+\rangle$ state, then, according to Eq. (8.3), the superposition of two asymptotically long sequences of identical channels $\widetilde{\mathcal{E}}$ is given by

$$\lim_{n\to\infty} \mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}(\widetilde{\mathcal{E}}^n, \widetilde{\mathcal{E}}^n)(\rho) = |\eta\rangle\langle\eta| \otimes \left[q_{\rho,\eta}|+\rangle\langle+| + (1 - q_{\rho,\eta})\frac{I}{2} \right] , \tag{8.7}$$

where $q_{\rho,\eta} = \langle\eta|\rho|\eta\rangle$. This channel is equivalent to a classical binary asymmetric channel[2], which has classical capacity $\log_2(5/4) \approx 0.32$ [163].

Thus, by sending a single particle through a superposition of alternative paths through a communication network subject to erasure errors, one can achieve classical communication with a non-zero capacity of 0.32 bits, even in the extreme limit of an asymptotically long sequence of channels. This possibility relies upon the fact that the vacuum interference operator F has an eigenvalue of 1, so that the terms $\pm F^n \rho F^{\dagger n}$ do not exponentially decay to zero with the sequence length n.

8.4 Communication through a superposition of asymptotically long sequences of non-identical erasure channels with repeaters

So far, we have considered only networks whose errors arise from identical erasure channels, each with the same erasure probability p and error state $|\eta\rangle$. However, in practice, it is reasonable to expect variations within the error parameters between different parts of the network. For any chosen path through the network, the sequence of errors acting on the particle could be modelled as a sequence of (possibly) non-identical erasure channels $\mathcal{E}_n \circ \mathcal{E}_{n-1} \circ \cdots \circ \mathcal{E}_1$, with

$$\mathcal{E}_k(\rho) = p_k\, |\eta_k\rangle\langle\eta_k| + (1 - p_k)\rho\,, \tag{8.8}$$

where $k \in \{1, \ldots, n\}$, p_k is a probability and $|\eta_k\rangle$ an error state.

Let $\widetilde{\mathcal{E}}_k$ be a vacuum extension of the channel $\mathcal{E}_k$. A superposition of two identical sequences $\widetilde{\mathcal{E}}_n \circ \widetilde{\mathcal{E}}_{n-1} \circ \cdots \circ \widetilde{\mathcal{E}}_1$ yields a concatenated vacuum interference operator $F = F_n F_{n-1} \ldots F_1$, where each $F_k = |\eta_k\rangle\langle\eta_k| + (1 - p_k)(I_M - |\eta_k\rangle\langle\eta_k|)$ is the vacuum interference operator for $\widetilde{\mathcal{E}}_k$. Assuming that for all k, $p_k > 0$ and $|\langle\eta_{k+1}|\eta_k\rangle| < 1$, F approaches zero as $n \to \infty$. This decay of the vacuum interference operator can be circumvented by inserting, at each node in the network, an intermediate repeater $\widetilde{\mathcal{R}}_k$ engineered to have its own vacuum interference operator

[2]Interestingly, this expression is equal to that of a superposition of two single erasure channels (i.e., $n = 1$) with erasure probability $p = 1$ (cf. §5.2)

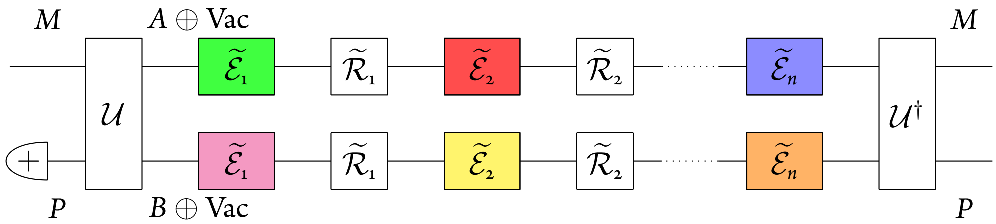


Figure 8.3: *A circuit diagram of a superposition of two sequences of n channels $(\mathcal{E}_1, \ldots \mathcal{E}_n)$, specified by the vacuum extensions $(\widetilde{\mathcal{E}}_1, \ldots, \widetilde{\mathcal{E}}_n)$, with intermediate repeaters $(\widetilde{\mathcal{R}}_1, \ldots, \widetilde{\mathcal{R}}_{n-1})$.* The message is encoded in system $M \simeq A \simeq B$.

$G_k = |\eta_{k+1}\rangle\langle\eta_k| + G_k^{\text{rest}}$, where G_k^{rest} is any operator such that G_k is a valid vacuum interference operator satisfying $G_k|\eta_k\rangle = |\eta_{k+1}\rangle$. It is always possible to construct such a repeater $\widetilde{\mathcal{R}}_k$ by defining the unitary channel $\widetilde{\mathcal{R}}_k = \widetilde{R}_k(\cdot)\widetilde{R}_k^\dagger$, with $\widetilde{R}_k = \left(|\eta_{k+1}\rangle\langle\eta_k| + G_k^{\text{rest}}\right) \oplus |\text{vac}\rangle\langle\text{vac}|$. This construction yields an effective vacuum interference operator for the whole sequence of erasure channels and repeaters: $F_{\text{eff}} = F_n G_{n-1} F_{n-1} \ldots G_1 F_1$.

The superposition of two identical sequences of non-identical erasure channels concatenated with repeaters is depicted in Figure 8.3. If the path is initialised in the $|+\rangle$ state, then the output state of the superposition channel is

$$\begin{aligned} &\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}\left[\left(\widetilde{\mathcal{E}}_n\widetilde{\mathcal{R}}_{n-1}\widetilde{\mathcal{E}}_{n-1}\cdots\widetilde{\mathcal{R}}_1\widetilde{\mathcal{E}}_1\right), \left(\widetilde{\mathcal{E}}_n\widetilde{\mathcal{R}}_{n-1}\widetilde{\mathcal{E}}_{n-1}\cdots\widetilde{\mathcal{R}}_1\widetilde{\mathcal{E}}_1\right)\right](\rho) \\ &= \frac{\mathcal{E}_n \circ \mathcal{R}_{n-1} \circ \mathcal{E}_{n-1} \circ \cdots \circ \mathcal{R}_1 \circ \mathcal{E}_1(\rho) + F_{\text{eff}}\rho F_{\text{eff}}^\dagger}{2} \otimes |+\rangle\langle+| \\ &+ \frac{\mathcal{E}_n \circ \mathcal{R}_{n-1} \circ \mathcal{E}_{n-1} \circ \cdots \circ \mathcal{R}_1 \circ \mathcal{E}_1(\rho) - F_{\text{eff}}\rho F_{\text{eff}}^\dagger}{2} \otimes |-\rangle\langle-|. \end{aligned} \tag{8.9}$$

If $p_k > 0$ for all k and if the sequence $\{|\eta_k\rangle\}_{k=1}^\infty$ has a well-defined limit $|\eta_*\rangle := \lim_{k\to\infty}|\eta_k\rangle$, then, in the asymptotic limit $n \to \infty$, the output state in Eq. (8.9) becomes

$$|\eta_*\rangle\langle\eta_*| \otimes \left[q_{\rho,\eta_1}|+\rangle\langle+| + (1 - q_{\rho,\eta_1})\frac{I}{2}\right], \tag{8.10}$$

where $q_{\rho,\eta_1} = \langle\eta_1|\rho|\eta_1\rangle$. This expression has the same form as Eq. (8.7), enabling classical communication at a non-zero rate through asymptotically long sequences of non-identical erasure channels.

8.5 General case

In the previous two sections, we showed that for erasure channels, superposing asymptotically long sequences of quantum channels enables the transmission of classical information at a non-zero rate. In this section, we present necessary and sufficient conditions for this phenomenon to be realised with a generic sequence of identical quantum channels.

Recall, that without repeater channels, we expect to be able to communicate non-zero classical information through a superposition of sequences of identical channels if and only if the vacuum interference operator F has eigenvalue 1. If this expectation is correct, then an even stronger statement should also hold: with repeater channels, one can communicate non-zero classical information through a superposition of sequences of identical channels if and only if F has singular value 1. Specifically, if the vacuum interference operator G of a repeater channel renders an effective vacuum interference operator $F_{\text{eff}} := GF$ with eigenvalue 1, then the stronger statement reduces to the weaker one. An example of such a pair F and G is given in the following.

Consider a qubit amplitude damping channel $\mathcal{A}$ whose vacuum extension $\widetilde{\mathcal{A}}$ has Kraus operators $\{(|0\rangle\langle 0|+\sqrt{1-\gamma}|1\rangle\langle 1|)\oplus\alpha_0 |\text{vac}\rangle\langle\text{vac}|, \sqrt{\gamma}|0\rangle\langle 1|\oplus\alpha_1 |\text{vac}\rangle\langle\text{vac}|\}$. For complete damping $\gamma = 1$, we recover a qubit erasure channel with vacuum interference operator $F = |0\rangle\langle\alpha|$, where $|\alpha\rangle := \alpha_0|0\rangle + \alpha_1|1\rangle$. F does not have eigenvalue 1 (unless $\alpha_0 = 1$) but does have singular value 1, since $||F^\dagger F||_\infty = 1$. Now consider a repeater channel $\mathcal{R}$ whose vacuum extension $\widetilde{\mathcal{R}}$ has Kraus operators $\{|\alpha\rangle\langle 0| \oplus |\text{vac}\rangle\langle\text{vac}|, |\alpha^\perp\rangle\langle 1| \oplus 0\}$, where $|\alpha^\perp\rangle := \bar{\alpha}_1|0\rangle - \bar{\alpha}_0|1\rangle$ is a state orthogonal to $|\alpha\rangle$. Observe that $\mathcal{R}(|0\rangle\langle 0|) = |\alpha\rangle\langle\alpha|$ and that the vacuum interference operator G of $\widetilde{\mathcal{R}}$ satisfies $G = |\alpha\rangle\langle 0| = F^\dagger$. Importantly, $GF = |\alpha\rangle\langle\alpha|$ has eigenvalue 1. Furthermore, for all n,

$$\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[(\widetilde{\mathcal{R}}\widetilde{\mathcal{E}})^n, (\widetilde{\mathcal{R}}\widetilde{\mathcal{E}})^n](\rho) = |0\rangle\langle 0| \otimes \left[q_{\rho,\alpha}|+\rangle\langle+| + (1-q_{\rho,\alpha})\frac{I}{2}\right], \tag{8.11}$$

where $q_{\rho,\alpha} = \langle\alpha|\rho|\alpha\rangle$. Since this output state has the same form as Eq. (8.7), we conclude that the channel $(\widetilde{\mathcal{R}}\widetilde{\mathcal{E}})^n$ has non-zero classical capacity for all n.

Motivated by the above discussion, we establish the following theorem for asymptotically long sequences of identical quantum channels $\mathcal{E}$.

Theorem 21. *Let $\mathcal{E}$: $\mathsf{Chan}(X \to Y)$ be a quantum channel. Let $\widetilde{\mathcal{E}}$ be a vacuum extension of $\mathcal{E}$, with vacuum interference operator F and Kraus representation $\{\widetilde{E}_i := E_i \oplus \alpha_i |\text{vac}\rangle\langle\text{vac}|\}_{i=0}^{r-1}$, where $\{E_i\}_{i=0}^{r-1}$ is a Kraus representation of $\mathcal{E}$, and $\{\alpha_i\}_{i=0}^{r-1}$ are vacuum amplitudes. Assume that for every repeater chan-*

nel $\mathcal{Q} : \mathsf{Chan}(Y \to X)$*, the classical capacity of the concatenated channel* $\mathcal{E} \circ (\mathcal{Q} \circ \mathcal{E})^{n-1}$ *tends to zero as* $n \to \infty$*. Then, the following are equivalent:*

1. *There exists a repeater channel* $\mathcal{R} : \mathsf{Chan}(Y \to X)$ *with vacuum extension* $\widetilde{\mathcal{R}}$ *and vacuum interference operator G, such that the superposition of two independent identical sequences of channels* $\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}, \widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}] : \mathsf{Chan}(X \to Y \otimes C)$ *has classical capacity strictly greater than zero as* $n \to \infty$*.*

2. *The vacuum interference operator F has singular value 1.*

Moreover, the above conditions hold only if

3. *There exist two pure states* $|\varphi\rangle \in \mathcal{H}_X$ *and* $|\zeta\rangle \in \mathcal{H}_Y$ *such that* $\mathcal{E}(|\varphi\rangle\langle\varphi|) = |\zeta\rangle\langle\zeta|$*, and there exists* $\theta \in \mathbb{R}$ *such that* $\alpha_i = e^{i\theta}\sqrt{\langle\varphi|E_i^\dagger E_i|\varphi\rangle}$ *for all* $i = 0, 1, \ldots, r-1$*.*

The proof is given in Appendix A.6.

8.6 Relaxation of assumptions and practical communication advantages

In the above, we have provided a general characterisation of the quantum channels which allow communication at a finite rate when used in asymptotically long paths traversed in a coherent superposition. However, the above results relied on three crucial assumptions: (a) the path degree of freedom remains completely noiseless throughout the whole sequence, (b) there is no loss of particles, and (c) we are only interested in *asymptotically* long sequences of channels. However, in practice, these assumptions are only justified as idealisations of more complex scenarios; realistic transmission lines are not lossless and do not in general retain perfect coherence on the path degree of freedom, nor will they always have a vacuum interference operator with singular value of exactly 1.

In the following, we examine what happens when the idealisation (c) relaxed, whilst the relaxation of idealisations (a)–(b) are examined in Chapter 9. We perform numerical simulations of various superpositions of channels to calculate lower bounds to their classical capacity. We do this by computing the maximum Holevo capacity of the superposition of channels over all input ensembles consisting of two orthogonal states (cf. §6.2.4). This gives a lower bound to the Holevo capacity, which in turn gives a lower bound to the classical capacity [104]. We compare this with upper and lower bounds of the classical capacity of the same sequence of channels without superposition. The lower bounds are found in the same way as for the superposition case. The upper bounds are found from the analytically known capacity of the quantum erasure channel with erasure state orthogonal to the input Hilbert space [31], given by $(1-p)^n$, which always gives an upper bound to the capacity of the corresponding erasure channel of the form of Eq. (8.1).

Non-unit singular value F

Here we relax the initial requirement of obtaining finite-capacity communication in the *asymptotic limit* of infinitely many channels. We instead turn our attention to using a superposition of paths to transmit information through a *finite* number of channels.

For illustration, we again consider an erasure channel $\mathcal{E}$ of the form (8.1), but this time consider a general form of possible vacuum extensions (without assuming a specific implementation as above). Such a general vacuum extension $\widetilde{\mathcal{E}}$ has Kraus operators

$$\begin{aligned} \widetilde{E}_i &= \sqrt{p}|\eta\rangle\langle i| \oplus \sqrt{p}\,\langle i|\alpha\rangle\, I_{\mathrm{Vac}} \quad \text{for } i = 0, \ldots, r-1\,, \\ \widetilde{E}_r &= \sqrt{1-p}I_{\widetilde{M}}\,, \end{aligned} \tag{8.12}$$

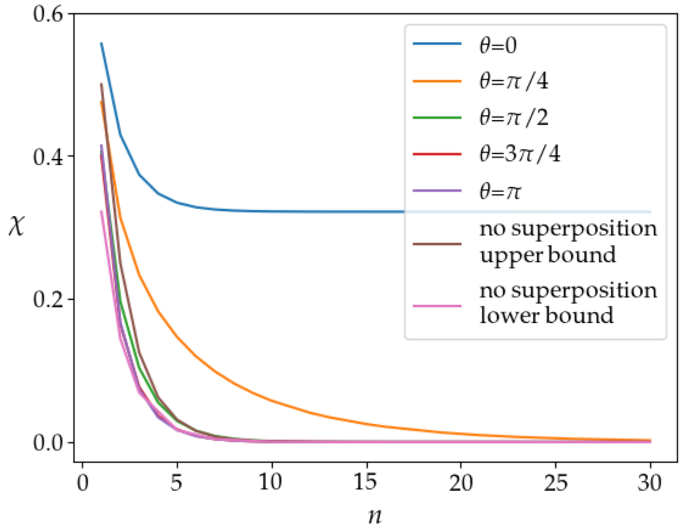
where $|\alpha\rangle := \sum_{i=0}^{r-1} \alpha_i\,|i\rangle$ (note that the global phase of $|\alpha\rangle$ is a physical parameter here), and has vacuum interference operator $F = p\,|\eta\rangle\langle\alpha| + (1-p)I_M$. Therefore, we have that

$$F^n = f(n,\mathcal{E})\,|\eta\rangle\langle\alpha| + (1-p)^n I\,, \tag{8.13}$$

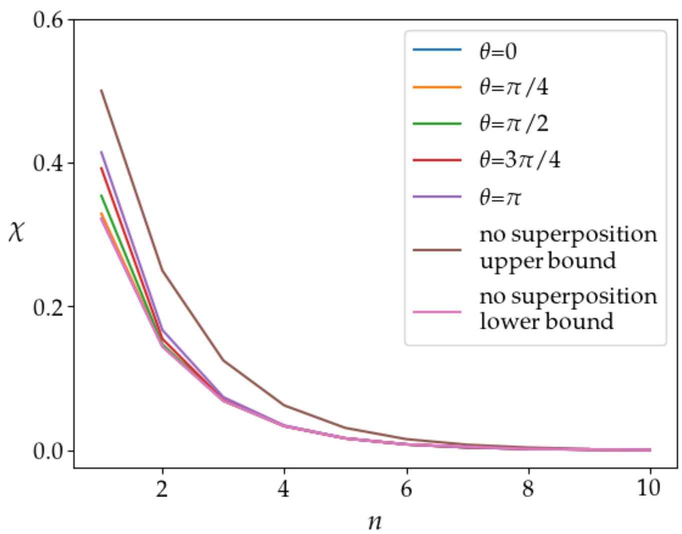
where $f(n,\mathcal{E}) := \sum_{k=0}^{n-1} \begin{pmatrix} n \\ k \end{pmatrix} p^{n-k}(1-p)^k\,\langle\alpha|\eta\rangle^{n-k-1}$ is the *vacuum decay function* that determines the amount of vacuum interference left after n channels.

We investigate the use of qubit erasure channels with non-unit singular values of F numerically by varying the parameters p, α_0, α_1. These are re-parametrised as $\alpha_0 = e^{i\varphi_0}\cos(\theta/2)$, $\alpha_1 = e^{i\varphi_1}\sin(\theta/2)$, $0 \leq \theta \leq \pi$, $0 \leq \varphi_0 \leq 2\pi$, $0 \leq \varphi_1 \leq 2\pi$. When $|\alpha\rangle = |\eta\rangle$ (including the global phase), then F has a singular value of 1. Without loss of generality, we set $|\eta\rangle = |0\rangle$.

Figure 8.4 shows the (lower bound to) the classical capacity as a function of sequence length n for the parameters chosen as $p = 0.5$, $\varphi_1 = 0$, for various values of θ. A stark contrast is seen between (a) $\varphi_0 = 0$ (in which case F has a singular value of 1 when $\theta = 0$) and (b) $\varphi_0 = \pi$ (in which case F never has a singular value of 1). In the latter case, no significant advantage is observed, which is due to the terms in the sum in $f(n, \mathcal{E})$ cancelling each other out as the sign of $\langle \alpha|\eta\rangle^{n-k-1}$ alternates. In the former case, $\langle \alpha|\eta\rangle^{n-k-1}$ is always positive, and the superposition of paths provides a non-trivial increase in the coherence terms $\pm\gamma(F^n \rho F^{\dagger^n})$. With no superposition of paths, the upper bound of the classical capacity goes below 0.01 for $n \geq 7$, whilst the lower bound of the classical capacity stays above 0.01 until $n = 21$ with a superposition of paths, for $\theta \lesssim \pi/4$.



(a)



(b)

Figure 8.4: *Lower bounds to the classical capacity against sequence length n for a superposition of two identical sequences of qubit erasure channels, with erasure probability $p = 0.5$, for (a) $\varphi_0 = 0, \varphi_1 = 0$ and (b) $\varphi_0 = \pi, \varphi_1 = 0$, with various values of θ.* Both figures also show a comparison with using the same sequence of n qubit erasure channels without a superposition of paths. It can be seen that the parameters in (a) enable a significant enhancement of the classical capacity for $\theta \lesssim \pi/4$, whilst in (b) no value of θ provides a significant advantage over no superposition.

9

Practical considerations and experimental implementation

In this chapter, we discuss some of the additional considerations that need to be taken into account when a superposition of channels is implemented in practice. First, we discuss the possibility of dephasing noise on the path degree of freedom, both for independent and correlated channels. Then, we discuss the possibility of particle loss on each path. Finally, we present some key features of the experimental design in the proof-of-principle experimental collaborations of Refs. [166] and [173].

9.1 Dephasing on the path

9.1.1 Independent channels

Here, we relax the assumption that the path is completely noiseless in the superposition of channels. In particular, we consider a dephasing error on the path. That is, between each pair of nodes, the path qubit undergoes the dephasing channel

$$\mathcal{P}(\omega) = sZ\omega Z + (1-s)\omega\,, \tag{9.1}$$

where $s \in\ ]0, 1/2[$ is a probability and $\omega \in \mathsf{St}(C)$ is the initial state of the path. This is equivalent to applying a dephasing channel between the one-particle sector B and the vacuum sector Vac on one of the two paths: $\widetilde{\mathcal{E}} \to \mathcal{Q}_s \circ \widetilde{\mathcal{E}} = \widetilde{\mathcal{E}} \circ \mathcal{Q}_s$, where

$$\begin{aligned} &\mathcal{Q}_s \in \mathsf{Chan}(B \oplus \mathrm{Vac}) \\ &\mathcal{Q}_s = s(I \oplus -|\mathrm{vac}\rangle\langle\mathrm{vac}|)\rho(I \oplus -|\mathrm{vac}\rangle\langle\mathrm{vac}|) + (1-s)\rho\,, \end{aligned} \tag{9.2}$$

with no dephasing on the other path.

Accordingly, we define a placement supermap called the *dephased superposition placement* by

$$\mathcal{S}^{\omega}_{\text{deph sup}}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}; \mathcal{Q}_s)(\rho) := \mathcal{U}^{\dagger} \circ [\widetilde{\mathcal{A}} \otimes (\mathcal{Q}_s \circ \widetilde{\mathcal{B}})] \circ \mathcal{U}\,(\rho \otimes \omega)\,, \tag{9.3}$$

where $\mathcal{U}$ is defined by Eq. (3.16). (Note, that $\mathcal{Q}_s$ commutes with any vacuum extended channel $\widetilde{\mathcal{B}}$, because any vacuum extended channel is block diagonal with respect to the one-particle/vacuum sector partition.) The evolution experienced by a single particle under the dephased superposition of channels is described by the effective channel

$$\mathcal{C}'_{\omega} = (\mathcal{I}_M \otimes \mathcal{P}_C)\mathcal{C}_{\omega}\,, \tag{9.4}$$

where $\mathcal{C}_{\omega}$ is the output of the corresponding standard superposition placement.

Now consider the case where the message travels in a superposition of two independent identical channels $\mathcal{A}$, with vacuum extensions $\widetilde{\mathcal{A}}$ and the control qubit is initialised in the $|+\rangle$ state. A direct calculation (e.g. by substituting the Kraus operators $\{A_i \oplus \alpha_i\,|\mathrm{vac}\rangle\langle\mathrm{vac}|\}_{i=0}^{r-1}$ of $\widetilde{\mathcal{A}}$ into Eq. (9.3)), reveals that

$$\begin{aligned} \mathcal{S}^{|+\rangle\langle+|}_{\text{deph sup}}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{A}}; \mathcal{Q}_s)(\rho) = &\frac{\mathcal{A}(\rho) + (1-2s)F\rho F^{\dagger}}{2} \otimes |+\rangle\langle+| \\ &+ \frac{\mathcal{A}(\rho) - (1-2s)F\rho F^{\dagger}}{2} \otimes |-\rangle\langle-|\,, \end{aligned} \tag{9.5}$$

where $F := \sum_i \overline{\alpha}_i A_i$.

For a network of channels as considered in Chapter 8, we apply Eq. (9.3) n times, to obtain the dephased superposition of channels

$$\mathcal{S}^{|+\rangle\langle+|}_{\text{deph sup}}(\widetilde{\mathcal{A}}^n, \widetilde{\mathcal{A}}^n; \mathcal{Q}^n_s)(\rho) = \frac{\mathcal{A}^n(\rho) + \gamma F^n \rho F^{\dagger^n}}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{A}^n(\rho) - \gamma F^n \rho F^{\dagger^n}}{2} \otimes |-\rangle\langle-| \,, \tag{9.6}$$

where $\gamma = (1 - 2s)^n$.

This shows that the magnitude of the coherence terms $\pm\gamma(F^n \rho F^{\dagger^n})$ decreases exponentially with sequence length n. However, for finite n and small enough dephasing probability s, communication is still possible – in stark contrast to using the channels individually without a superposition of paths.

To illustrate the effect of dephasing, Figure 9.1 shows a numerical plot of (lower bounds to) the classical capacity against sequence length n for a dephased superposition of two identical sequences of qubit erasure channels $\widetilde{\mathcal{E}}$ with error probability $p = 0.5$, for various dephasing parameters s on the path. Note that $s = 0$ corresponds to no dephasing, while $s = 0.5$ corresponds to complete dephasing on the path so that the final state of the path is simply the maximally mixed state; this is equivalent to using a single sequence of n channels without a superposition of paths. We see that for every value of n, the lower bound to the capacity monotonically increases as s decreases from 0.5 to 0, and an observable advantage over no superposition of paths is still present for low n when $s \lesssim 0.1$.

9.1.2 Correlated channels

For completeness, we briefly consider the effects of dephasing on the correlated superposition of channels presented in Chapter 6. For simplicity, here we focus on the communication scenario involving a single transmission line, specifically, the superposition of times of a correlated random unitary channel, with vacuum extension $\widetilde{\mathcal{R}}$, given by Eq. (6.4). By substituting Eq. (6.4) into Eq. (9.4), it is immediate to see that the effect of dephasing is to dampen the interference term $\mathcal{G}$ in the effective channel (6.7): specifically, the interference term changes from $\mathcal{G}$ to $(1 - 2s)\,\mathcal{G}$, and the

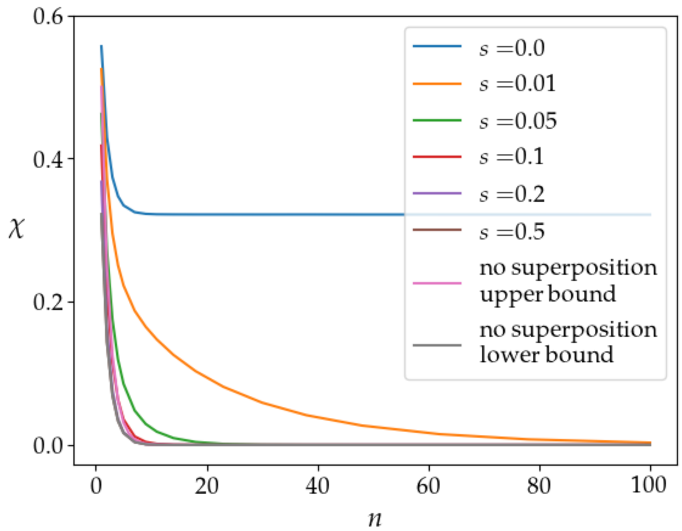


Figure 9.1: *Lower bounds to the classical capacity against sequence length n for a dephased superposition of two identical sequences of qubit erasure channels, with erasure probability $p = 0.5$, for dephasing on the path with parameter s.* The numerical values are calculated as described in §8.6.

effective channel, with the control initialised in the $|+\rangle$ state, becomes

$$\begin{aligned}\mathcal{C}_{|+\rangle\langle+|}(\rho) &= \mathcal{S}^{|+\rangle\langle+|}_{\text{deph sup}}(\widetilde{\mathcal{R}})(\rho) \\ &= \frac{\mathcal{R}_1(\rho) + (1-2s)\mathcal{G}(\rho)}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{R}_1(\rho) - (1-2s)\mathcal{G}(\rho)}{2} \otimes |-\rangle\langle-| \,.\end{aligned} \tag{9.7}$$

This is the same as the damping by factor $(1-2s)$ for the case of independent channels (9.6).

In the case of completely depolarising channels on the message degree of freedom, the presence of a non-zero interference term means that, as long as the dephasing of the control is not complete ($s \neq 1/2$), the superposition of evolutions can still allow for a non-zero amount of classical information to be transmitted, thereby offering an advantage over the transmission at a definite moment of time.

Figure 9.2 shows the behaviour of the classical capacity as a function of the dephasing parameter s. The figure shows that correlations between two uses of the channel offer an enhancement of the classical capacity. To make this point, we first evaluate numerically the maximum capacity achievable in the lack of correlations, with arbitrary vacuum extensions of the completely depolarising channel (blue curve). Notably, the maximum capacity for every fixed value of s is achieved by the

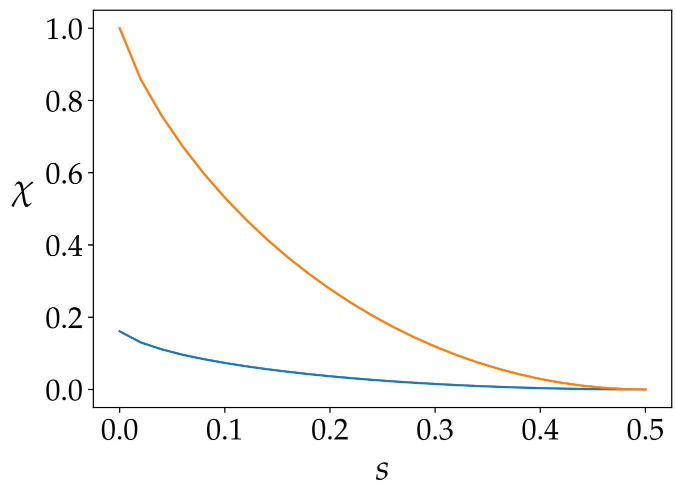


Figure 9.2: *Blue: Maximum classical capacity in the absence of correlations, as a function of the dephasing parameter s.* The maximum is computed over all possible vacuum extensions of the completely depolarising channel, and is achieved by the random unitary vacuum extension with the choice of phases $\{\varphi_m\}$ that give the maximum capacity of 0.16 bits when $s = 0$. *Orange: Lower bound to the maximal classical capacity in the presence of correlations, as a function of the dephasing parameter s.* The lower bound is computed by considering the correlated probability distribution $p(m, n) = \delta_{n,\sigma(m)}/4$, where σ is the permutation that exchanges 0 with 1, and 2 with 3, and $\varphi_0 = \varphi_1 = \varphi_2 = 0, \varphi_3 = \pi/2$.

same vacuum extension of the completely depolarising channel that achieves the maximum capacity in the ideal $s = 0$ case. We then show that a higher capacity can be achieved with the dephased version of the superposition of times of the correlated channel described in Eqs. (6.11) and (6.14). To this purpose, we numerically evaluate a lower bound to the Holevo capacity (and therefore to classical capacity), obtained by restricting the maximisation to orthogonal input ensembles (orange curve). Note that both the blue and orange curves are above 0 for every non-maximal amount of dephasing ($s \neq 1/2$), meaning that the single particle transmission at a coherent superposition of times always offers an advantage over the transmission at a definite time.

9.2 Loss of particles

In this section, we briefly discuss the performance of the superposition of channels in the presence of particle loss. In realistic communication channels, the probability of particle loss increases exponentially with distance. This means that even if the amount of loss is negligible in individual channels, such as in the protocols considered in Chapters 5–6, loss will inevitably play a role in realistic implementations of protocols involving long sequences of channels, such as those considered in Chapter 8.

We shall model loss of particles in the following way. Consider a particle propagating through a channel $\mathcal{A} \in \mathsf{Chan}(A)$, for some internal degree of freedom A, with vacuum extension $\widetilde{\mathcal{A}} \in \mathsf{Chan}(\widetilde{A})$. In the presence of loss with probability p, the evolution of the particle is given by

$$\widetilde{\mathcal{A}}' := \mathcal{L}_p \circ \widetilde{\mathcal{A}} = (1-p)\widetilde{\mathcal{A}} + p\mathcal{L}_1\,, \tag{9.8}$$

where $\mathcal{L}_p$ is the loss channel given by

$$\begin{aligned} &\mathcal{L}_p \in \mathsf{Chan}(A \oplus \mathrm{Vac}) \\ &\mathcal{L}_p(\rho) := (1-p)\rho + p\,\mathrm{Tr}(\rho)\,|\mathrm{vac}\rangle\langle\mathrm{vac}|\ . \end{aligned} \tag{9.9}$$

In order to describe the use of two such channels in a superposition, we need to update the definition of the isomorphism $\mathcal{U}$ in Eq. (3.16) between the composite message-path system $M \otimes P$ and the one-particle sector $(A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes \mathrm{B})$. With the added possibility of loss of particles, we are no longer restricted to the one-particle sector. Instead, the output of the superposition of channels is constrained to lie in the direct sum of the one-particle sector $(A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes \mathrm{B})$

and the no-particle sector (Vac $\otimes$ Vac). As such, we define a new isomorphism

$$U' : (\mathcal{H}_M \otimes \mathcal{H}_P) \oplus \mathcal{H}_{\text{fail}} \to (\mathcal{H}_A \otimes \mathcal{H}_{\text{Vac}}) \oplus (\mathcal{H}_{\text{Vac}} \otimes \mathcal{H}_B) \oplus (\mathcal{H}_{\text{Vac}} \otimes \mathcal{H}_{\text{Vac}}) \tag{9.10}$$

defined by

$$\begin{aligned} U'(|\psi\rangle_M \otimes |0\rangle_P) &:= |\psi\rangle_A \otimes |\text{vac}\rangle_B \\ U'(|\psi\rangle_M \otimes |1\rangle_P) &:= |\text{vac}\rangle_A \otimes |\psi\rangle_B \\ U'(|\text{fail}\rangle) &:= |\text{vac}\rangle_A \otimes |\text{vac}\rangle_B\,, \end{aligned} \tag{9.11}$$

where $\mathcal{H}_{\text{fail}}$ is a one-dimensional Hilbert space spanned by the unique failure state $|\text{fail}\rangle$, corresponding to the failure of having any particle at all.

With this, we define the placement supermap called the *lossy superposition placement* by

$$\mathcal{S}^{\omega}_{\text{loss sup}}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}; \mathcal{L}^A_p, \mathcal{L}^B_p)(\rho) := (\mathcal{U}')^\dagger \circ (\mathcal{L}^A_p \circ \widetilde{\mathcal{A}}) \otimes (\mathcal{L}^A_p \circ \widetilde{\mathcal{B}}) \circ \mathcal{U}\,(\rho \otimes \omega)\,, \tag{9.12}$$

where both $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are vacuum extensions of some channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$, and $\mathcal{L}^A_p, \mathcal{L}^B_p$ are some loss channels of the form (9.9).

For two independent identical vacuum-extended channels $\widetilde{\mathcal{A}}$, with independent and identical loss processes on both paths, and with the path initialised in the $|+\rangle$ state, the lossy superposition of channels gives

$$\begin{aligned} \mathcal{S}^{|+\rangle\langle+|}_{\text{loss sup}}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{A}}; \mathcal{L}_p, \mathcal{L}_p)(\rho) = (1-p)\left[\frac{\mathcal{A}(\rho) + F\rho F^\dagger}{2} \otimes |+\rangle\langle+| + \frac{\mathcal{A}(\rho) - F\rho F^\dagger}{2} \otimes |-\rangle\langle-|\right] \\ + p\,|\text{fail}\rangle\langle\text{fail}| + \text{off-diagonal terms}\,, \end{aligned} \tag{9.13}$$

where 'off-diagonal terms' refer to terms proportional to $|\text{fail}\rangle\langle\pm|$ or its Hermitian conjugate. The off-diagonal terms are not important, because the receiver will typically perform a non-demolition measurement in the particle-number basis, before proceeding to decode the message.

This result can be generalised to superpositions of sequences of multiple independent channels, and to superpositions of correlated channels, in a natural way.

This shows that for a probability of loss p at each port, the lossy superposition of channels returns the corresponding superposition of channels without loss with probability $(1-p)$ and no particle with probability p. That is, the superposition of channels is affected by loss exactly as much as the original channels themselves. This means that although (a) the superposition of channels itself cannot be used as a protocol to circumvent loss, (b) the superposition of channels still provides an advantage over using the channels in a classical configuration for every fixed value of loss probability p. In particular, it provides the same advantage as without loss, but this can now only be obtained by postselection on the one-particle sector with probability p.

9.3 Experimental design

In this section, we provide a brief overview of the methods used to experimentally implement the superposition of channels and its associated communication advantages presented in earlier chapters. The following is based on the experimental collaboration with the University of Vienna [166] and the experimental collaboration with Imperial College London [173].

Since this is a theoretical work, we shall only discuss the conceptual elements of the experimental implementation. The details of the experimental components and methods, as well as the detailed results including discussions of experimental errors, are out of the scope of this work, and can be found in the above references. We only consider implementations based on the superposition of paths, as the superposition of times is more challenging to implement in experiment. The superposition of times is an important task for future experimental work on this topic. All of the results on time-correlated channels presented in Chapter 6 can also be formulated in terms of spatially correlated channels (since the correlated channels we considered were all no-signalling channels), which is indeed how those results have been experimentally implemented in the first instance.

Ref. [166] performed an implementation of the superposition of (various combinations of) two independent random unitary channels. Each channel consists of a randomisation over (a subset of) the four Pauli unitaries, with the identity occurring with probability p. The figure of merit for communication was chosen to be the coherent information of the channel, with respect to inputting one half of the maximally entangled state. This is a lower bound to the coherent information of the channel, which in turn is a lower bound to the quantum capacity of the channel (cf. §6.4).

By combining the two channels in a superposition of paths, the theoretical value of the coherent information with respect to the maximally entangled state is larger than that of each of the constituent channels individually [166]. These communication advantages are similar to that of

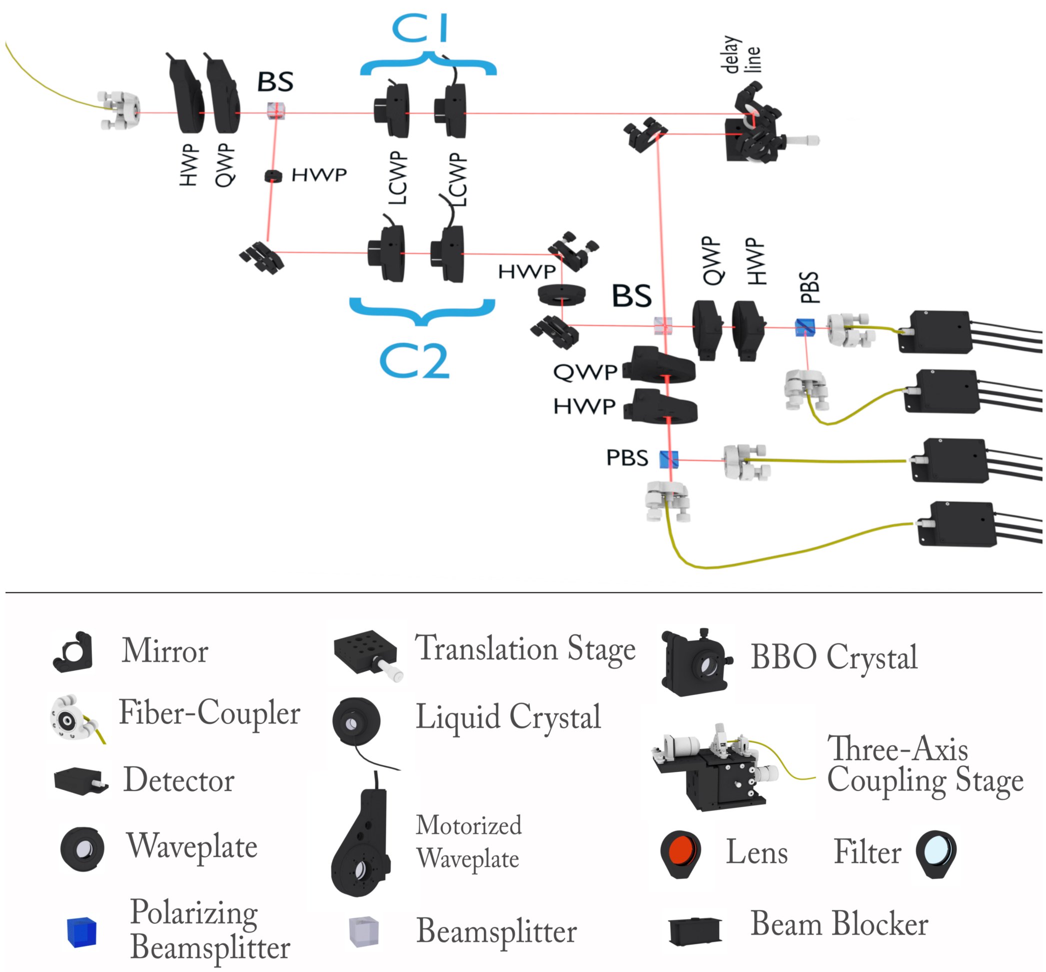


Figure 9.3: *Diagram of the experimental setup of the collaboration with the University of Vienna.* Figure modified from Ref. [166]. The noisy channels are labelled C1 and C2 in the diagram. HWP is a half wave plate, QWP is a quarter wave plate, LCWP is a liquid-crystal wave plates; all three are optical devices that rotate the polarisation. BS denotes a beamsplitter and PBS denotes a polarising beamsplitter.

the communication protocols presented in §5.3 and §6.4.1. The experiment verified the theoretical results, up to experimental errors. However, each unitary operator V_m was simply vacuum-extended to another unitary with the same phase (of 0) between the one-particle and vacuum sectors: $\widetilde{V}_m = V_m \oplus |\text{vac}\rangle\langle\text{vac}|$. This means that the whole space of possible vacuum-extensions was not explored, and as a result, the optimal communication capacity could not be retrieved.

The experimental setup of Ref. [166] is based on a simple Mach-Zehnder interferometer, as described earlier in §3.4.1. Single photons were generated using a type-II spontaneous parametric down-conversion source. The noisy channels were realised by performing the appropriate randomisation over the Pauli unitaries, which were individually implemented using liquid-crystal wave plates (LCWP). LCWPs can be tuned to the appropriate unitary within 100 ms, meaning that the randomisation can be done on-the-fly, without averaging during data analysis. A diagram of the full experimental setup is given in Figure 9.3. The coherent information of the superposition channel, with respect to the maximally entangled state, was found by performing quantum process tomography [159] on the output of the interferometer.

Ref. [173] performed an implementation of the superposition of two random unitary channels, both independent and correlated, corresponding exactly to the communication protocols presented in §6.1 and §6.4. Crucially, all possible unitary vacuum extensions $\widetilde{V}_m = V_m \oplus e^{i\varphi_m} |\text{vac}\rangle\langle\text{vac}|$ of each unitary V_m were explored. This enabled the communication enhancements for both classical and quantum communication to be formally quantified as a function of the choice of the vacuum amplitudes $\{e^{i\varphi_m}\}$.

The experimental setup of Ref. [173] is based on a displaced Sagnac interferometer, as shown in Figure 9.4. This is structurally analogous to a Mach-Zehnder interferometer but was chosen because it provides better phase stability. Single photons were again generated by a type-II spontaneous parametric down-conversion source, and sent to the input of the interferometer. The Holevo capacity and coherent information, with respect to the maximally entangled state, can be calculated using quantum process tomography on the output of the interferometer.

The noisy channels were realised by performing the appropriate randomisation over the Pauli unitaries, this time individually implemented using combinations of quarter and half wave plates. This results in unitaries with a fixed phase with the vacuum, of the form $\widetilde{V}'_m = V_m \oplus e^{i\xi_m} |\text{vac}\rangle\langle\text{vac}|$. In order to scan over all possible phases, an adjustable phase-shifter was inserted on one of the arms of the interferometer. If the phase-shifter is described by a unitary $U = I \oplus e^{i\zeta} |\text{vac}\rangle\langle\text{vac}|$, then its concatenation with the vacuum-extended unitary $\widetilde{V}'_m$ on the same arm is described by the overall

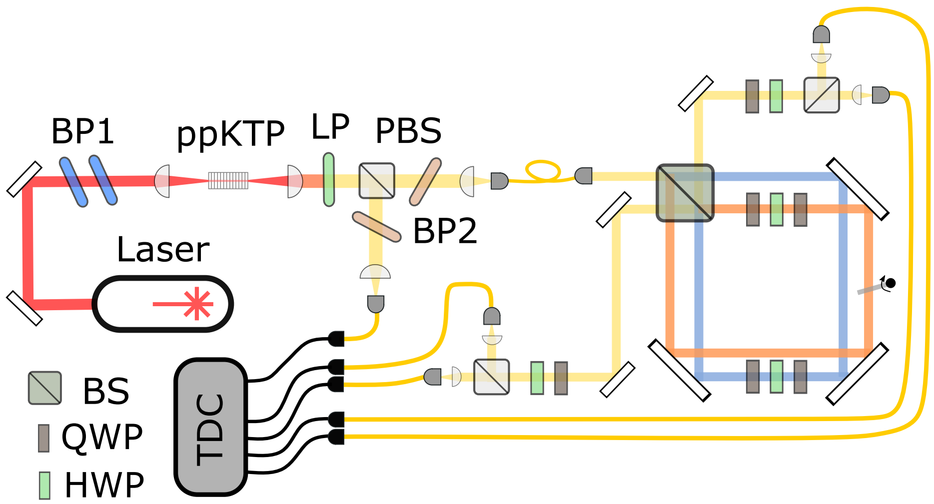


Figure 9.4: *Schematic diagram of the experimental setup of the collaboration with Imperial College London.* Figure taken from Ref. [173]. The acronyms for the various experimental components are as follows: BP1 and BP2 are band-pass filters, LP is a longpass filter, ppKTP is a periodically poled potassium titanyl phosphate, TDC is a time-to-digital converter, and the other acronyms are as in the previous figure. The black pin-shaped object on the orange arm is a small piece of glass mounted on a motorised rotation stage, which implements an adjustable phase shifter; this is used to scan over all possible phases between the one-particle and vacuum sectors of each unitary.

unitary $\widetilde{V}_m = V_m \oplus e^{i\varphi_m} |\text{vac}\rangle\langle\text{vac}|$, where $\varphi_m = \xi_m + \zeta$. Scanning over all $\zeta \in [0, 2\pi[$ is equivalent to scanning over all possible phases φ_m of the original vacuum amplitude $e^{i\varphi_m}$.

Moreover, note that the superposition of two unitaries $\widetilde{V}_m = V_m \oplus e^{i\varphi_m} |\text{vac}\rangle\langle\text{vac}|$ and $\widetilde{V}_n = V_n \oplus e^{i\varphi_n} |\text{vac}\rangle\langle\text{vac}|$ has, by Eq. (3.17), the form

$$\begin{aligned} \mathcal{S}^{\omega}_{\text{sup}}(\widetilde{\mathcal{V}}_m, \widetilde{\mathcal{V}}_n)(\rho) =& \langle 0|\omega|0\rangle\, \mathcal{V}_m(\rho) \otimes |0\rangle\langle 0| + \langle 1|\omega|1\rangle\, \mathcal{V}_n(\rho) \otimes |1\rangle\langle 1| \\ &+ \langle 0|\omega|1\rangle\, e^{i(\varphi_n - \varphi_m)} V_m \rho V_n^{\dagger} \otimes |0\rangle\langle 1| + \langle 1|\omega|0\rangle\, e^{i(\varphi_m - \varphi_n)} V_n \rho V_m^{\dagger} \otimes |1\rangle\langle 0| \,, \end{aligned} \tag{9.14}$$

which depends only on the phase differences $\varphi_m - \varphi_n$. Since in this experiment each run consists of a fixed choice of unitaries (with the randomisation performed in post-processing), the above equation implies that a phase-shifter only needs to be used on one of the two arms in order to scan over all possible vacuum extensions of both channels.

The data analysis for this experiment is currently in progress.

10 Discussions and conclusions

In this section, we summarise the main results of the thesis and discuss the connections with related works, in particular, highlighting the use of the methods we have developed in other fields. We comment on possible research directions for future work in the second-quantised Shannon theory itself, as well as in other areas including quantum computation, foundations, gravity and causality.

10.1 Summary

In this thesis, we formalised a second level of quantisation of Shannon's theory of information, where both the information carriers and their trajectories in spacetime are treated in a quantum manner. Using the tools of higher-order transformations, we defined the notion of a superposition of two quantum channels, constructed by coherently routing a single particle in a superposition of travelling through one channel or the other. We constructed a resource theory of communication, which formalises the resources available to the different parties in a communication scenario, and specified a minimal requirement that all meaningful communication models should satisfy. We cast both standard quantum Shannon theory and its second level of quantisation in this form, as well as a related paradigm where communication channels can be combined in an indefinite causal order.

We provided several examples of the novel possibilities for communication achievable in this extended setting, including enhancements in both the classical and quantum capacity of communication channels, with both independent and correlated noise. We extended the framework to examples of communication networks, where further advantages can be found in the asymptotic limit of an infinite sequence of channels. Finally, we studied the robustness of communication in a superposition of trajectories to various sources of errors, and outlined the design of two recent experiments which demonstrated our results in practice.

10.2 Outlook for the second-quantised Shannon theory

We envisage that our framework for a second-quantised Shannon theory opens up a new paradigm for the study of communication in a quantum setting. Various directions for further research are apparent.

First, it would be interesting to explore the possibilities arising when *multiple* particles are simultaneously transmitted on a superposition of (the same set of) paths. This would lead to multi-particle interference between the different paths, as described by the Hong-Ou-Mandel effect [106], in addition to the interference caused by a single particle on multiple paths. The study of such a physical scenario would be important in applying the noise-reduction techniques of the second-quantised Shannon theory to practical applications in photonic quantum computation, where multiple interacting single photons are typically used [37].

Second, the extent to which the results of Chapter 8 could be applied to *quantum* communication across long sequences of channels would be an interesting avenue of study. In the case of quantum communication, unlike for classical communication, it is not sufficient for the vacuum interference operator to have a singular value of 1. However, it may be that a weaker result holds, which could potentially still, to some extent, suppress the exponential degradation of quantum information over large distances.

Third, it is important to extend the tabletop proof-of-principle experiments of the novel communication advantages to practical demonstrations of communication through actual long-distance optical fibres. Additionally, the use of correlated channels at a superposition of times would be interesting task to explore in practice. In conjunction with such demonstrations, an important task would be to pinpoint the precise scenarios where communication with a superposition of trajectories can provide real-life practical advantages for communication through noisy transmission lines.

10.3 Related works and possible future directions

We begin by reviewing two related communication protocols, where the amount of information a single particle can convey is enhanced by considering the quantum properties of an external degree of freedom. This is followed by a discussion on work completed in parallel to this thesis, in part guided by the results presented in the previous chapters, on topics in quantum gravity and quantum foundations. Finally, we discuss potential applications in quantum computation.

10.3.1 Single-particle communication

In Ref. [200], Zhang, Chen and Chitambar constructed a multiple-access channel using a single quantum particle, with a Shannon-theoretic study provided in Ref. [46]. In this scheme, multiple senders transmit information to a single receiver, using only a single quantum particle between them. This is done by first transmitting a single particle through a superposition of N paths, and letting N senders act locally, one on each path. The authors showed that the multiple senders can simultaneously transmit classical information to the receiver in this setting, which is impossible with a single classical particle [46].

The formalism of these works has several similarities to that of this thesis, in particular the description of a single particle coherently routed to the inputs of multiple channels, with the absence of an input described by the vacuum. However, a fundamental difference is that the internal degrees of freedom of the particle are not considered at all. Instead, the senders encode information in the phase between the different paths, and are also allowed to destroy the particle.

The protocol of two-way communication with a single quantum particle, due to del Santo and Dakić [72, 140], shares similar features to Refs. [46, 200]. In this scheme, two parties simultaneously send information to one another, with only the transmission of a single particle overall. The authors showed that such two-way communication is possible, in stark contrast to the possibilities offered by a single classical particle. Again, the protocol relies on the particle being initialised in a superposition of locations between the two parties, with the absence of the particle being described by the vacuum.

As in Refs. [46, 200], the protocol of Refs. [72, 140] does not consider an internal degree of freedom of the particle, and in fact does not consider noise in transmission at all.

An interesting avenue for future research would be to apply the methods developed in this thesis to study the scenarios of the above works, including their formalisation in a resource-theoretic setting. In particular, it would be interesting to extend the above scenarios to cases where information is encoded in an internal degree of freedom of the particle, and where this degree of freedom undergoes a noisy evolution.

10.3.2 Quantum communication with a superposition of spacetimes

The superposition of quantum channels has so far mostly been applied to physical scenarios involving single-photon evolutions. However, the formalism can, in principle, be applied to any physical processes described by quantum channels. In Ref. [110], we considered the performance of quantum communication protocols near a black hole. Traditionally, a black hole described by classical physics had been shown to cause decoherence on quantum states in its vicinity, thus degrading the performance of any quantum communication protocol [10, 139]. Yet, a black hole is expected to be a fundamentally quantum object. In this work, we showed that by considering a black hole in a superposition of alternative states (for example, mass states), the resulting superposition of evolutions, or equivalently the superposition of spacetimes induced by the black hole, can allay the decoherence, and therefore the degradation of quantum information protocols in its vicinity. Thus, using the methods of superpositions of channels, we showed that quantum gravity can be used as a resource for communication.

10.3.3 Quantum circuits with superpositions of wires

The standard formalism of quantum circuits [5, 6, 65, 73, 145], which lies at the heart of quantum computation in both theory and practice, enables quantum channels to be combined in parallel and in sequence. Until recently, this structure of parallel and sequential composition appeared to be sufficient for describing quantum processes for all practical purposes. However, as the work of this thesis shows, the placement of devices can itself be described in a quantum manner, which goes beyond standard placements in parallel and sequence. Ultimately, a second level of quantisation of Shannon theory calls for a second level of quantisation of the circuit model: not only should the gates of the circuit describe quantum operations, but the wires between the gates should be able to describe the coherent routing of information through alternative paths.

In Ref. [191] we provide such an extended circuit model called *routed quantum circuits*. In this framework, the composition of quantum systems in any combination of direct sums and tensor products can be depicted in a graphical manner, with the addition of Boolean matrices specifying appropriate constraints.

The constructions of routed quantum circuits are motivated in part by Definition 7 of a superposition of channels. In Definition 7, the isomorphism U between the message-path system $M \otimes P$ and the one particle sector $(A \otimes \mathrm{Vac}) \oplus (\mathrm{Vac} \otimes B)$, with $M \simeq A \simeq B$, is in fact an isometry onto the full Hilbert space $\widetilde{A} \otimes \widetilde{B}$ describing the independent action of the two vacuum-extended channels. Thus, even though only the one-particle sector can be populated, this cannot be seen from its representation in standard quantum circuits: the circuit in Figure 3.2 makes it look as though the two vacuum-extended channels are used in parallel.

In routed quantum circuits, the same physical scenario is depicted in Figure 10.1 (we shall stick to the case of unitary channels for simplicity; non-unitary channels can be described in an analogous way with a slightly more elaborate version of the formalism). In this framework, a system is explicitly decomposed into its sectors, depending on a physically meaningful partition. For example, the vacuum-extended system $\widetilde{A}$ is written as $A^k = A^0 \oplus A^1$, where $A^0 := \mathrm{Vac}$ and $A^1 := A$; similarly $\widetilde{B}$ is written as $B^l = B^0 \oplus B^1$. Each channel is appended with a Boolean matrix called the *route*, which describes which sectors of the incoming Hilbert spaces can be connected to which sectors of the outgoing Hilbert spaces. For example, the isomorphism U is appended with the Kroenecker delta route $\delta^{k\bar{l}}$, (with $\bar{l}$ meaning 'not l') which specifies that only the composite sectors $A^0 \otimes B^1$ or $A^1 \otimes B^0$ can be populated at once, i.e. the one-particle sector. Similarly, the fact that each unitary $\widetilde{V}$ preserves the number of particles is depicted by appending it with the Kroenecker delta route δ_k^m between its input system A^k and output system A^m. More generally, a route will be given by a Boolean matrix that can relate any combination of input sectors to any combination of output sectors.

The advantage of routed quantum circuits is that they make clear which combination of sectors in a large composite Hilbert space can be simultaneously populated, thus giving a better understanding of the compositional structure of various physical scenarios. In turn, a more detailed compositional structure could potentially result in a more efficient design of future quantum algorithms or communication protocols, for example, in the spirit the second-quantised Shannon

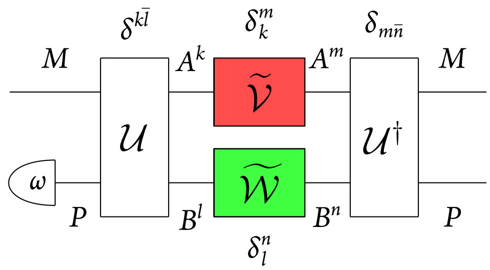


Figure 10.1: *The superposition of two independent unitary channels $\mathcal{V}$ and $\mathcal{W}$ specified by the vacuum extensions $\widetilde{\mathcal{V}}$ and $\widetilde{\mathcal{W}}$, depicted in the formalism of routed quantum circuits.* Each gate is appended with a Boolean matrix that determines which sectors of its input systems (lower indices) are consistent with which sectors of its output systems (upper indices).

theory.

Quantum causality

On the foundational side, routed quantum circuits were also inspired by the earlier work of Lorenz and Barrett [134], and Barrett, Lorenz and Oreshkov [19] on causal decompositions of unitaries. These works studied the causal structure of unitary transformations and attempted to decompose unitaries into a unitary quantum circuit that explicitly depicts their causal structure. It was found that the causal structure of some unitaries cannot be depicted as a standard circuit diagram where the compositional structure is equivalent to the causal structure. That is, it is not always possible to draw a standard circuit diagram where the connectivity of wires between two unitaries determines whether or not signalling can happen between them. Yet, it turns out that such causally faithful circuit diagrams are possible, at least for a large class of unitaries, using an extended circuit formalism [134], which is encompassed by routed quantum circuits. (In essence, the formalism of Ref. [134] can be viewed as routed quantum circuits where only Kroenecker delta routes are allowed.)

With the development of the full formalism of routed quantum circuits, which was guided by the mathematical structures encountered in the superposition of quantum channels, an interesting avenue for future research would be to see if there exist unitary transformations for which the full machinery of routed quantum circuits is necessary to depict the causal structure, i.e. where the formalism of Ref. [134] is not sufficient. An answer to this question would constitute a significant gain in our understanding of the causal structures of quantum theory.

In turn, an understanding of the causal structures of quantum theory can be expected to play an important role in understanding the ultimate potential of quantum physics for information-processing applications such as quantum machine learning [34, 194], just as current research in classical causality [153] has been instrumental in the understanding of classical machine learning algorithms [16, 164, 170, 171, 175].

Indefinite causal order

In Refs. [53, 57], Chiribella, D'Ariano, Perinotti and Valiron first defined the quantum SWITCH. They noted that the quantum SWITCH supermap cannot be represented in the standard quantum circuit formalism, yet could potentially be depicted in a more general circuit formalism where the wires are themselves movable. Of course, any process with indefinite causal order (cf. §2.4.1) can be drawn as a standard quantum circuit with the addition of a feedback loop; however, only a small subclass of processes that can be drawn with feedback loops are logically consistent. Moreover, the ability to draw processes as circuits is an important step in understanding why they are consistent. Therefore, an ideal tool for the study of indefinite causal order is a circuit formalism that enables the construction of such processes in a way that ensures, and explains why, they are logically consistent.

Routed quantum circuits provides precisely such a formalism. In Ref. [192], we applied the framework of routed quantum circuits to processes with indefinite causal order. We provided a general method to construct (unitarily extendable [14]) processes with indefinite causal order, starting from elementary graphs depicting a given causal structure. Our method defines a set of basic rules to check that any elementary graph is logically consistent, and therefore any (routed) quantum circuit formed from it. The method yields an intuitive understanding of the causal relations within a process, and why they do not lead to logical inconsistencies.

Again, a principal construction in this formalism is directly based on the work of this thesis. That is, the way in which wires can be split into combinations of direct sums and tensor products (typically, representing alternative possible causal structures) is by using the unitary $\mathcal{U}$ of Eq. (3.16) and its generalisations. A related work by Ormrod, Vanrietvelde and Barrett uses similar techniques to study the causal structure of the quantum SWITCH itself [150]. From a historical perspective, this brings us full circle: the study of communication advantages arising from indefinite causal order [58, 80, 168] was the initial inspiration to study communication with a superposition of trajectories, the methods of which have now in turn paved the way for a deeper understanding of indefinite causal order itself.

10.3.4 Quantum computation with superpositions of gates

In classical computation, a fundamental ingredient is the IF clause, which performs an operation conditional on the state of a control register. In quantum computation, however, the direct generalisation to a quantum IF clause, which coherently performs arbitrary unitary operations depending on a quantum control state, is not possible, as described earlier in §2.5.2. Yet, within the framework of superposition of channels specified by vacuum extensions, the coherent control of arbitrary operations is possible, given the additional parameters defined by the vacuum extension. This enables us to build a form of a quantum IF clause [85, 144] with the appropriate resources. Experimental realisations of the coherent control of unknown unitaries have been demonstrated in both photonic [202] and superconducting [86] systems.

A recent work by Dong, Nakayama, Soeda and Murao has extended earlier studies on the control of unknown unitaries to the coherent control of both quantum channels and quantum supermaps [76]. Another recent work by Vanrietvelde and Chiribella [190] has further developed some of the formal superposition of channels constructions in this thesis, from a more computational perspective. In particular, they formalised the coherent control of arbitrary quantum channels as a quantum supermap on routed quantum channels, in the spirit of Refs. [134, 191].

At the same time, several recent works have proposed algorithms, for example in quantum machine learning, whose advantages are purported to arise from sending information on a superposition of alternative trajectories [130, 151]. An interesting avenue of research would be to formalise these works using the methods of higher-order transformations and resource theories applied to superpositions of channels, in an analogous way for computation as was done for communication in this thesis.

We envisage that the formal methods developed in this thesis and subsequent works, enhanced by further research in this area, could lead to a new 'second-quantised' paradigm of computation, where resource theoretically well-defined advantages are proven in tasks ranging from simple algorithms to quantum machine learning.

References

[1] Alastair A Abbott, Christina Giarmatzi, Fabio Costa, and Cyril Branciard. Multipartite causal correlations: Polytopes and inequalities. *Physical Review A*, 94(3):032131, 2016.

[2] Alastair A Abbott, Julian Wechs, Dominic Horsman, Mehdi Mhalla, and Cyril Branciard. Communication through coherent control of quantum channels. *Quantum*, 4:333, 2020.

[3] Johan Åberg. Operations and single-particle interferometry. *Physical Review A*, 70(1): 012103, 2004.

[4] Johan Åberg. Subspace preservation, subspace locality, and gluing of completely positive maps. *Annals of Physics*, 313(2):326–367, 2004.

[5] Samson Abramsky and Bob Coecke. A categorical semantics of quantum protocols. In *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, 2004.*, pages 415–425. IEEE, 2004. doi: 10.1109/LICS.2004.1319636.

[6] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 20–30, 1998. doi: 10.1145/276698.276708.

[7] Yakir Aharonov and Lev Vaidman. Complete description of a quantum system at a given time. *Journal of Physics A: Mathematical and General*, 24(10):2315, 1991.

[8] Yakir Aharonov, Peter G Bergmann, and Joel L Lebowitz. Time symmetry in the quantum process of measurement. *Physical Review*, 134(6B):B1410, 1964.

[9] Yakir Aharonov, Jeeva Anandan, Sandu Popescu, and Lev Vaidman. Superpositions of time evolutions of a quantum system and a quantum time-translation machine. *Physical Review Letters*, 64(25):2965, 1990.

[10] Paul M Alsing, Ivette Fuentes-Schuller, Robert B Mann, and Tracey E Tessier. Entanglement of Dirac fields in noninertial frames. *Physical Review A*, 74(3):032326, 2006.

[11] Mateus Araújo, Fabio Costa, and Časlav Brukner. Computational advantage from quantum-controlled ordering of gates. *Physical Review Letters*, 113(25):250402, 2014.

[12] Mateus Araújo, Adrien Feix, Fabio Costa, and Časlav Brukner. Quantum circuits cannot control unknown operations. *New Journal of Physics*, 16(9):093026, 2014.

[13] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. Witnessing causal nonseparability. *New Journal of Physics*, 17(10):102001, 2015.

[14] Mateus Araújo, Adrien Feix, Miguel Navascués, and Časlav Brukner. A purification postulate for quantum mechanics with indefinite causal order. *Quantum*, 1:10, 2017.

[15] Mateus Araújo, Philippe Allard Guérin, and Ämin Baumeler. Quantum computation with indefinite causal structures. *Physical Review A*, 96(5):052315, 2017.

[16] Martin Arjovsky, Léon Bottou, Ishaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019.

[17] Jonathan Ball, Andrzej Dragan, and Konrad Banaszek. Exploiting entanglement in communication channels with correlated noise. *Physical Review A*, 69(4):042324, 2004.

[18] Konrad Banaszek, Andrzej Dragan, Wojciech Wasilewski, and Czesław Radzewicz. Experimental demonstration of entanglement-enhanced classical communication over a quantum channel with correlated noise. *Physical Review Letters*, 92(25):257901, 2004.

[19] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. Cyclic quantum causal models. *Nature Communications*, 12(1):1–15, 2021.

[20] Ämin Baumeler and Stefan Wolf. Perfect signaling among three parties violating predefined causal order. In *2014 IEEE International Symposium on Information Theory*, pages 526–530. IEEE, 2014.

[21] Ämin Baumeler and Stefan Wolf. The space of logically consistent classical processes without causal order. *New Journal of Physics*, 18(1):013036, 2016.

[22] Ämin Baumeler, Adrien Feix, and Stefan Wolf. Maximal incompatibility of locally classical behavior and global causal order in multiparty scenarios. *Physical Review A*, 90(4):042106, 2014.

[23] Jessica Bavaresco, Mateus Araújo, Časlav Brukner, and Marco Túlio Quintino. Semi-device-independent certification of indefinite causal order. *Quantum*, 3:176, 2019.

[24] Abolfazl Bayat, Daniel Burgarth, Stefano Mancini, and Sougato Bose. Memory effects in spin-chain channels for information transmission. *Physical Review A*, 77(5):050306, 2008.

[25] John S Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[26] Khaled Ben Dana, María García Díaz, Mohamed Mejatty, and Andreas Winter. Resource theory of coherence: Beyond states. *Physical Review A*, 95:062327, 2017.

[27] Jean Bénabou. Algebre homologique-categories avec multiplication. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences*, 256(9):1887, 1963.

[28] Charles H Bennett and Gilles Brassard. Quantum cryptography: public key distribution and coin tossing. In *International Conference on Computers, Systems and Signal Processing (Bangalore, India, Dec. 1984)*, pages 175–9, 1984.

[29] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69:2881–2884, 1992.

[30] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996.

[31] Charles H Bennett, David P DiVincenzo, and John A Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217, 1997.

[32] Charles H Bennett, Peter W Shor, John A Smolin, and Ashish V Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83(15):3081, 1999.

[33] Some Sankar Bhattacharya and Manik Banik. Biased non-causal game. *arXiv preprint arXiv:1509.02721*, 2015.

[34] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[35] Garrett Birkhoff. Three observations on linear algebra. *Univ. Nac. Tacuman, Rev. Ser. A*, 5: 147–151, 1946.

[36] Alessandro Bisio and Paolo Perinotti. Theoretical framework for higher-order quantum theory. *Proceedings of the Royal Society A*, 475(2225):20180706, 2019.

[37] Hector Bombin, Isaac H Kim, Daniel Litinski, Naomi Nickerson, Mihir Pant, Fernando Pastawski, Sam Roberts, and Terry Rudolph. Interleaving: Modular architectures for fault-tolerant photonic quantum computing. *arXiv preprint arXiv:2103.08612*, 2021.

[38] Sougato Bose. Quantum Communication through an Unmodulated Spin Chain. *Physical Review Letters*, 91(20):207901, 2003.

[39] Garry Bowen and Stefano Mancini. Quantum channels with a finite memory. *Physical Review A*, 69:012306, Jan 2004.

[40] Cyril Branciard, Mateus Araújo, Adrien Feix, Fabio Costa, and Časlav Brukner. The simplest causal inequalities and their violation. *New Journal of Physics*, 18(1):013008, 2015.

[41] Heinz-Peter Breuer, Elsi-Mari Laine, Jyrki Piilo, and Bassano Vacchini. Colloquium: Non-markovian dynamics in open quantum systems. *Reviews of Modern Physics*, 88:021002, 2016.

[42] H-J Briegel, Wolfgang Dür, Juan I Cirac, and Peter Zoller. Quantum repeaters: the role of imperfect local operations in quantum communication. *Physical Review Letters*, 81(26): 5932, 1998.

[43] Filippo Caruso, Vittorio Giovannetti, Cosmo Lupo, and Stefano Mancini. Quantum channels and memory effects. *Reviews of Modern Physics*, 86(4):1203, 2014.

[44] Esteban Castro-Ruiz, Flaminia Giacomini, and Časlav Brukner. Dynamics of quantum causal structures. *Physical Review X*, 8:011047, 2018.

[45] Nicolas J Cerf, Julien Clavareau, Chiara Macchiavello, and Jérémie Roland. Quantum entanglement enhances the capacity of bosonic channels with memory. *Physical Review A*, 72 (4):042330, 2005.

[46] Xinan Chen, Yujie Zhang, Eric Chitambar, Virginia Lorenz, and Andreas Winter. *Information Carried by a Single Particle in Multiple-Access Channels*. Poster at the 25th Annual Conference on Quantum Information Processing (QIP 2022), Pasadena, California, 2022.

[47] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Physical Review A*, 86(4):040301, 2012.

[48] Giulio Chiribella. Agents, subsystems, and the conservation of information. *Entropy*, 20 (5):358, 2018.

[49] Giulio Chiribella and Daniel Ebler. Optimal quantum networks and one-shot entropies. *New Journal of Physics*, 18(9):093053, 2016.

[50] Giulio Chiribella and Hlér Kristjánsson. Quantum Shannon theory with superpositions of trajectories. *Proceedings of the Royal Society A*, 475(2225):20180903, 2019.

[51] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Quantum circuit architecture. *Physical Review Letters*, 101:060401, 2008.

[52] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Transforming quantum operations: quantum supermaps. *EPL (Europhysics Letters)*, 83(3):30004, 2008.

[53] Giulio Chiribella, Giacomo Mauro D'Ariano, P Perinotti, and B Valiron. Beyond quantum computers. *arXiv preprint arXiv:0912.0195*, 2009.

[54] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009.

[55] Giulio Chiribella, Michele Dall'Arno, Giacomo Mauro D'Ariano, Chiara Macchiavello, and Paolo Perinotti. Quantum error correction with degenerate codes for correlated noise. *Physical Review A*, 83(5):052305, 2011.

[56] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Informational derivation of quantum theory. *Physical Review A*, 84(1):012311, 2011.

[57] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2):022318, 2013.

[58] Giulio Chiribella, Manik Banik, Some Sankar Bhattacharya, Tamal Guha, Mir Alimuddin, Arup Roy, Sutapa Saha, Sristy Agrawal, and Guruprasad Kar. Indefinite causal order enables perfect quantum communication with zero capacity channels. *New Journal of Physics*, 23(3): 033039, 2021.

[59] Giulio Chiribella, Matt Wilson, and HF Chau. Quantum and classical data transmission through completely depolarizing channels in a superposition of cyclic orders. *Physical Review Letters*, 127(19):190502, 2021.

[60] Eric Chitambar and Gilad Gour. Quantum resource theories. *Reviews of Modern Physics*, 91 (2):025001, 2019.

[61] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[62] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 339–354. The Royal Society, 1998.

[63] Bob Coecke. Introducing categories to the practicing physicist. In *What is category theory*, pages 45–74. Polimetrica Publishing Milan, Italy, 2006.

[64] Bob Coecke. Quantum picturalism. *Contemporary physics*, 51(1):59–83, 2010.

[65] Bob Coecke and Aleks Kissinger. *Picturing quantum processes*. Cambridge University Press, 2017.

[66] Bob Coecke and Eric Oliver Paquette. Categories for the practising physicist. In *New structures for physics*, pages 173–286. Springer, 2010.

[67] Bob Coecke, Tobias Fritz, and Robert W Spekkens. A mathematical theory of resources. *Information and Computation*, 250:59–86, 2016.

[68] Timoteo Colnaghi, Giacomo Mauro D'Ariano, Stefano Facchini, and Paolo Perinotti. Quantum computation with programmable connections between gates. *Physics Letters A*, 376 (45):2940–2943, 2012.

[69] Giacomo Mauro D'Ariano, Giulio Chiribella, and Paolo Perinotti. *Quantum theory from first principles: an informational approach*. Cambridge University Press, 2017.

[70] Antonio D'Arrigo, Giuliano Benenti, and Giuseppe Falci. Transmission of classical and quantum information through a quantum memory channel with damping. *The European Physical Journal D*, 66(6):147, 2012.

[71] Edward Davies. Information and quantum measurement. *IEEE Transactions on Information Theory*, 24(5):596–599, 1978.

[72] Flavio del Santo and Borivoje Dakić. Two-way communication with a single quantum particle. *Physical Review Letters*, 120(6):060503, 2018.

[73] David Elieser Deutsch. Quantum computational networks. *Proceedings of the Royal Society of London: A. Mathematical and Physical Sciences*, 425(1868):73–90, 1989.

[74] Igor Devetak, Aram W Harrow, and Andreas J Winter. A resource framework for quantum shannon theory. *IEEE Transactions on Information Theory*, 54(10):4587–4618, 2008.

[75] Paul Adrien Maurice Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 114(767):243–265, 1927.

[76] Qingxiuxiong Dong, Shojun Nakayama, Akihito Soeda, and Mio Murao. Controlled quantum operations and combs, and their applications to universal controllization of divisible unitary operations. *arXiv preprint arXiv:1911.01645*, 2019.

[77] John M Donohue, Megan Agnew, Jonathan Lavoie, and Kevin J Resch. Coherent ultrafast measurement of time-bin encoded photons. *Physical Review Letters*, 111(15):153602, 2013.

[78] John M Donohue, Jonathan Lavoie, and Kevin J Resch. Ultrafast time-division demultiplexing of polarization-entangled photons. *Physical Review Letters*, 113(16):163602, 2014.

[79] Fay Dowker. Causal sets as discrete spacetime. *Contemporary Physics*, 47(1):1–9, 2006.

[80] Daniel Ebler, Sina Salek, and Giulio Chiribella. Enhanced communication with the assistance of indefinite causal order. *Physical Review Letters*, 120(12):120502, 2018.

[81] Artur K Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67 (6):661, 1991.

[82] Adrien Feix, Mateus Araújo, and Časlav Brukner. Quantum superposition of the order of parties as a communication resource. *Physical Review A*, 92(5):052326, 2015.

[83] Richard P Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, 1982.

[84] Vladimir Fock. Konfigurationsraum und zweite quantelung. *Zeitschrift für Physik*, 75(9): 622–647, 1932.

[85] Nicolai Friis, Vedran Dunjko, Wolfgang Dür, and Hans J Briegel. Implementing quantum control for unknown subroutines. *Physical Review A*, 89(3):030303, 2014.

[86] Nicolai Friis, Alexey A Melnikov, Gerhard Kirchmair, and Hans J Briegel. Coherent controlization using superconducting qubits. *Scientific Reports*, 5:18036, 2015.

[87] Christopher A Fuchs. Nonorthogonal quantum states maximize classical information capacity. *Physical Review Letters*, 79(6):1162, 1997.

[88] Farzad Ghafari, Nora Tischler, Carlo Di Franco, Jayne Thompson, Mile Gu, and Geoff J Pryde. Interfering trajectories in experimental quantum-enhanced stochastic simulation. *Nature Communications*, 10(1):1–8, 2019.

[89] Vittorio Giovannetti and Rosario Fazio. Information-capacity description of spin-chain correlations. *Physical Review A*, 71(3):032314, 2005.

[90] Vittorio Giovannetti and Stefano Mancini. Bosonic memory channels. *Physical Review A*, 71(6):062304, 2005.

[91] Nicolas Gisin, Noah Linden, Serge Massar, and S Popescu. Error filtration and entanglement purification for quantum communication. *Physical Review A*, 72(1):012338, 2005.

[92] James P Gordon. "Noise at optical frequencies: Information theory," in Quantum Electronics and Coherent Light. *Proceedings of the International School of Physics "Enrico Fermi," Course XXXI, 1964*, 1964.

[93] K Goswami, Christina Giarmatzi, M Kewming, Fabio Costa, Cyril Branciard, Jacquiline Romero, and AG White. Indefinite causal order in a quantum switch. *Physical Review Letters*, 121(9):090503, 2018.

[94] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White. Increasing communication capacity via superposition of order. *Physical Review Research*, 2:033292, 2020.

[95] Philippe Allard Guérin, Adrien Feix, Mateus Araújo, and Časlav Brukner. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Physical Review Letters*, 117(10):100502, 2016.

[96] Philippe Allard Guérin, Giulia Rubino, and Časlav Brukner. Communication through quantum-controlled noise. *Physical Review A*, 99(6):062317, 2019.

[97] Yu Guo, Xiao-Min Hu, Zhi-Bo Hou, Huan Cao, Jin-Ming Cui, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, and Giulio Chiribella. Experimental transmission of quantum information using a superposition of causal orders. *Physical Review Letters*, 124(3):030502, 2020.

[98] Lucien Hardy. Quantum gravity computers: On the theory of computation with indefinite causal structure. In *Quantum reality, relativistic causality, and closing the epistemic circle*, pages 379–401. Springer, 2009.

[99] Masahito Hayashi, Hiroshi Imai, Keiji Matsumoto, Mary Beth Ruskai, and Toshiyuki Shimono. Qubit channels which require four inputs to achieve capacity: Implications for additivity conjectures. *Quantum Information & Computation*, 5(1):13–31, 2005. ISSN 1533-7146.

[100] Alexander S Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[101] Alexander S Holevo. Statistical problems in quantum physics. In *Proceedings of the Second Japan-USSR Symposium on Probability Theory*, pages 104–119. Springer, 1973.

[102] Alexander S Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998.

[103] Alexander S Holevo. Remarks on the classical capacity of quantum channel. *arXiv preprint quant-ph/0212025*, 2002.

[104] Alexander Semenovich Holevo. On the mathematical theory of quantum communication channels. *Problemy Peredachi Informatsii*, 8(1):62–71, 1972.

[105] Junya Honda and Hirosuke Yamamoto. Polar coding without alphabet extension for asymmetric models. *IEEE Transactions on Information Theory*, 59(12):7829–7838, 2013.

[106] Chong-Ki Hong, Zhe-Yu Ou, and Leonard Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59(18):2044, 1987.

[107] Michael Horodecki, Peter W Shor, and Mary Beth Ruskai. Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(06):629–641, 2003.

[108] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps. *Physics Letters A*, 283(1-2):1–7, 2001.

[109] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865, 2009.

[110] Richard Howl, Ali Akil, Hlér Kristjánsson, Xiaobin Zhao, and Giulio Chiribella. Quantum gravity as a communication resource. *arXiv preprint arXiv:2203.05861*, 2022.

[111] Peter C Humphreys, Benjamin J Metcalf, Justin B Spring, Merritt Moore, Xian-Min Jin, Marco Barbieri, W Steven Kolthammer, and Ian A Walmsley. Linear optical quantum computing in a single spatial mode. *Physical Review Letters*, 111(15):150501, 2013.

[112] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[113] Evgueni Karpov, David Daems, and NJ Cerf. Entanglement-enhanced classical capacity of quantum communication channels with memory in arbitrary dimensions. *Physical Review A*, 74(3):032320, 2006.

[114] Christopher King. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory*, 49(1):221–229, 2003.

[115] Christopher King, Michael Nathanson, and Mary Beth Ruskai. Qubit channels can require more than two inputs to achieve capacity. *Physical Review Letters*, 88(5):057901, 2002.

[116] Aleks Kissinger and Sander Uijlen. A categorical semantics for causal structure. In *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–12. IEEE, 2017.

[117] Dennis Kretschmann and Reinhard F Werner. Quantum channels with memory. *Physical Review A*, 72(6):062323, 2005.

[118] Hlér Kristjánsson, Yan Zhong, Anthony Munson, and Giulio Chiribella. Quantum networks with coherent routing of information through multiple nodes. *In preparation*.

[119] Hlér Kristjánsson, Giulio Chiribella, Sina Salek, Daniel Ebler, and Matthew Wilson. Resource theories of communication. *New Journal of Physics*, 22(7):073014, 2020.

[120] Hlér Kristjánsson, Robert Gardner, and Giulio Chiribella. *Quantum Communications Report for Ofcom*. Ofcom, 2021. URL https://www.ofcom.org.uk/research-and-data/technology/general/quantum-communications.

[121] Hlér Kristjánsson, Wenxu Mao, and Giulio Chiribella. Witnessing latent time correlations with a single quantum particle. *Physical Review Research*, 3(4):043147, 2021.

[122] Thaddeus D Ladd, Fedor Jelezko, Raymond Laflamme, Yasunobu Nakamura, Christopher Monroe, and Jeremy Lloyd O'Brien. Quantum computers. *Nature*, 464(7285):45, 2010.

[123] L-P Lamoureux, Edouard Brainis, NJ Cerf, Ph Emplit, Marc Haelterman, and Serge Massar. Experimental error filtration for quantum communication over highly noisy channels. *Physical Review Letters*, 94(23):230501, 2005.

[124] Rolf Landauer. Is quantum mechanics useful? *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 353(1703):367–376, 1995.

[125] Benjamin P Lanyon, Marco Barbieri, Marcelo P Almeida, Thomas Jennewein, Timothy C Ralph, Kevin J Resch, Geoff J Pryde, Jeremy L O'Brien, Alexei Gilchrist, and Andrew G White. Simplifying quantum logic using higher-dimensional Hilbert spaces. *Nature Physics*, 5(2):134–140, 2009.

[126] Felix Leditzky, Debbie Leung, and Graeme Smith. Quantum and private capacities of low-noise channels. In *2017 IEEE Information Theory Workshop (ITW)*, pages 484–488. IEEE, 2017.

[127] Felix Leditzky, Debbie Leung, and Graeme Smith. Dephrasure channel and superadditivity of coherent information. *Physical review letters*, 121(16):160501, 2018.

[128] Lev B Levitin. On quantum measure of information. In *Proceedings of the Fourth All-Union Conference on Information Transmission and Coding Theory, Tashkent (1969)*, 1969.

[129] Xi-Han Li and Shohini Ghose. Hyperentanglement concentration for time-bin and polarization hyperentangled photons. *Physical Review A*, 91(6):062302, 2015.

[130] Yidong Liao, Daniel Ebler, Feiyang Liu, and Oscar Dahlsten. Quantum speed-up in global optimization of binary neural nets. *New Journal of Physics*, 23(6):063013, 2021.

[131] Yunchao Liu and Xiao Yuan. Operational resource theory of quantum channels. *Physical Review Research*, 2(1):012035, 2020.

[132] Zi-Wen Liu and Andreas Winter. Resource theories of quantum channels and the universal role of resource erasure. *arXiv preprint arXiv:1904.04201*, 2019.

[133] Nicolas Loizeau and Alexei Grinbaum. Channel capacity enhancement with indefinite causal order. *Physical Review A*, 101(1):012340, 2020.

[134] Robin Lorenz and Jonathan Barrett. Causal and compositional structure of unitary transformations. *Quantum*, 5:511, 2021.

[135] Chiara Macchiavello and G Massimo Palma. Entanglement-enhanced information transmission over a quantum channel with correlated noise. *Physical Review A*, 65:050301, 2002.

[136] Chiara Macchiavello, G Massimo Palma, and S Virmani. Transition behavior in the channel capacity of two-quibit channels with memory. *Physical Review A*, 69(1):010303(R), 2004.

[137] Wenxu Mao. *Quantum Superposition of Communication Lines* (BSc Thesis). The University of Hong Kong, 2019.

[138] Ivan Marcikic, Hugues de Riedmatten, Wolfgang Tittel, Valerio Scarani, Hugo Zbinden, and Nicolas Gisin. Time-bin entangled qubits for quantum communication created by femtosecond pulses. *Physical Review A*, 66(6):062308, 2002.

[139] Eduardo Martín-Martínez, Luis J Garay, and Juan León. Unveiling quantum entanglement degradation near a Schwarzschild black hole. *Physical Review D*, 82(6):064006, 2010.

[140] Francesco Massa, Amir Moqanaki, Ämin Baumeler, Flavio del Santo, Joshua A Kettlewell, Borivoje Dakić, and Philip Walther. Experimental two-way communication with one photon. *Advanced Quantum Technologies*, 2(11):1900050, 2019.

[141] Laleh Memarzadeh, Chiara Macchiavello, and Stefano Mancini. Recovering quantum information through partial access to the environment. *New Journal of Physics*, 13(10):103031, 2011.

[142] Nikolai Miklin, Alastair A Abbott, Cyril Branciard, Rafael Chaves, and Costantino Budroni. The entropic approach to causal correlations. *New Journal of Physics*, 19(11):113041, 2017.

[143] Anthony Munson. *Communication Advantages for the Superposition of Asymptotically Long Sequences of Quantum Channels* (MSc Thesis). University of Oxford, 2020.

[144] Mio Murao. *Higher-order quantum operations of blackbox unitaries*. Talk at Q-Turn, 2020.

[145] Michael A Nielsen and Isaac L Chuang. *Quantum information and quantum computation*. Cambridge University Press, 2000.

[146] Daniel KL Oi. Interference of quantum channels. *Physical Review Letters*, 91(6):067902, 2003.

[147] Ognyan Oreshkov. Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics. *Quantum*, 3:206, 2019.

[148] Ognyan Oreshkov and Christina Giarmatzi. Causal and causally separable processes. *New Journal of Physics*, 18(9):093020, 2016.

[149] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092, 2012.

[150] Nick Ormrod, Augustin Vanrietvelde, and Jonathan Barrett. Causal structure in the presence of sectorial constraints, with application to the quantum switch. *arXiv preprint arXiv:2204.10273*, 2022.

[151] Daniel K Park, Ilya Sinayskiy, Mark Fingerhuth, Francesco Petruccione, and June-Koo Kevin Rhee. Parallel quantum trajectories via forking for sampling without redundancy. *New Journal of Physics*, 21(8):083024, 2019.

[152] Nikola Paunković and Marko Vojinović. Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders. *Quantum*, 4:275, 2020.

[153] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[154] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413, 1996.

[155] Paolo Perinotti. Causal structures and the classification of higher order quantum computations. In *Time in Physics*, pages 103–127. Springer, 2017.

[156] MB Plenio and S Virmani. Spin chains and channels with memory. *Physical Review Letters*, 99(12):120504, 2007.

[157] Felix A. Pollock, César Rodríguez-Rosario, Thomas Frauenheim, Mauro Paternostro, and Kavan Modi. Non-markovian quantum processes: Complete framework and efficient characterization. *Physical Review A*, 97:012127, 2018.

[158] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[159] JF Poyatos, J Ignacio Cirac, and Peter Zoller. Complete characterization of a quantum process: the two-bit quantum gate. *Physical Review Letters*, 78(2):390, 1997.

[160] Lorenzo M Procopio, Amir Moqanaki, Mateus Araújo, Fabio Costa, Irati Alonso Calafell, Emma G Dowd, Deny R Hamel, Lee A Rozema, Časlav Brukner, and Philip Walther. Experimental superposition of orders of quantum gates. *Nature Communications*, 6:7913, 2015.

[161] Lorenzo M Procopio, Francisco Delgado, Marco Enríquez, Nadia Belabas, and Juan Ariel Levenson. Communication enhancement through quantum coherent control of n channels in an indefinite causal-order scenario. *Entropy*, 21(10):1012, 2019.

[162] Lorenzo M. Procopio, Francisco Delgado, Marco Enríquez, Nadia Belabas, and Juan Ariel Levenson. Sending classical information via three noisy channels in superposition of causal orders. *Physical Review A*, 101:012346, 2020.

[163] Tom Richardson and Ruediger Urbanke. *Modern Coding Theory*. Cambridge University Press, 2008.

[164] Jonathan G Richens, Ciarán M Lee, and Saurabh Johri. Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications*, 11(1):1–9, 2020.

[165] Giulia Rubino, Lee A Rozema, Adrien Feix, Mateus Araújo, Jonas M Zeuner, Lorenzo M Procopio, Časlav Brukner, and Philip Walther. Experimental verification of an indefinite causal order. *Science Advances*, 3(3):e1602589, 2017.

[166] Giulia Rubino, Lee A Rozema, Daniel Ebler, Hlér Kristjánsson, Sina Salek, Philippe Allard Guérin, Alastair A Abbott, Cyril Branciard, Časlav Brukner, Giulio Chiribella, et al. Experimental quantum communication enhancement by superposing trajectories. *Physical Review Research*, 3(1):013093, 2021.

[167] Giovanna Ruggeri, Giulio Soliani, Vittorio Giovannetti, and Stefano Mancini. Information transmission through lossy bosonic memory channels. *EPL (Europhysics Letters)*, 70(6): 719, 2005.

[168] Sina Salek, Daniel Ebler, and Giulio Chiribella. Quantum communication in a superposition of causal orders. *arXiv preprint arXiv:1809.06655*, 2018.

[169] Gaurav Saxena, Eric Chitambar, and Gilad Gour. Dynamical resource theory of quantum coherence. *Physical Review Research*, 2(2):023298, 2020.

[170] Bernhard Schölkopf. Causality for machine learning. In *Probabilistic and Causal Inference: The Works of Judea Pearl*, pages 765–804. 2022.

[171] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

[172] Benjamin Schumacher and Michael D Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131, 1997.

[173] Santiago Sempere-Llagostera, Robert Gardner, Kwok Ho Wan, Raj B. Patel, Hlér Kristjánsson, Giulio Chiribella, and Ian A. Walmsley. Experimental perfect correction of noisy channels via quantum superposition of paths. *In preparation*.

[174] Claude Elwood Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[175] Claudia Shi, David Blei, and Victor Veitch. Adapting neural networks for the estimation of treatment effects. *Advances in neural information processing systems*, 32, 2019.

[176] Peter W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134. IEEE, 1994.

[177] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52:R2493–R2496, 1995.

[178] Peter W Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65. IEEE, 1996.

[179] Peter W Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002.

[180] Ralph Silva, Yelena Guryanova, Nicolas Brunner, Noah Linden, Anthony J Short, and Sandu Popescu. Pre-and postselected quantum states: Density matrices, tomography, and kraus operators. *Physical Review A*, 89(1):012121, 2014.

[181] Ralph Silva, Yelena Guryanova, Anthony J Short, Paul Skrzypczyk, Nicolas Brunner, and Sandu Popescu. Connecting processes with indefinite causal order and multi-time quantum states. *New Journal of Physics*, 19(10):103022, 2017.

[182] Graeme Smith and Jon Yard. Quantum communication with zero-capacity channels. *Science*, 321(5897):1812–1815, 2008.

[183] Graeme Smith, John A Smolin, and Andreas Winter. The quantum capacity with symmetric side channels. *IEEE Transactions on Information Theory*, 54(9):4208–4217, 2008.

[184] Akihito Soeda. *Limitations on quantum subroutine designing due to the linear structure of quantum operators*. Talk at the International Conference on Quantum Information and Technology (IC- QIT), 2013.

[185] Márcio M Taddei, Jaime Cariñe, Daniel Martínez, Tania García, Nayda Guerrero, Alastair A Abbott, Mateus Araújo, Cyril Branciard, Esteban S Gómez, Stephen P Walborn, et al. Experimental computational advantage from superposition of multiple temporal orders of quantum gates. *arXiv preprint arXiv:2002.07817*, 2020.

[186] Ryuji Takagi, Kun Wang, and Masahito Hayashi. Application of the resource theory of channels to communication scenarios. *Physical Review Letters*, 124(12):120502, 2020.

[187] Thomas Theurer, Dario Egloff, Lijian Zhang, and Martin B Plenio. Quantifying operations with an application to coherence. *Physical Review Letters*, 122(19):190405, 2019.

[188] Jayne Thompson, Kavan Modi, Vlatko Vedral, and Mile Gu. Quantum plug n' play: modular computation in the quantum regime. *New Journal of Physics*, 20(1):013004, 2018.

[189] William G Unruh. Maintaining coherence in quantum computers. *Physical Review A*, 51 (2):992, 1995.

[190] Augustin Vanrietvelde and Giulio Chiribella. Universal control of quantum processes using sector-preserving channels. *Quantum Information and Computation*, 21(15&16):1320–1352, 2021.

[191] Augustin Vanrietvelde, Hlér Kristjánsson, and Jonathan Barrett. Routed quantum circuits. *Quantum*, 5:503, 2021.

[192] Augustin Vanrietvelde, Nick Ormrod, Hlér Kristjánsson, and Jonathan Barrett. Consistent circuits for indefinite causal order. *arXiv preprint arXiv:2206.10042*, 2022.

[193] Lorenza Viola, Emanuel Knill, and Raymond Laflamme. Constructing qubits in physical systems. *Journal of Physics A: Mathematical and General*, 34(35):7067, 2001.

[194] Kwok Ho Wan, Oscar Dahlsten, Hlér Kristjánsson, Robert Gardner, and MS Kim. Quantum generalisation of feedforward neural networks. *npj Quantum Information*, 3(1):36, 2017.

[195] Julian Wechs, Hippolyte Dourdent, Alastair A Abbott, and Cyril Branciard. Quantum circuits with classical versus quantum control of causal order. *PRX Quantum*, 2(3):030335, 2021.

[196] Mark M Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[197] William K Wootters and Wojciech H Zurek. Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of bohr's principle. *Physical Review D*, 19(2):473, 1979.

[198] Xing Xiao, Yao Yao, Ying-Mao Xie, Xing-Hua Wang, and Yan-Ling Li. Protecting entanglement from correlated amplitude damping channel using weak measurement and quantum measurement reversal. *Quantum Information Processing*, 15(9):3881–3891, 2016.

[199] Paolo Zanardi, Daniel A Lidar, and Seth Lloyd. Quantum tensor product structures are observable induced. *Physical Review Letters*, 92(6):060402, 2004.

[200] Yujie Zhang, Xinan Chen, and Eric Chitambar. Building multiple access channels with a single particle. *Quantum*, 6:653, 2022.

[201] Xiaobin Zhao, Yuxiang Yang, and Giulio Chiribella. Quantum metrology with indefinite causal order. *Physical Review Letters*, 124(19):190503, 2020.

[202] Xiao-Qi Zhou, Timothy C Ralph, Pruet Kalasuwan, Mian Zhang, Alberto Peruzzo, Benjamin P Lanyon, and Jeremy L O'Brien. Adding control to arbitrary unknown quantum operations. *Nature Communications*, 2(1):1–8, 2011.

[203] Magdalena Zych, Fabio Costa, Igor Pikovski, and Časlav Brukner. Bell's theorem for temporal order. *Nature Communications*, 10(1):1–10, 2019.

A
Proofs of Theorems

A.1 Proof of Theorem 4 and Lemma 1

Proof of Lemma 4. The No Leakage Condition can be rewritten as

$$\mathrm{Tr}\left[\left(\sum_i P_A \widetilde{C}_i^\dagger P_A \widetilde{C}_i P_A\right)\rho\right] = \mathrm{Tr}[P_A\rho] \qquad \forall \rho \in \mathsf{St}(A)\,,$$

which is equivalent to

$$\sum_i P_A \widetilde{C}_i^\dagger P_A \widetilde{C}_i P_A = P_A\,. \tag{A.1}$$

The normalisation of Kraus operators implies the inequality

$$\begin{aligned} P_A &= P_A \left(\sum_i \widetilde{C}_i^\dagger \widetilde{C}_i\right) P_A \\ &= \sum_i P_A \widetilde{C}_i^\dagger \left(P_A + P_A^\perp\right) \widetilde{C}_i P_A \qquad P_A^\perp := I - P_A \\ &\leq \sum_i P_A \widetilde{C}_i^\dagger P_A \widetilde{C}_i P_A\,, \end{aligned} \tag{A.2}$$

where the equality sign holds if and only if

$$\sum_i P_A \widetilde{C}_i^\dagger P_A^\perp \widetilde{C}_i P_A = 0 \,, \tag{A.3}$$

or equivalently

$$\sum_i \left(P_A^\perp \widetilde{C}_i P_A\right)^\dagger \left(P_A^\perp \widetilde{C}_i P_A\right) = 0 \,, \tag{A.4}$$

The fact that every term in the sum is a positive semidefinite operator implies that the equality holds if and only if each term is zero, that is, if and only if $P_A^\perp \widetilde{C}_i P_A = 0$ for all i.

Putting these results together, we obtain

$$\forall i : \quad \widetilde{C}_i P_A = (P_A + P_A^\perp)\widetilde{C}_i P_A = P_A \widetilde{C}_i P_A \,, \tag{A.5}$$

as stated in Equation (3.4). □

Proof of Theorem 4. 1 $\Longrightarrow$ 2. Let $\mathcal{S} \in \mathsf{Chan}(A \oplus B)$ be a superposition of channels $\mathcal{A} \in \mathsf{Chan}(A)$ and $\mathcal{B} \in \mathsf{Chan}(B)$, and let $\mathcal{S}(\rho) = \sum_i S_i \rho S_i^\dagger$ be a Kraus decomposition of $\mathcal{S}$. Since $\mathcal{S}$ satisfies the No Leakage Condition for A, we have by Lemma 1 that

$$S_i P_A = P_A S_i P_A \qquad \forall i \in \{1, \dots r\} \,. \tag{A.6}$$

Similarly, since $\mathcal{S}$ satisfies the No Leakage Condition for B, we have that

$$S_i P_B = P_B S_i P_B \qquad \forall i \in \{1, \dots r\} \,. \tag{A.7}$$

Combining Equations (A.6) and (A.7) we obtain $S_i = S_i(P_A + P_B) = A_i \oplus B_i$, with $A_i := P_A S_i P_A$ and $B_i := P_B S_i P_B$. By the definition of a superposition of channels, the restriction of $\mathcal{S}$ to sector A is the channel $\mathcal{A}$, and thus $\mathcal{S}(P_A \rho P_A) = \mathcal{A}(P_A \rho P_A)$. Hence, we conclude that $\{A_i\}_{i=1}^r$ is a Kraus representation of $\mathcal{A}$. Similarly, since the restriction of $\mathcal{S}$ to sector B is the channel $\mathcal{B}$, and therefore $\{B_i\}_{i=1}^r$ is a Kraus representation of $\mathcal{B}$.

2 $\Longrightarrow$ 1 is immediate.

2 $\Longrightarrow$ 3. Consider the Stinespring representation of the channel $\mathcal{S}$, obtained by introducing an environment E of dimension r, equal to the number of Kraus operators of $\mathcal{S}$. Explicitly, the Stinespring representation is given by the isometry $V = \sum_{i=1}^r S_i \otimes |i\rangle$, where $\{|i\rangle\}_{i=1}^r$ is the canonical basis for E. Since each S_i is of the form $S_i = A_i \oplus B_i$, the isometry V is of the form $V = V_A \oplus V_B$,

where $V_A : \mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_E$ and $V_B : \mathcal{H}_B \to \mathcal{H}_B \otimes \mathcal{H}_E$ are the isometries

$$V_A := \sum_{i=1}^{r} A_i \otimes |i\rangle \tag{A.8}$$

$$V_B := \sum_{i=1}^{r} B_i \otimes |i\rangle \, . \tag{A.9}$$

Each isometry V_A and V_B can be extended to a unitary U_A and U_B, so that $V_A = U_A(I_A \otimes |\eta_A\rangle)$ and $V_B = U_B(I_B \otimes |\eta_B\rangle)$, where $|\eta_A\rangle$ and $|\eta_B\rangle$ are unit vectors in $\mathcal{H}_E$.

Without loss of generality, we can choose $|\eta_A\rangle = |\eta_B\rangle = |\eta\rangle$. Each unitary U_A and U_B can be realised as a time evolution for a time T with Hamiltonian H_{AE} and H_{BE}, respectively. Therefore, we can define the unitary evolutions $U_A := \exp[-iH_{AE}\,T/\hbar]$, $U_B := \exp[-iH_{BE}\,T/\hbar]$, and

$$U := \exp[-i(H_{AE} \oplus H_{BE})T/\hbar] = U_A \oplus U_B \, . \tag{A.10}$$

With these definitions, we obtain

$$\mathrm{Tr}_E \left[U_{AE}(\rho \otimes |\eta\rangle\langle\eta|) U_{AE}^\dagger \right] = \sum_i K_i \rho K_i^\dagger \, , \tag{A.11}$$

where

$$\begin{aligned} K_i &:= \big(I_A \otimes \langle i|\big)\, U_{AE}\, \big(I_A \otimes |\eta\rangle\big) \\ &= \big(I_A \otimes \langle i|\big) V_A \\ &= A_i \, , \end{aligned} \tag{A.12}$$

using Eq. (A.8) in the last equality. Similarly, we obtain

$$\mathrm{Tr}_E \left[U_{BE}(\rho \otimes |\eta\rangle\langle\eta|) U_{BE}^\dagger \right] = \sum_i L_i \rho L_i^\dagger \, , \tag{A.13}$$

with

$$\begin{aligned} L_i &:= \big(I_B \otimes \langle i|\big)\, U_{BE}\, \big(I_B \otimes |\eta\rangle\big) \\ &= \big(I_B \otimes \langle i|\big) V_B \\ &= B_i \, , \end{aligned} \tag{A.14}$$

using Eq. (A.9) in the last equality, and finally

$$\mathrm{Tr}_E\left[U(\rho\otimes|\eta\rangle\langle\eta|)U^\dagger\right] = \sum_i M_i\rho M_i^\dagger\,, \tag{A.15}$$

with

$$\begin{aligned} M_i &:= \big(I_S\otimes\langle i|\big)\,U\,\big(I_S\otimes|\eta\rangle\big) \\ &= \big(I_A\otimes\langle i|\big)\,U_{AE}\,\big(I_A\otimes|\eta\rangle\big)\;\oplus\;\big(I_B\otimes\langle i|\big)\,U_{BE}\,\big(I_B\otimes|\eta\rangle\big) \\ &= L_i\oplus K_i \\ &= A_i\oplus B_i\,, \end{aligned} \tag{A.16}$$

where $S := A\oplus B$.

3 $\Longrightarrow$ 1. Let E be an environment, let $|\eta\rangle\in\mathcal{H}_E$ be a pure state, and let $\mathcal{H}_{AE}$, $\mathcal{H}_{BE}$ be Hamiltonians with supports in $\mathcal{H}_A\otimes\mathcal{H}_E$ and $\mathcal{H}_B\otimes\mathcal{H}_E$, respectively, such that

$$\begin{aligned} \mathcal{A}(\rho) &= \mathrm{Tr}_E[U_{AE}(\rho\otimes\eta)U_{AE}^\dagger] & U_{AE} &= \exp[-iH_{AE}\,T/\hbar] \\ \mathcal{B}(\rho) &= \mathrm{Tr}_E[U_{BE}(\rho\otimes\eta)U_{BE}^\dagger] & U_{BE} &= \exp[-iH_{BE}\,T/\hbar] \\ \mathcal{S}(\rho) &= \mathrm{Tr}_E[U(\rho\otimes\eta)U^\dagger] & U &= \exp[-i(H_{AE}\oplus H_{BE})\,T/\hbar]\equiv U_{AE}\oplus U_{BE}\,. \end{aligned} \tag{A.17}$$

By construction, if ρ has support in $\mathcal{H}_A$, then

$$\begin{aligned} \mathcal{S}(\rho) &= \mathrm{Tr}_E[UP_{AE}\,(\rho\otimes\eta)\,P_{AE}U^\dagger] \qquad P_{AE} := P_A\otimes I_E \\ &= \mathrm{Tr}_E[U_{AE}\,(\rho\otimes\eta)\,U_{AE}^\dagger] \\ &= \mathcal{A}(\rho)\,. \end{aligned} \tag{A.18}$$

Similarly, if ρ has support in $\mathcal{H}_B$, then

$$\begin{aligned} \mathcal{S}(\rho) &= \mathrm{Tr}_E[UP_{BE}\,(\rho\otimes\eta)\,P_{BE}U^\dagger] \qquad P_{BE} := P_B\otimes I_E \\ &= \mathrm{Tr}_E[U_{BE}\,(\rho\otimes\eta)\,U_{BE}^\dagger] \\ &= \mathcal{B}(\rho)\,. \end{aligned} \tag{A.19}$$

Therefore, $\mathcal{S}$ is a superposition of $\mathcal{A}$ and $\mathcal{B}$. □

A.2 Proof of Proposition 15

The proof uses the following lemma:

Lemma 22. *Let $\mathcal{D}$ be a completely depolarising channel with vacuum extension $\widetilde{\mathcal{D}}$ and vacuum interference operator F. Then, the operator norm of F satisfies the inequality* $||F||_\infty \leq \frac{1}{\sqrt{d}}$.

Proof. Let the Kraus operators and vacuum amplitudes of $\mathcal{D}$ be given by $\{A_i\}$, $\{\alpha_i\}$, respectively. By definition,

$$||F||_\infty = \max_{\{|v\rangle:|||v\rangle||=1\}} \max_{\{|w\rangle:|||w\rangle||=1\}} \langle v|F|w\rangle \tag{A.20}$$

and

$$\begin{aligned} |\langle v|F|w\rangle| &= \left|\sum_i \overline{\alpha_i}\langle v|A_i|w\rangle\right| \\ &\leq \sqrt{\left(\sum_i |\alpha_i|^2\right)\left(\sum_j \langle v|A_j|w\rangle\langle w|A_j^\dagger|v\rangle\right)} \\ &= \sqrt{\langle v|\mathcal{D}\left(|w\rangle\langle w|\right)|v\rangle} \end{aligned} \tag{A.21}$$

If $\mathcal{D}$ is the completely depolarising channel, then $\mathcal{D}(|w\rangle\langle w|) = I/d$ and therefore the bound becomes $|\langle v|F|w\rangle| \leq \sqrt{1/d}$ which implies $||F||_\infty \leq \sqrt{1/d}$. □

We are now ready to provide the proof of Proposition 15.

Proof of Proposition 15. To prove that a channel is entanglement-breaking, it is sufficient to show that it transforms a maximally entangled state into a separable state [107]. We use the maximally entangled state $|\Phi^+\rangle := \sum_{k=0}^{d-1} |k\rangle \otimes |k\rangle/\sqrt{d}$. When the channel $\mathcal{C}_{\omega,F}$ is applied to this state, the output state is

$$(\mathcal{C}_{\omega,F} \otimes \mathcal{I})(|\Phi^+\rangle\langle\Phi^+|) = \left(\frac{I\otimes I}{d^2} + G_F\right) \otimes \frac{\omega}{2} + \left(\frac{I\otimes I}{d^2} - G_F\right) \otimes \frac{Z\omega Z}{2}\,, \tag{A.22}$$

where $G_F := (F\otimes I)(|\Phi^+\rangle\langle\Phi^+|)(F\otimes I)^\dagger$.

We now show that the operators $I\otimes I/d^2 \pm G_F$ are proportional to states with a positive partial transpose. To show this, we find that the partial transpose of G_F on the second space is

$$G_F^{T_2} = (F\otimes I)\frac{\texttt{SWAP}}{d}(F\otimes I)^\dagger\,. \tag{A.23}$$

Therefore, for every unit vector $|\Psi\rangle$, we obtain the following bound:

$$\begin{aligned}\langle\Psi|G_F^{T_2}|\Psi\rangle &= \frac{\langle\Psi|(F\otimes F^\dagger)|\Psi\rangle}{d}\\ &\leq \frac{\langle\Psi|(FF^\dagger\otimes I)|\Psi\rangle}{d}\\ &\leq \frac{\|FF^\dagger\|_\infty}{d}\\ &= \frac{\|F\|_\infty^2}{d}\\ &\leq \frac{1}{d^2}\,,\end{aligned} \tag{A.24}$$

where the first inequality follows from the Cauchy-Schwarz inequality and the last inequality follows from Lemma 22.

Eq. (A.24) implies that

$$\langle\Psi|\left(\frac{I\otimes I}{d^2}\pm G_F\right)^{T_2}|\Psi\rangle \geq \frac{1}{d^2}-\langle\Psi|G_F^{T_2}|\Psi\rangle \geq 0\,. \tag{A.25}$$

Since $|\Psi\rangle$ is an arbitrary vector, we conclude that the operator $\left(\frac{I\otimes I}{d^2}\pm G_F\right)^{T_2}$ has positive partial transpose. For $d=2$, the Peres-Horodecki criterion [108, 154], guarantees that $\frac{I\otimes I}{4}\pm G_F$ is proportional to a separable state. Therefore, the output state (A.22) is separable. □

A.3 Proofs of Proposition 16

Proof of Proposition 16. For a fixed vacuum extension, and therefore for a fixed vacuum interference operator F, the Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ is upper bounded by the Holevo capacity of the channel $\mathcal{C}_{|+\rangle\langle+|,F}$ (Lemma 14). Hence, it is enough to prove the bound for the channel $\mathcal{C}_{|+\rangle\langle+|,F}$.

Note that the output of channel $\mathcal{C}_{|+\rangle\langle+|,F}$ has dimension $2d$. For a generic channel $\mathcal{E}$ with $(2d)$-dimensional output, the Holevo capacity is upper bounded as [103]

$$\chi(\mathcal{E}) \leq \log(2d) - \min_\rho H\left[\mathcal{E}(\rho)\right]\,, \tag{A.26}$$

where $H(\rho) := -\operatorname{Tr}[\rho\log\rho]$ is the von Neumann entropy, and the minimisation can be restricted without loss of generality to pure states.

We now upper bound the right-hand-side of Eq. (A.26) for $\mathcal{E} = \mathcal{C}_{|+\rangle\langle+|,F}$. The action of the channel $\mathcal{C}_{|+\rangle\langle+|,F}$ on a generic input state ρ is

$$\mathcal{C}_{|+\rangle\langle+|,F}(\rho) = \frac{\frac{I}{d}+F\rho F^\dagger}{2} \otimes |+\rangle\langle+| + \frac{\frac{I}{d}-F\rho F^\dagger}{2} \otimes |-\rangle\langle-| \,. \tag{A.27}$$

In the case of a pure state $\rho = |\psi\rangle\langle\psi|$, we write $F|\psi\rangle = k\,|\phi\rangle$, where $|\phi\rangle$ is a unit vector and k is a normalisation constant. With this notation, we obtain

$$\begin{aligned} \mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|) = &\frac{\left(\frac{1}{d}+k^2\right)|\phi\rangle\langle\phi| + \frac{1}{d}P_\perp}{2} \otimes |+\rangle\langle+| \\ &+ \frac{\left(\frac{1}{d}-k^2\right)|\phi\rangle\langle\phi| + \frac{1}{d}P_\perp}{2} \otimes |-\rangle\langle-| \,, \end{aligned} \tag{A.28}$$

with $P_\perp := I - |\phi\rangle\langle\phi|$. The von Neumann entropy of this state is

$$\begin{aligned} H\left[\mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|)\right] &= -\frac{\frac{1}{d}+k^2}{2}\log\frac{\frac{1}{d}+k^2}{2} - \frac{d-1}{2d}\log\frac{1}{2d} \\ &\quad -\frac{\frac{1}{d}-k^2}{2}\log\frac{\frac{1}{d}-k^2}{2} - \frac{d-1}{2d}\log\frac{1}{2d} \\ &= \frac{d-1}{d}\log(2d) - \frac{\frac{1}{d}+k^2}{2}\log\frac{\frac{1}{d}+k^2}{2} - \frac{\frac{1}{d}-k^2}{2}\log\frac{\frac{1}{d}-k^2}{2} \end{aligned} \tag{A.29}$$

Now, note that

$$k = \|F\,|\psi\rangle\| \leq \|F\|_\infty \leq \frac{1}{\sqrt{d}} \,, \tag{A.30}$$

where the last inequality follows from Lemma 22. The expression (A.29) is monotonically decreasing for k in the interval $[0, 1/\sqrt{d}]$. Hence, we obtain the lower bound

$$\begin{aligned} H\left[\mathcal{C}_{|+\rangle\langle+|,F}(|\psi\rangle\langle\psi|)\right] \geq &\frac{d-1}{d}\log(2d) \\ &- \frac{\frac{1}{d}+\|F\|_\infty^2}{2}\log\frac{\frac{1}{d}+\|F\|_\infty^2}{2} \\ &- \frac{\frac{1}{d}-\|F\|_\infty^2}{2}\log\frac{\frac{1}{d}-\|F\|_\infty^2}{2} \,. \end{aligned} \tag{A.31}$$

Inserting this expression into Eq. (A.26) with $\mathcal{E} = \mathcal{C}_{|+\rangle\langle+|,F}$, we obtain Eq. (6.18). □

A.4 Lemmas for proof of Theorem 18

Lemma 23. *Let F be a vacuum interference operator. Let $\mathcal{C}_{\omega,F}$ be a superposition of two independent qubit completely depolarising channels with vacuum interference operator F, as given by Eq. (6.17). Let a, b be the singular values of F. Then $F' = a\,|0\rangle\langle 0| + b\,|1\rangle\langle 1|$ is also a vacuum interference operator associated with the qubit completely depolarising channel, with $\mathcal{C}_{\omega,F'}$ the corresponding superposition channel, and furthermore the classical capacity of the channel $\mathcal{C}_{\omega,F}$ is equal to that of the channel $\mathcal{C}_{\omega,F'}$.*

Proof. Using the singular value decomposition, F can be written as $F = UF'V$, where U and V are suitable unitary matrices, and F' is diagonal in the basis $\{|0\rangle, |1\rangle\}$. From Ref. [2], we know that any linear operator on a d-dimensional Hilbert space is a valid vacuum interference operator of the completely depolarising channel if it satisfies $\operatorname{Tr} F^\dagger F \leq 1/d$. Therefore, if F is a vacuum interference operator of the completely depolarising channel, then so is F', since $\operatorname{Tr} {F'}^\dagger F' = \operatorname{Tr}\left[(U^\dagger F V^\dagger)^\dagger U^\dagger F V^\dagger\right] = \operatorname{Tr} F^\dagger F \leq 1/d$.

Now, we have that the channel $\mathcal{C}_{\omega,F'}$ can be obtained from $\mathcal{C}_{\omega,F}$ via the unitary transformation $\mathcal{C}_{\omega,F'} = (\mathcal{U} \otimes \mathcal{I}_C)^\dagger \circ \mathcal{C}_{\omega,F} \circ \mathcal{V}^\dagger$, where $\mathcal{U}^\dagger$ and $\mathcal{V}^\dagger$ are the inverses of the unitary channels associated to the unitary matrices U and V, respectively, and $\mathcal{I}_C$ is the identity channel on the control system. The classical capacity of a channel is preserved under unitary transformations; therefore $\mathcal{C}_{\omega,F}$ and $\mathcal{C}_{\omega,F'}$ have equal classical capacity. □

Lemma 24. *When the operator F is diagonal in the computational basis, the Holevo capacity of the channel $\mathcal{C}_{\omega,F}$ is given by*

$$\chi(\mathcal{C}_{\omega,F}) = \max_{\{p_x, |\psi_x\rangle\}} \left\{ H\left[\mathcal{C}_{\omega,F}\left(\sum_{x,m} p_x\, |\langle m|\psi_x\rangle|^2\, |m\rangle\langle m|\right)\right] - \sum_x p_x\, H\left[\mathcal{C}_{\omega,F}(|\psi_x\rangle\langle\psi_x|)\right] \right\}, \tag{A.32}$$

where the maximum is over the ensembles of d pure states with positive coefficients in the computational basis.

The proof of this result, which relies on group-theoretic methods, is beyond the scope of this thesis, and is provided in Appendix C of Ref. [121].

A.5 Proof of Proposition 20

Here we provide a proof of Proposition 20. The proof follows from the following Lemmas, proven at the end of this Appendix.

Lemma 25. *Let $\mathcal{N}_1 \in \mathsf{Chan}(X)$ and $\mathcal{N}_2 \in \mathsf{Chan}(X)$ be two quantum channels, and let $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$ be their vacuum extensions. Then, $\widetilde{\mathcal{N}}_1 \circ \widetilde{\mathcal{N}}_2$ is a vacuum extension of $\mathcal{N}_1 \circ \mathcal{N}_2$, and its vacuum interference operator is F_1F_2, where F_1 (F_2) is the vacuum interference operator of $\widetilde{\mathcal{N}}_1$ ($\widetilde{\mathcal{N}}_2$).*

Proof. Using Equation (3.11) for the vacuum-extended channels $\widetilde{\mathcal{N}}_1$ and $\widetilde{\mathcal{N}}_2$, we obtain the relation

$$
\begin{aligned}
(\widetilde{\mathcal{N}}_1 \circ \widetilde{\mathcal{N}}_2)(\rho) =& (\mathcal{N}_1 \circ \mathcal{N}_2)(P_X \rho P_X) + \langle \mathrm{vac}|\rho|\mathrm{vac}\rangle \, |\mathrm{vac}\rangle\langle \mathrm{vac}| \\
&+ F_1 F_2 \rho |\mathrm{vac}\rangle\langle \mathrm{vac}| + |\mathrm{vac}\rangle\langle \mathrm{vac}|\rho F_2^\dagger F_1^\dagger \,,
\end{aligned}
\tag{A.33}
$$

which is valid for every $\rho \in \mathsf{St}(\widetilde{X})$. From Equation (A.33) we see that $\widetilde{\mathcal{N}}_1 \circ \widetilde{\mathcal{N}}_2$ is a vacuum extension of $\mathcal{N}_1 \circ \mathcal{N}_2$. Moreover, comparison of Equation (A.33) with Equation (3.11) shows that the vacuum interference operator of $\widetilde{\mathcal{N}}_1 \circ \widetilde{\mathcal{N}}_2$ is $F_1 F_2$. □

Lemma 26. *Let $F \in L(\mathcal{H}_X)$ be the vacuum interference operator associated to a generic vacuum-extended channel $\widetilde{\mathcal{C}} \in \mathsf{Chan}(\widetilde{X})$. Then, $\|F\|_\infty \le 1$.*

Proof. By the definition (3.12), the vacuum interference operator of the vacuum-extended channel $\widetilde{\mathcal{C}}$ can be expressed as $F = \sum_i \overline{\gamma}_i C_i$, where $\{\widetilde{C}_i := C_i \oplus \gamma_i |\mathrm{vac}\rangle\langle \mathrm{vac}|\}$ is an arbitrary Kraus representation of $\widetilde{\mathcal{C}}$.

By definition,

$$
\|F\|_\infty := \max_{\{|\varphi\rangle \in \mathcal{H}_X, \, \||\varphi\rangle\| = 1\}} \|F\,|\varphi\rangle\| \ . \tag{A.34}
$$

Let $|\varphi\rangle \in \mathcal{H}_X$ be a unit vector such that $\|F\|_\infty = \|F|\varphi\rangle\|$. Then, we obtain the following series of (in)equalities:

$$
\begin{aligned}
\|F\|_\infty^2 &= \|F|\varphi\rangle\|^2 \\
&= \langle\varphi|\, F^\dagger F \,|\varphi\rangle \\
&= \sum_{ij} \gamma_i \bar{\gamma}_j \,\langle\varphi|\, C_i^\dagger C_j \,|\varphi\rangle \\
&\le \sum_{ij} |\gamma_i| \left|\gamma_j\right| \left|\langle\varphi|\, C_i^\dagger C_j \,|\varphi\rangle\right| \\
&\le \sum_{ij} |\gamma_i| \left|\gamma_j\right| \sqrt{\langle\varphi|\, C_i^\dagger C_i \,|\varphi\rangle \,\langle\varphi|\, C_j^\dagger C_j \,|\varphi\rangle} \\
&= \left(\sum_i |\gamma_i| \sqrt{\langle\varphi|\, C_i^\dagger C_i \,|\varphi\rangle} \right)^2 \\
&\le \left(\sum_i \langle\varphi|\, C_i^\dagger C_i \,|\varphi\rangle \right)^2 \\
&= 1 \,,
\end{aligned}
\tag{A.35}
$$

where the first inequality is the triangle inequality for the modulus, and the second and third inequalities are Cauchy-Schwarz inequalities. $\square$

Lemma 27. *Let $\mathcal{C} \in \mathsf{Chan}(X)$ be a quantum channel on a quantum system of dimension $d \geq 2$, and let $\widetilde{\mathcal{C}} \in \mathsf{Chan}(\widetilde{X})$ be an arbitrary vacuum extension of $\mathcal{C}$. If the Choi operator*

$$c := \sum_{i,j} |i\rangle\langle j| \otimes \mathcal{C}(|i\rangle\langle j|)\,, \tag{A.36}$$

has full rank, then the vacuum interference operator F associated to $\widetilde{\mathcal{C}}$ satisfies the strict inequality $\|F\|_\infty < 1$.

Proof. Lemma 26 shows that the norm of F is smaller than or equal to 1. The equality $\|F\|_\infty = 1$ holds if and only if all the inequalities in Equation (A.35) hold with the equality sign. In the following we show that saturating the second inequality is impossible when c has full rank.

The second inequality in Equation (A.35) is saturated if and only if

$$C_i\,|\varphi\rangle \propto C_j\,|\varphi\rangle \qquad \forall i,j\,, \tag{A.37}$$

that is, if and only if

$$C_i\,|\varphi\rangle = \lambda_i|\varphi_0\rangle \qquad \forall i\,, \tag{A.38}$$

where $|\varphi_0\rangle \in \mathcal{H}_X$ is a fixed unit vector, and $\{\lambda_i\}$ are complex numbers.

Let $A = \sum_i a_i\, C_i$ be an arbitrary linear combination of the operators $\{C_i\}$, with complex coefficients $\{a_i\}$. Then, one has

$$\begin{aligned} A\,|\varphi\rangle &= \sum_i a_i\, C_i|\varphi\rangle \\ &= \left(\sum_i a_i\lambda_i\right)\,|\varphi_0\rangle\;. \end{aligned} \tag{A.39}$$

In other words, every linear combination of the operators $\{C_i\}$ must map $|\varphi\rangle$ into a vector proportional to $|\varphi_0\rangle$.

Now, the Choi operator c has full rank if and only if the operators C_i are a spanning set for the vector space $L(\mathcal{H}_X)$. This means, in particular, that there exist coefficients $\{a_i\}$ such that $A = \sum_i a_i\, C_i = |\varphi_0^\perp\rangle\langle\varphi|$, where $|\varphi_0^\perp\rangle$ is a unit vector orthogonal to $|\varphi_0\rangle$ (such a vector exists because the Hilbert space $\mathcal{H}_X$ is at least two-dimensional). In this case, $A|\varphi\rangle = |\varphi_0^\perp\rangle$, meaning that Equation (A.39) cannot be satisfied. This implies that Equation (A.37) cannot be satisfied, and that the bound $\|F\|_\infty \leq 1$ cannot hold with the equality sign. $\square$

Proof of Proposition 20. Let $\widetilde{\mathcal{N}}_{\text{dep}}$ be an arbitrary vacuum extension of the completely depolarising channel $\mathcal{N}_{\text{dep}}$. By Lemma 25, the relation

$$\widetilde{\mathcal{N}}_{\text{dep}} \circ \widetilde{\mathcal{N}}_{\text{dep}} = \widetilde{\mathcal{N}}_{\text{dep}} \tag{A.40}$$

implies the relation $F^2 = F$, where F is the vacuum interference operator of $\widetilde{\mathcal{N}}_{\text{dep}}$. In turn, the relation $F^2 = F$ implies the relation $F = F^n$ for every integer $n \in \mathbb{N}$. In terms of the norm, this condition gives the bound

$$\begin{aligned} \|F\|_\infty &= \|F^n\|_\infty \\ &\leq \|F\|_\infty^n \qquad \forall n \in \mathbb{N}\,. \end{aligned} \tag{A.41}$$

Now, the Choi operator of the completely depolarising channel $\mathcal{N}_{\text{dep}}$ is $c = I \otimes I/d$ and has full rank. Hence, Lemma 27 implies $\|F\|_\infty < 1$. This means that Eq. (A.41) implies that $\|F\|_\infty = 0$, and therefore $F = 0$. In summary, the only vacuum extension satisfying the condition (A.40) is the incoherent one. □

A.6 Proof of Theorem 21

Proof. Recall that the largest singular value of an operator H is equal to its operator norm $||H||_\infty$. Recall also that the superposition of two independent identical sequences of channels $\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}, \widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}] : \mathcal{L}(\mathcal{H}_X) \to \mathcal{L}(\mathcal{H}_Y \otimes \mathcal{H}_P)$ is given by

$$\begin{aligned} &\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}, \widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}](\rho) \\ &= \frac{\mathcal{E} \circ (\mathcal{R} \circ \mathcal{E})^{n-1}(\rho) + F(GF)^{n-1}\rho F(GF)^{n-1}}{2} \otimes |+\rangle\langle+| \\ &+ \frac{\mathcal{E} \circ (\mathcal{R} \circ \mathcal{E})^{n-1}(\rho) - F(GF)^{n-1}\rho F(GF)^{n-1}}{2} \otimes |-\rangle\langle-|\,. \end{aligned} \tag{A.42}$$

$1 \Rightarrow 2$: We prove the contrapositive. First, note that for any vacuum interference operator H, $||H||_\infty \leq 1$ (see Lemma 26, or Eq. (A.44) below). Let G be a vacuum interference operator of a quantum channel $\mathcal{R} : \mathcal{L}(\mathcal{H}_Y) \to \mathcal{L}(\mathcal{H}_X)$. Suppose $||F||_\infty \neq 1$. Then $||F||_\infty < 1$. Since the operator norm is sub-multiplicative, $||GF||_\infty \leq ||G||_\infty ||F||_\infty < 1$. It follows that as $n \to 0$, $(GF)^{n-1} \to 0$ and hence $F(GF)^{n-1} \to 0$. Therefore, the off-diagonal path terms vanish, and $\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}, \widetilde{\mathcal{E}} \circ (\widetilde{\mathcal{R}} \circ \widetilde{\mathcal{E}})^{n-1}](\rho) \to \mathcal{E} \circ (\mathcal{R} \circ \mathcal{E})^{n-1}(\rho) \otimes I/2$, which has capacity zero as $n \to \infty$.

$2 \Rightarrow 3$: By definition,

$$||F||_\infty := \max_{\{|\varphi\rangle \in \mathcal{H}_X,\, |||\varphi\rangle||=1\}} ||F\,|\varphi\rangle||\ , \tag{A.43}$$

where, as usual, $\|\cdot\|$ denotes the Euclidean norm. Now, letting $|\varphi\rangle \in \mathcal{H}_X$ be a unit vector such that $\|F\|_\infty = \|F|\varphi\rangle\|$, we obtain the following (in)equalities:

$$\begin{aligned}
||F||_\infty &= ||\bar{a}_0 E_0|\varphi\rangle + \bar{a}_1 E_1|\varphi\rangle + \cdots + \bar{a}_{r-1}E_{r-1}|\varphi\rangle|| \\
&\leq ||\bar{a}_0 E_0|\varphi\rangle|| + ||\bar{a}_1 E_1|\varphi\rangle|| + \cdots + ||\bar{a}_{r-1}E_{r-1}|\varphi\rangle|| \\
&= \left|\bar{a}_0\sqrt{\langle\varphi|E_0{}^\dagger E_0|\varphi\rangle}\right| + \cdots + \left|\bar{a}_{r-1}\sqrt{\langle\varphi|E_{r-1}{}^\dagger E_{r-1}|\varphi\rangle}\right| \\
&\leq \sqrt{(|a_0|^2 + |a_1|^2 + \cdots + |a_{r-1}|^2)} \\
&\cdot \sqrt{\langle\varphi|E_0{}^\dagger E_0|\varphi\rangle + \langle\varphi|E_1{}^\dagger E_1|\varphi\rangle + \cdots + \langle\varphi|E_{r-1}{}^\dagger E_0|\varphi\rangle} \\
&= 1 \cdot \sqrt{\langle\varphi|\sum_{i=0}^{r-1} E_i^\dagger E_i|\varphi\rangle} \\
&= 1
\end{aligned} \tag{A.44}$$

The first inequality follows from the triangle inequality and is saturated if and only if there exists a pure state $|\zeta\rangle \in \mathcal{L}(\mathcal{H}_Y)$ such that

$$\frac{\bar{a}_0 E_0|\varphi\rangle}{||\bar{a}_0 E_0|\varphi\rangle||} = \frac{\bar{a}_1 E_1|\varphi\rangle}{||\bar{a}_1 E_1|\varphi\rangle||} = \cdots = \frac{\bar{a}_{r-1}E_{r-1}|\varphi\rangle}{||\bar{a}_{r-1}E_{r-1}|\varphi\rangle||} = |\zeta\rangle\,. \tag{A.45}$$

If this equation is satisfied, then for all $i \in \{0, 1, \ldots, r-1\}$,

$$|\zeta\rangle\langle\zeta| = \frac{\bar{a}_i E_i|\varphi\rangle}{||\bar{a}_i E_i|\varphi\rangle||} \cdot \frac{a_i\langle\varphi|E_i^\dagger}{||a_i\langle\varphi|E_i^\dagger||} = \frac{E_i|\varphi\rangle\langle\varphi|E_i^\dagger}{\langle\varphi|E_i^\dagger E_i|\varphi\rangle}\,. \tag{A.46}$$

This implies that

$$\mathcal{E}(|\varphi\rangle\langle\varphi|) = \sum_{i=0}^{r-1} E_i|\varphi\rangle\langle\varphi|E_i^\dagger = \left(\sum_{i=0}^{r-1}\langle\varphi|E_i^\dagger E_i|\varphi\rangle\right)|\zeta\rangle\langle\zeta| = |\zeta\rangle\langle\zeta|\,. \tag{A.47}$$

That is, $||F||_\infty = 1$ only if there exists a pure state in $\mathcal{H}_X$ which is mapped to a pure state in $\mathcal{H}_Y$ by $\mathcal{E}$.

The second inequality follows from the Cauchy-Schwarz inequality. It is saturated if and only if there exists $\theta \in \mathbb{R}$ such that $a_i = e^{i\theta}\sqrt{\langle\varphi|E_i^\dagger E_i|\varphi\rangle}$ for all $i \in \{0, 1, \ldots, r-1\}$.

$2 \Rightarrow 1$: If $||F||_\infty = 1$, then there exist two states $|\varphi\rangle \in \mathcal{H}_X$ and $|\zeta\rangle \in \mathcal{H}_Y$ such that $F\,|\varphi\rangle = |\zeta\rangle$. Thus, $(|\varphi\rangle\,\langle\zeta|\,F)^n\,|\varphi\rangle = |\varphi\rangle$ for all $n \in \mathbb{N}$.

Let $\mathcal{R} : \mathsf{Chan}(Y \to X)$ be a quantum channel with vacuum extension $\widetilde{\mathcal{R}}$. Let the Kraus operators of $\widetilde{\mathcal{R}}$ be $\{\widetilde{R}_0 = |\varphi\rangle\langle\zeta| \oplus |\text{vac}\rangle\langle\text{vac}|\} \cup \{\widetilde{R}_i = |\varphi\rangle\langle i| \oplus 0\}_{i=1}^{r-1}$, where $\{|i\rangle\}_{i=0}^{r-1}$ is a basis of $\mathcal{H}_Y$

with the definition $|0\rangle := |\zeta\rangle$.

Let G be the vacuum interference operator of $\widetilde{\mathcal{R}}$. Then $G = |\varphi\rangle\langle\zeta|$ and thus $GF|\varphi\rangle = |\varphi\rangle$. It follows that the effective vacuum interference operator $F_{\text{eff}} = F(GF)^{n-1}$ satisfies $F_{\text{eff}}|\varphi\rangle = |\zeta\rangle$ and hence $||F_{\text{eff}}||_\infty = 1$.

Furthermore, there exists a state $|\xi\rangle \in \mathcal{H}_X$ orthogonal to $|\varphi\rangle$ such that $F_{\text{eff}}|\xi\rangle \neq |\zeta\rangle$; otherwise, $F_{\text{eff}}\frac{|\varphi\rangle+|\xi\rangle}{\sqrt{2}} = \sqrt{2}\,|\zeta\rangle$, which implies $||F_{\text{eff}}||_\infty \neq 1$, a contradiction.

Since $\mathcal{E}\circ(\mathcal{R}\circ\mathcal{E})^{n-1}$ has classical capacity that approaches zero as $n \to \infty$, then for all $\rho \in \mathsf{St}(X)$, $(\mathcal{E}\circ(\mathcal{R}\circ\mathcal{E})^{n-1})(\rho) = \sigma$ for some $\sigma \in \mathsf{St}(Y\otimes C)$. Therefore

$$\mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}}\circ(\widetilde{\mathcal{R}}\circ\widetilde{\mathcal{E}})^{n-1},\widetilde{\mathcal{E}}\circ(\widetilde{\mathcal{R}}\circ\widetilde{\mathcal{E}})^{n-1}](|\varphi\rangle\langle\varphi|) \neq \mathcal{S}_{\text{sup}}^{|+\rangle\langle+|}[\widetilde{\mathcal{E}}\circ(\widetilde{\mathcal{R}}\circ\widetilde{\mathcal{E}})^{n-1},\widetilde{\mathcal{E}}\circ(\widetilde{\mathcal{R}}\circ\widetilde{\mathcal{E}})^{n-1}](|\xi\rangle\langle\xi|)\,, \tag{A.48}$$

i.e., the superposition channel yields distinct output states for two distinct input states. This result proves that the superposition channel has a classical capacity strictly greater than zero as $n \to \infty$. □

B
Mathematical subtleties

B.1 Placed and unplaced channels

Here we specify the mathematical structure of placed and unplaced channels. In the following, the subset of completely positive (CP) maps will be denoted by $\mathsf{CP}(A \to B) \subset \mathsf{Map}(A \to B)$, and the subset of trace-preserving (TP) maps will be denoted by $\mathsf{TP}(A \to B) \subset \mathsf{Map}(A \to B)$.

For a single use of a single device, the distinction between placed and unplaced channels concerns only the type of inputs and outputs. Mathematically, placed and unplaced channels are both described by completely positive trace-preserving maps.

When multiple devices are involved, the distinction is more substantial. As described in §4.1.3–4.1.4, k devices can be placed in parallel, giving rise to a multipartite quantum channel, or in sequence, giving rise to a k-step quantum channel, or, more generally, in any combination of parallel and sequence, giving rise to an l-step quantum channel, with any l from 1 to k.

In contrast, unplaced channels belong to a different set of completely positive maps. We will now specify this set explicitly. So far, we represented unplaced channels by lists, such as $(\mathcal{N}_1, \dots, \mathcal{N}_k)$. However, the set of lists is not closed under probabilistic mixtures, which arise quite naturally when some of the parameters of the communication devices are subject to random fluctuations.

Probabilistic mixtures can be described by convex combinations of the form $\sum_{i=1}^{L} p_i \, (\mathcal{N}_{1,i}, \dots, \mathcal{N}_{k,i})$, where $(p_i)_{i=1}^{L}$ is a probability distribution, and, for every $i \in \{1, \dots, L\}$, $(\mathcal{N}_{1,i}, \dots, \mathcal{N}_{k,i}) \in \mathsf{Chan}(X_1 \to Y_1) \times \cdots \times \mathsf{Chan}(X_k \to Y_k)$ is a list of channels. The convex combination $\sum_{i=1}^{L} p_i \, (\mathcal{N}_{1,i}, \dots, \mathcal{N}_{k,i})$ must satisfy a basic consistency requirement: if a channel in the list $(\mathcal{N}_1, \dots, \mathcal{N}_k)$ is a convex combination of channels, say $\mathcal{N}_1 = \sum_i p_i \mathcal{N}_{1,i}$, then the list $(\mathcal{N}_1, \dots, \mathcal{N}_k)$ should be equal to the corresponding convex combination $\sum_i p_i (\mathcal{N}_{1,i}, \mathcal{N}_2, \dots, \mathcal{N}_k)$.

Requiring this consistency property to hold for every entry of the list implies that the convex combinations $\sum_{i=1}^{L} p_i\,(\mathcal{N}_{1,i},\ldots,\mathcal{N}_{k,i})$ can be represented as elements of the tensor product space $\mathsf{TP}(X_1 \to Y_1) \otimes \mathsf{TP}(X_2 \to Y_2) \otimes \cdots \otimes \mathsf{TP}(X_k \to Y_k)$, which consists of all linear combinations of the form

$$\mathcal{M} = \sum_{i=1}^{L} c_i\,\mathcal{N}_{1,i} \otimes \mathcal{N}_{2,i} \otimes \cdots \otimes \mathcal{N}_{k,i}\,, \tag{B.1}$$

where $(c_i)_{i=1}^{L}$ are real coefficients, and each $\mathcal{N}_{j,k}$ is a trace-preserving map in $\mathsf{TP}(X_j \to Y_j)$.

In summary, the unplaced channels can be regarded as elements of the tensor product space $\mathsf{TP}(X_1 \to Y_1) \otimes \mathsf{TP}(X_2 \to Y_2) \otimes \cdots \otimes \mathsf{TP}(X_k \to Y_k)$. Precisely, the list $(\mathcal{N}_1,\ldots,\mathcal{N}_k)$ can be regarded as the product channel $\mathcal{N}_1 \otimes \cdots \otimes \mathcal{N}_k$. In this work, we use the list notation $(\mathcal{N}_1,\ldots,\mathcal{N}_k)$ and the tensor product notation $\mathcal{N}_1 \otimes \cdots \otimes \mathcal{N}_k$, interchangeably, depending on which representation is more convenient.

As can be seen from Equation (B.1), the tensor product space $\mathsf{TP}(X_1 \to Y_1) \otimes \mathsf{TP}(X_2 \to Y_2) \otimes \cdots \otimes \mathsf{TP}(X_k \to Y_k)$ also contains channels that are not of the product form. In fact, the set of all channels in this tensor product space is in one-to-one correspondence with the set of k-partite no-signalling channels $\mathsf{NSChan}(X_1 \to Y_1,\ldots,X_k \to Y_k)$ (cf. §2.7.2).

This means that unplaced channels can, most generally, represent correlated devices. Such a multi-partite unplaced channel shall be represented in the usual way as a channel $\mathcal{N} \in \mathsf{NSChan}(X_1 \to Y_1,\ldots,X_k \to Y_k)$. If some parts of the overall channel factorise, we can write it as a product of no-signalling channels, e.g. $\mathcal{N}_{1,2} \otimes \mathcal{N}_3 \in \mathsf{NSChan}(X_1 \to Y_1, X_2 \to Y_2) \times \mathsf{Chan}(X_3 \to Y_3)$, or as the list $(\mathcal{N}_{1,2},\mathcal{N}_3)$.

This suggests, that multi-partite unplaced channels should be able to be placed either between multiple parties at separate locations, or between a single party at different moments in time, as long as these correlations do not lead to signalling between the different systems. Such scenarios are considered in §4.1.7.