# Quantum channel simulation and discrimination with applications to quantum communications

Jason Luke Pereira

Doctor of Philosophy

University of York

Computer Science

October 2020

# Abstract

Quantum channel discrimination is a highly versatile task in quantum information. Almost any physical process can be modelled as a quantum channel, so discrimination between channels has broad applications across many fields of science. Since the processes modelled by quantum channels and the contexts in which we want to discriminate between them are so wide-ranging, it should come as no surprise that the possible protocols are equally varied. The most general discrimination protocols can use any sequence of operations allowed by physics.

This is what makes protocol stretching such a powerful mathematical tool. It allows the calculation of bounds on the performance of any discrimination protocol and can be applied to any situation in which the channels involved are jointly simulable by some quantum processor.

Certain channels, such as the amplitude damping channels, cannot be simulated using standard teleportation. Others, like the lossy bosonic channels, can be (individually) simulated using teleportation, but two lossy channels with different losses cannot be jointly simulated.

In this thesis, we characterise port-based teleportation so that it can be used as a tool for channel simulation. Port-based teleportation is a variant of quantum teleportation that can simulate any channel in the asymptotic limit of infinite ports. For a finite number of ports, we can find resource states that simulate amplitude damping channels well. By combining our channel simulations with the technique of protocol stretching, we are able to tighten existing bounds on the discrimination of amplitude-damping channels.

We also address the relatively unstudied subfield of channel position finding. We use channel simulation to bound the performance of environment localisation, and we show the viability of idler-free channel position finding over pure loss channels. Finally, we calculate the secret key rate for a scenario of quantum hacking based on a side channel in the sender's device.

# Contents

# List of Figures

# Acknowledgements

I would not be where I am today, were it not for the ongoing love and support given to me by my family. My sincerest thanks go to Melwyn, Tina, and Aaron and to my grandparents: Yvonne, Selby, Angelo, and Lucy.

This thesis would not have been possible were it not for the help, support, and opportunities given to me by my supervisor, Prof. Stefano Pirandola. He has guided my research throughout my PhD and has given me the best possible start in the world of academia. I am also grateful for my collaboration with my co-authors: Leonardo Banchi and Quntao Zhuang.

My thanks go also to the others in the Quantum Information group, past and present, and especially to my office mates - Athena Karsa, Kieran Wilkinson, Alasdair Fletcher, and Cillian Harney - for the help, stimulating discussions, and friendship they have given me over the years. You have truly taught me the meaning of the phrase "Live, laugh, love". I thank also Timothy Atkinson and Chaitanya Kaul, who have welcomed me since my first day in the department.

I am sincerely grateful for all of my friends, both from York and from UCL, for their support, encouragement, and memes.

*A quantum physicist moved to a new house and found that the doorbell had an extraordinarily loud chime. So much so that it was the loudest doorbell that he, or any of his friends, had ever heard. Eventually, he decided to submit it to the Guinness book of World Records. He called them up and the woman on the phone was amazed by the sound of his doorbell, even over the call.*

*"We'll submit it right away," she said. "We'll send a carpenter around to remove it from the door frame so that we can measure exactly how loud it is at our lab."*

*"Wow!" exclaimed the physicist. "That seems excessive. Can't you just measure it here?"*

*"Oh no," replied the woman. "We can't perform a Bell measurement locally."*

# Declaration

I declare that the research described in this thesis is original work, which I undertook at the University of York during 2017 - 2020. Except where stated, all of the work contained within this thesis represents the original contribution of the author.

Some parts of this thesis have been published in journals or preprinted on the arXiv; where items were published jointly with collaborators, the author of this thesis is responsible for the material presented here. For each published item the primary author is the first listed author.

- *Hacking Alice's box in continuous-variable quantum key distribution*, Jason Pereira and Stefano Pirandola. Published in Physical Review A **98**, 062319, (2018). [8]
- *Optimal environment localization*, Jason Pereira, Quntao Zhuang, and Stefano Pirandola. Published in Physical Review Research **2**, (2020). [9]
- *Bounds on amplitude-damping-channel discrimination*, Jason Pereira and Stefano Pirandola. Published in Physical Review A **103**, (2021). [10]
- *Characterising port-based teleportation as a universal simulator of qubit channels*, Jason Pereira, Leonardo Banchi, and Stefano Pirandola. Accepted for publication in Journal of Physics A: Mathematical and Theoretical, (2021). [11]
- *Idler-free channel position finding*, Jason Pereira, Leonardo Banchi, Quntao Zhuang, and Stefano Pirandola. Preprinted on the arXiv, (2020). [12]

# Chapter 1

# Introduction

## 1.1 Quantum channels

Classically, when a bit is sent from one place to another, the map describing how it is transformed is called a channel. Physically, an example of a channel is a communication line through which information must be transmitted. In an ideal case, the receiver would receive exactly what the sender is sending (this is called an identity channel), but this is not always the case. Part of the signal could be lost or corrupted during transmission, resulting in the receiver's output being different from the sender's input. For example, an erasure channel is a type of channel that either faithfully transmits a bit or (with some probability) transmits an erasure state, which carries no information about the input bit other than that it was lost. Another example of a classical channel is a bit-flip channel: for such channels, a bit with value 0 is mapped to a bit with value 1 with probability $p_{0\to1}$, a bit with value 1 is mapped to a bit with value 0 with probability $p_{1\to0}$, and the bit is faithfully transmitted in all other cases. Channels can also describe how the value of a bit changes over time. An example is information storage on a disk. If the information on a disk is slowly being corrupted over time, due to, for instance, physical wearing-down of the disk or the disk being stored in an area with a magnetic field, the input-output relations between the information that was originally stored on the disk and the information that would be read from the disk at some later point in time define a classical channel. Classical channels can be symmetric (meaning that bits of either value are affected identically) or not. More complicated channels can have a memory [13]; for a memory channel, the input-output map for the $i$-th transmission may depend on the $(i-1)$-th transmission (rather than being independent and identical for each channel use).

The maximum rate at which information can be reliably transmitted through a classical channel

(in terms of bits per transmission) is called the classical channel capacity. The channel capacity for a number of basic channels is well-known [13, 14].

Quantum channels are the quantum analogue of classical channels; they are mathematical objects describing the transformation of a quantum state [15]. Any map that sends valid quantum states to valid quantum states is a quantum channel. As such, they model a vast number of physical processes, and their study is of great interest in the field of quantum information. A relevant example for the field of quantum communications is the optical fibre connecting two parties in a communication scheme.

Quantum channels are more varied than classical channels, since there are more basic transformations that can be enacted on a quantum state than on a bit. For instance, whilst a bit can be lost with some probability or flipped from 0 to 1 (or vice versa) with some probability, a qubit can undergo the same transformations, but it can also have its phase changed or be transformed into some arbitrary superposition of the 0 and 1 states. Any transformation that can be applied to a bit by a classical channel can be applied to a qubit by a quantum channel, along with a number of transformations that cannot be applied to a bit.

As well as there being a broader variety of quantum channels than classical channels, there are also a larger number of quantities of interest for quantum channels [16]. As well as the classical channel capacity, as defined for classical channels (the maximum rate at which classical information can be faithfully transmitted), quantum channels also have a secret key capacity, which is the maximum rate at which private key distribution can be carried out over such a channel (without an eavesdropper gaining information about the key whilst remaining undetectable). Quantum channels also have a quantum capacity, which is the maximum rate at which they can reliably transmit quantum information. These capacities can be calculated assuming the aid of one or two-way classical communications. The broadest definition of the quantum capacity is the two-way capacity, which allows unlimited local operations and classical communications (LOCCs) between channel uses [5]. The maximum rate at which entanglement can be distributed over a quantum channel (the entanglement-distribution capacity) is also of interest. Accordingly, quantum channels are less well understood in general than classical channels.

The classical channel capacity is of interest for applications such as sending classical information over quantum networks, where the goal is to transfer as much information as possible. The classical capacity of a wide variety of channels, including the quantum erasure channel [17], the phase erasure channel [16], and the generalised Pauli channels [18], are known. Classical information transfer can be aided by pre-shared entanglement; in this case, we must consider the

entanglement-assisted classical capacity. This is the relevant parameter for applications such as super-dense coding [19].

The secret key capacity is of interest for quantum key distribution (QKD) scenarios [20]. It gives the best possible secret key rate achievable by any possible QKD protocol over a given channel. As such, it can be used to quantify how well a protocol performs, in terms of how close it comes to achieving the maximum possible key rate. The secret key capacity is less than or equal to the classical channel capacity, since a secret key is a particular type of classical information that must be distributed from sender to receiver.

The entanglement-distribution capacity is important because pre-shared entanglement is an important resource for a variety of protocols, including super-dense coding, quantum teleportation and QKD. It is equal to the two-way quantum capacity, which is the maximum rate at which quantum information can be transferred over a quantum channel (with unlimited LOCCs between channel uses). The quantum capacity is useful for scenarios in which the goal is to distribute quantum information over a quantum channel or network. These quantities are less than the secret key capacity, since a maximally entangled pair of qubits can be used to faithfully send a bit from the sender to the receiver, without an eavesdropper being able to obtain any information about it.

A hierarchy of the various two-way bounds was given by Pirandola et al. in Ref. [6].

## 1.2 Channel discrimination and parameter estimation

Suppose we are presented with a black box that contains some channel from a set of possible options. Quantum channel discrimination is, as the name suggests, the task of determining which channel we have. Specifically, by sending probes into the black box, collecting the outputs, carrying out operations on them, and carrying out measurements, we want to determine which channel is in the box. The method by which we attempt to discriminate between the channels is called our protocol. More generally, we can refer to any algorithm by which we may attempt to achieve a task (such as secret key distribution or parameter estimation) as a protocol for that task. The probability of us correctly guessing which channel we have is called the success probability and the probability of us guessing the channel incorrectly is called the error probability. These probabilities obviously depend on the protocol we use. If we are allowed to probe the box an unlimited number of times, we can always find a protocol with an error probability that tends to $0$ with the number of channel uses. Suppose, however, that we are only allowed a finite number of channel uses (which we will sometimes refer to as rounds of the protocol), $N$. We then want to find the

protocol that maximises the success probability for a given $N$.

In fact, there are two major routes that we can go down. We can construct a protocol that minimises the error probability - this is called minimum error discrimination - or we can construct a protocol that never misidentifies the channel when it succeeds but has some chance of failing - this is called unambiguous discrimination [21]. We will focus on minimum error discrimination.

The most general $N$-round protocol we can construct consists of us preparing some initial quantum state, probing the channel with some part of it, and then carrying out arbitrary quantum operations between rounds [2, 22]. This can include measurements. This has the structure of a quantum comb [23].

Adaptive protocols, where subsequent probes can be dependent on measurements carried out on previous probes, have proven to be more powerful than non-adaptive protocols [24]. Harrow et al. found a pair of channels that cannot be perfectly discriminated between by any non-adaptive protocol with a finite number of rounds, but that can be perfectly distinguished between by a 2-round adaptive protocol. This has necessitated the study of the most general adaptive protocols, in order to establish ultimate bounds on the minimum achievable error probability for quantum channel discrimination [25]. Quantum channel simulation and protocol stretching (see Chapter 2) are powerful tools for establishing these ultimate bounds [2, 5, 26–28].

One well-studied task within quantum channel discrimination is binary discrimination. This is discrimination between only two possible channels. For equal prior probabilities, the error probability in distinguishing between the two possible output states of a discrimination protocol is known exactly, in terms of the trace norm (the Helstrom bound) [29].

A related task is that of parameter estimation. In this scenario, the possible channels are continuously parametrised by a variable, $\theta$, and our task is to estimate the value of $\theta$ as precisely as possible. We again have a quantum comb structure (in the most general case). Channel simulation and protocol stretching can once again be used to reduce the protocol to a block form in a number of important cases [30].

### 1.2.1 Applications

One application of binary discrimination is in quantum illumination [31–44], where a device must discriminate between the presence and the absence of an object. Another application is the protocol of quantum reading, in which classical information that is encoded in the reflectivity of memory cells is read off by quantum states [25, 45]. By describing the input-output relations of the probing systems as quantum channels, both of these tasks can be treated as problems of quan-

tum channel discrimination (with the different channels corresponding to the different possible outcomes).

An example in quantum communications is quantum hacking [20, 46, 47], where Eve may wish to determine aspects of the settings of Alice's and Bob's devices, by probing them via side-channels. If the settings affect the quantum channel that the probes would pass through, Eve could carry out a discrimination protocol between the possible channels and therefore the possible settings.

Quantum metrology [48] uses quantum states to make more precise measurements than are possible classically or makes equally precise measurements whilst using less energy. Many tasks within quantum metrology involve parameter estimation. For instance, suppose we want to find the transmissivity of a delicate sample. Our task would then be to probe it with a limited number of photons in order to determine the transmissivity: this is parameter estimation.

Spedalieri et al. considered the use of quantum states of light (correlated thermal states and superpositions of number states) to probe delicate biological samples in order to both estimate the transmissivity (parameter estimation) and discriminate between the presence and absence of bacteria (channel discrimination) [3].

## 1.3 Motivation and structure of the thesis

Since almost any physical process can be regarded as a quantum channel, discrimination between quantum channels is a task with relevance to many fields of science. As a result, protocols for channel discrimination have a very broad applicability. Any measurement task can be reduced to a task of parameter estimation and any scenario in which we want to decide which of a set of possible physical processes is occurring can be reconstructed into a task of channel discrimination.

Quantum protocols have been shown to be capable of improving the precision of measurements [48], the amount of classical information that can be read out from a memory cell (for fixed energy) [45], the energy required to detect targets [31, 32, 41], and the energy required to detect the presence of bacteria in a sample or discriminate between two types of bacteria [3].

As a result, a key question is: what is the ultimate performance limit of a quantum protocol? If we can establish tight bounds on the best possible success probability or precision of a quantum protocol, we can assess which tasks might be further improved by better protocols. Since quantum hacking can be modelled as a task of channel discrimination (or parameter estimation, in the case of a continuous alphabet), it is also vital from an information security perspective to assess to

what degree an eavesdropper can exploit vulnerabilities, taking into account how technology and discrimination protocols might improve in the future. This will become especially important as classical cryptography becomes insecure (due to Shor's algorithm [49] and the development of quantum computers), requiring either the development of post-quantum cryptography [50] or for quantum communication technology to be rolled out around the world.

Channel simulation and teleportation stretching are powerful techniques that allow the formulation of ultimate bounds on any discrimination or parameter estimation protocol. One issue they currently have is that certain channels do not commute with standard teleportation, meaning that protocols involving these channels cannot be stretched in the same way.

In this work, we want to improve the power of the techniques of channel simulation and teleportation stretching by utilising alternatives to standard teleportation. We aim to thereby tighten existing bounds on quantum channel discrimination, especially those relating to important channels, such as the phase-insensitive Gaussian channels and the amplitude damping (AD) channel.

This thesis will present a number of results in the fields of quantum channel simulation and channel discrimination, which have then been applied to physical scenarios. We improve the characterisation of qubit port-based teleportation (PBT) and the channels that it simulates, hence making it more useful as a tool for qubit channel simulation. We also look in some depth at a quantum hacking attack.

Chapter 2 introduces the reader to some basic concepts, relating to quantum information, that will be useful going forwards. We start by discussing some basic differences between DV and CV quantum information and then go on to define some important quantities, such as the trace norm and the Bures fidelity, and explain their properties. We then give a very brief introduction to the formalism of Gaussian quantum information and present some helpful formulae. Finally, we explain how the techniques of channel simulation and protocol stretching can be applied, by way of an example.

Channel position finding (CPF) is a little-studied subfield of channel discrimination. CPF is the task of finding the position of a target channel amongst a sequence of background channels. In Chapter 3, we approach CPF in two different ways. We start by considering a scenario in which we must discriminate between pure loss channels, with different transmissivities, using limited energy. In this case, it is not possible for us to stretch the protocol using standard teleportation stretching; instead we consider a specific input state for a one-shot protocol. More specifically, we find the output fidelity in an idler-free setting (i.e. without any entanglement between the signal states and any states that do not pass through the channels), using an input state that has correlations

between the signal states for each channel. We then consider a scenario that is on the other end of the spectrum: we allow the input states to have unbounded energy and unlimited entanglement with an idler and calculate ultimate bounds on the error probability of any possible protocol. In this case, the task is environment localisation (CPF over a sequence of phase-insensitive Gaussian channels with fixed transmissivity).

In order to carry out teleportation stretching on a range of protocols, we would like to be able to simulate channels that do not commute with standard teleportation. An important example is the AD channel. One option is PBT. Chapter 4 develops PBT as a useful tool for channel simulation. We calculate explicit analytical expressions for the channel enacted by PBT for an arbitrary resource state. We then characterise the PBT protocol (with the square-root measurement) itself, by finding the channel from a resource state to the Choi matrix of the qubit channel it enacts. We find improved resource states for AD channel simulations.

Chapter 5 applies the results of the previous chapter in order to obtain tighter lower bounds on the error probability of an AD channel discrimination protocol. The upper bound on the optimal error probability is also tightened, and a bound based on the quantum Cramér-Rao bound (which is approximate for low numbers of channel uses) is found to approach our new upper bound for large numbers of channel uses. We calculate the diamond norm between any two AD channels and thereby find the ultimate one-shot error probability for a discrimination protocol. We then apply our bounds to physical scenarios involving quantum hacking and biological quantum sensing.

Chapter 6 looks in detail at a specific quantum hacking attack on a coherent state CV-QKD protocol. We calculate the key distribution rate for a coherent state protocol where the sender has a side-channel in her device that allows Trojan states, with a bounded mean photon number, to enter. We see that the key rate rapidly drops, even when the eavesdropper's mean photon number is very low. Finally, we discuss how the side-channel attack could be mitigated using a passive architecture with active monitoring to characterise any vulnerabilities.

Chapter 7 summarises our results. We present our conclusions and discuss the direction that further study could take.

# Chapter 2

# Preliminaries

In this chapter, we introduce some of the basic concepts and techniques that will be used throughout this work. We start by briefly describing some differences between discrete and continuous variable states. We go on to define some quantities of interest, relating to quantum states and channels, that are important in quantum channel discrimination. We then give a very brief introduction to those aspects of Gaussian quantum information that will be of relevance to the reader, focusing on the unitaries, channels, and formulae that we will need in Chapters 3 and 6. Finally, we introduce and discuss the techniques of channel simulation and protocol stretching. We explain how they can help with calculation of channel quantities, using the example of the two-way entanglement distribution capacity, and then describe some other applications of the technique.

## 2.1 Discrete and continuous variables

The systems studied in quantum information science can be divided into two broad categories: discrete variable (DV) systems and continuous variable (CV) systems. DV systems encompass quantum states with finite-dimensional Hilbert spaces (e.g. qubits or qutrits), whilst CV quantum information science deals with infinite-dimensional Hilbert spaces.

Physical processes such as the decay of an atom from one energy level to another can be well-modelled as the evolution of DV systems (although the environment with which the atom is coupled is infinite-dimensional). Further, when operating in the number state basis, DV systems can be used to model low energy scenarios. This is because truncation of the higher energy states will have little effect on the states, in this case. For quantum communications applications, the states of a qubit can be encoded by photonic states. Implementations include encoding the information in the presence or absence of a photon, in the phase of a photon (this is called time-

bin encoding; the state of the qubit determines whether the photon arrives early or late) or in the polarisation of a photon.

In quantum communications, DV systems are important because a lot of existing quantum key distribution (QKD) protocols use qubits as the signal states. Examples include BB84 [51] and B92 [52]. In fact, many existing experimental implementations of QKD use DV protocols [53].

Important qubit channels include the amplitude damping (AD) channel and its generalisation - the generalised AD channel, the Pauli channels, and the quantum erasure channel [15].

The (qubit) Pauli channels carry out each of the Pauli unitary transformations on the input state with some probability of each. The erasure channel either faithfully transmits the input state or replaces it with an erasure state (which is orthogonal to both $|0\rangle$ and $|1\rangle$), indicating that the input state was lost. These are all relatively simple channels and are well-studied.

The AD channel is a channel that faithfully transmits the state $|0\rangle$, but causes the state $|1\rangle$ to decay to $|0\rangle$ with some probability. It is a good model for a variety of scenarios in which a quantum state may decay from a higher energy state to a lower energy state, such as when a particle decays from an excited state to the ground state. It can also model low energy imaging scenarios, in which a probe with an average photon number of much less than a photon per mode is used to image a sample. Calculating channel quantities for the AD channel is often complicated, due to its asymmetry. Notably, it is not teleportation-covariant and so cannot be simulated by the standard teleportation protocol, using its Choi matrix as a resource. The generalised AD channel is analogous to a (CV) thermal loss channel, in the same way that an AD channel is analogous to a (CV) pure loss channel.

CV quantum information science encompasses all scenarios in which the variables are not restricted to a finite set of values [54]. Examples of continuous variables that can parametrise quantum states include position, momentum and energy. Note that these can all be discrete variables (or very closely modelled by discrete variables), depending on the scenario. For instance, energy can be discrete for an electron in an atom, which may be restricted to a fixed set of energy levels, but the energy of a free particle is a continuous variable. In quantum communications scenarios (and also in many quantum metrology scenarios), the CV states represent bosonic modes of light. In this case, the continuous variables are the quadratures of the electric field for each mode.

CV systems can also be used for quantum communications, and a number of CV-QKD protocols exist [20, 55]. Some of these have been experimentally implemented [56]. An advantage of using CV systems for quantum information (over using DV systems) is that a lot of existing technology can be re-purposed for CV-QKD: homodyne and heterodyne detectors already exist and

are used in classical communications. It is also easy to generate coherent states (which many CV protocols use), whilst it is difficult to produce reliable single-photon sources and detectors. Often, DV-QKD systems approximate single-photon sources using strongly attenuated lasers, and this opens up the systems to quantum hacking attacks, such as photon-number splitting (PNS) [57,58], since strongly attenuated lasers will produce multiple-photon signal states some proportion of the time.

Gaussian states are an important subcategory of CV states [55, 59, 60]. They are completely characterised by the expected values of their quadratures and the quadrature covariance matrix (CM). This is an important quality, as it greatly simplifies calculations involving them, by allowing us to work in the phase space of bosonic modes, in which Gaussian unitaries, channels, and measurements take very simple forms. As a result, a number of tools have been developed for Gaussian states. See Section 2.3 for more information.

## 2.2 Quantities of interest

We will begin by discussing some basic quantities of interest that we can calculate for a quantum state. These introductory notions can be found in Ref. [15].

A pure $d$-dimensional quantum state can be described by a $d$-dimensional vector (either a bra or a ket). On the other hand, a mixed state requires a $d$ by $d$ dimensional density matrix, $\rho$, to fully characterise it. In order to represent a valid quantum state, $\rho$ must be positive semidefinite. As a result, an important quantity of a quantum state is its purity. This is given by

$$P(\rho) = \mathrm{Tr}[\rho^2]. \tag{2.1}$$

$P$ takes values between $\frac{1}{d}$ (for a maximally mixed state) and 1 (for a pure state).

The Von Neumann entropy of a quantum state is the quantum analog of the classical Shannon entropy. It is given by

$$S(\rho) = -\mathrm{Tr}[\rho \log \rho], \tag{2.2}$$

where the base of the logarithm determines the units. Generally, we will want to work in bits, so we will use base 2 for our logarithms.

The Shannon entropies of a pair of classical random variables give rise to its mutual information, which is a measure of the amount of information that one variable encodes about the other. Similarly, the Von Neumann entropies of a pair of quantum states give rise to its quantum mutual

information, which is a measure of the correlations between the states.[1] However, more relevant for our purposes is the Holevo bound, which gives an upper bound on the accessible classical information that a quantum state encodes about a classical variable. In other words, there is no possible measurement on a state $\rho$ that gives more information about a classical variable $X$ than $H(\rho)$, where

$$H(\rho) = S(\rho) - \sum_i p(X = x_i) S(\rho_i). \tag{2.3}$$

$\rho_i$ is the state $\rho$ conditioned on $X$ taking the value $x_i$. If $X$ has a continuous probability distribution, the sum in Eq. (2.3) becomes an integral [55].

A necessary condition for the $d$-dimensional state $\rho$ to be separable, with respect to a bipartition into a $d_1$-dimensional system and a $d_2$-dimensional system (where $d_1 + d_2 = d$), is that it has a positive partial transpose. This means that the partial transpose of $\rho$ must also be a valid density matrix in order for $\rho$ to be separable. The partial transpose of $\rho$ can be obtained by dividing it into a $d_1$ by $d_1$ block matrix, with blocks that each have dimension $d_2$ by $d_2$, and then transposing each block. For the cases in which $\rho$ represents a two-qubit state or a qubit-qutrit state, this condition is also sufficient to show separability [61].

### 2.2.1 Useful quantities for state discrimination

Distinguishing between two possible quantum states is an important task in quantum information. One measure of the distance between two quantum states, $\rho_1$ and $\rho_2$, is the trace norm (also called the Schatten 1-norm), which is defined by

$$\|\rho_1 - \rho_2\|_1 = \mathrm{Tr}(|\rho_1 - \rho_2|) = \mathrm{Tr}\left(\sqrt{(\rho_1 - \rho_2)(\rho_1 - \rho_2)^\dagger}\right). \tag{2.4}$$

The trace norm is a particularly important metric because it gives the optimal success probability, $p_{\text{success}}$, for a measurement discriminating between the two states $\rho_1$ and $\rho_2$. The Helstrom bound [29] states that

$$p_{\text{success}} = \frac{1}{2} + \frac{1}{4}\|\rho_1 - \rho_2\|_1. \tag{2.5}$$

The trace norm takes values between 0 and 2. It is invariant under unitary transformations, non-increasing under quantum operations, convex, and it obeys the triangle inequality, meaning

$$\|\rho_1 - \rho_2\|_1 + \|\rho_2 - \rho_3\|_1 \geq \|\rho_1 - \rho_3\|_1. \tag{2.6}$$

[1] The quantum mutual information of the state $\rho_{12}$ (with respect to systems 1 and 2) is given by $S(\rho_1) + S(\rho_2) - S(\rho_{12})$.

Combining this with the Helstrom bound, we can also write

$$p^{12}_{\text{success}} + p^{23}_{\text{success}} - \frac{1}{2} \geq p^{13}_{\text{success}}, \tag{2.7}$$

where $p^{ij}_{\text{success}}$ is the probability of successfully discriminating between states $\rho_i$ and $\rho_j$.

Whilst considering the Schatten norms, it is worth mentioning the Schatten $\infty$-norm, $\|\rho_1 - \rho_2\|_\infty$, which is simply the largest eigenvalue of $|\rho_1 - \rho_2|$. Like the 1-norm, the $\infty$-norm takes values between 0 and 2, is invariant under unitary transformations, is convex, and obeys the triangle inequality (although it is neither non-increasing nor non-decreasing under quantum operations).[2]

The Bures fidelity is another measure of the closeness of two quantum states. It is defined by

$$F(\rho_1, \rho_2) = \text{Tr}\left(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right). \tag{2.8}$$

Note that the quantum fidelity is sometimes defined as the square of the expression given here. Eq. (2.8) takes a simpler form if one or both of the states are pure. If $\rho_1 = |\sigma_1\rangle\langle\sigma_1|$ is a pure state, we can write

$$F(\rho_1, \rho_2) = \sqrt{\langle\sigma_1|\rho_2|\sigma_1\rangle} \tag{2.9}$$

and if both $\rho_1$ and $\rho_2$ are pure, we can write

$$F(\rho_1, \rho_2) = |\langle\sigma_1|\sigma_2\rangle|, \tag{2.10}$$

where $\rho_1 = |\sigma_1\rangle\langle\sigma_1|$ and $\rho_2 = |\sigma_2\rangle\langle\sigma_2|$. It takes values between 0 and 1. It is invariant under unitary transformations, non-decreasing under quantum operations, and concave. Another important property of the fidelity is that it is multiplicative with respect to tensor products. It is this property that often makes it much easier to calculate, for tensor product states, than the trace norm. It is possible to bound the trace norm in terms of the fidelity; we find

$$2(1 - F(\rho_1, \rho_2)) \leq \|\rho_1 - \rho_2\|_1 \leq 2\sqrt{1 - F(\rho_1, \rho_2)^2}, \tag{2.11}$$

with the upper bound becoming an equality when both states are pure.

We now consider useful bounds on the task of parameter estimation. We can give the Bures distance, $d_B$, in terms of the fidelity:

$$d_B(\rho_1, \rho_2) = \sqrt{2(1 - F(\rho_1, \rho_2))}. \tag{2.12}$$

[2]For a simple counterexample, consider the channel (which we denote as $\mathcal{C}$) with Kraus operators $\{|0\rangle\langle 0| + \frac{1}{\sqrt{2}}|1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3|, \frac{1}{\sqrt{2}}|0\rangle\langle 1|\}$, acting on a 4-dimensional input Hilbert space. Let $\sigma_1 = \text{diag}(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$ and let $\sigma_2 = \text{diag}(0, 1, -\frac{1}{2}, -\frac{1}{2})$. Then, $\|\sigma_1\|_\infty < \|\mathcal{C}(\sigma_1)\|_\infty$ and $\|\sigma_2\|_\infty > \|\mathcal{C}(\sigma_2)\|_\infty$.

This is important because we can calculate the quantum Fisher information (QFI) using the Bures distance. If a quantum state, $\rho_\theta$, encodes a parameter $\theta$, we can write

$$\mathrm{QFI}_\theta = \frac{4d_B^2(\rho_\theta, \rho_{\theta+\delta\theta})}{\delta\theta^2}. \tag{2.13}$$

The QFI is additive with respect to tensor products. Now suppose we have $N$ copies of the state $\rho_\theta$ and want to estimate $\theta$. The achievable variance for our measurement, $\mathrm{Var}(\theta)$, is lower bounded in the asymptotic (in terms of $N$) case by the quantum Cramér-Rao bound [62, 63]. This is given by

$$\mathrm{Var}(\theta) \geq (N\mathrm{QFI}_\theta)^{-1}. \tag{2.14}$$

### 2.2.2 Useful quantities for channel discrimination and parameter estimation

We now look at quantities that can be used to determine how well we can discriminate between two quantum channels or to find the variance of an estimator for estimating a parameter encoded in a quantum channel.

Let us begin by defining the Choi matrix of a channel. For a DV channel, $\mathcal{C}$, acting on $d$-dimensional input states, the Choi matrix is defined as

$$(\mathcal{I}_I \otimes \mathcal{C}_S)\left[B_{IS}^d\right], \tag{2.15}$$

$$B_{IS}^d = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_I |i\rangle_S \tag{2.16}$$

where $S$ labels the system that passes through the channel, $I$ labels an idler system, and $\mathcal{I}_I$ is the identity channel on the idler mode. In other words, the Choi matrix is the output state if the channel is applied to half of a generalised Bell (maximally entangled) state. We can also consider other definitions based on different choices for the initial generalised Bell state. The Choi matrix of a channel completely characterises it: a channel is one to one with its Choi matrix (for a fixed choice of initial generalised Bell state).

For a (one-mode) CV channel, the asymptotic Choi matrix is the infinite-squeezing limit of a sequence of two-mode squeezed vacuum states (see Subsection 2.3.1), one mode of which has been passed through the channel.

Since every channel is uniquely defined by its Choi matrix, it would be natural to use the trace norm between the Choi matrices of two channels as a measure of the distance between them.[3]

[3]We will sometimes refer to the trace norm between two channels; this should be understood as the trace norm between the Choi matrices of the channels.

This is, indeed, one way of quantifying the distance between a pair of channels, however in many circumstances, we are interested in finding an upper bound on the distinguishability of quantum channels (equivalently, a lower bound on the error probability of discriminating between them). In the one-shot scenario (i.e. the case in which we are allowed only one channel use), this reduces to finding the input state that maximises the trace norm between channel outputs. Except for in special cases, this is not the maximally entangled state.

Instead, we define the diamond norm[4] between channels $\mathcal{C}^1$ and $\mathcal{C}^2$ as

$$\left\|\mathcal{C}^1 - \mathcal{C}^2\right\|_\diamond = \sup_{\sigma^{\text{in}}} \left\| \left(\mathcal{I}_I \otimes (\mathcal{C}^1_S - \mathcal{C}^2_S)\right) \left[\sigma^{\text{in}}_{SI}\right] \right\|_1, \tag{2.17}$$

where the supremum is taken over all valid, pure input states on the signal and idler systems. One might wonder why the idler system is required; without an idler, the phase change enacted by the channel would not be measurable, because the global phase of a quantum state cannot be measured, only the relative phase of one state to another (in this case, the relative phase of the signal and the idler). Note that we do not lose generality by restricting to pure input states, because the convexity of the trace norm guarantees that there exists a pure state that maximises the trace norm. The diamond norm therefore gives the ultimate one-shot bound on the probability of successfully discriminating between two channels.

The diamond norm can be found numerically using semidefinite programming [64]. Let us call the Hilbert space of the input states $\mathcal{X}$ and the Hilbert space of the output states $\mathcal{Y}$ and define $J(\mathcal{C}^1_S - \mathcal{C}^2_S)$ as

$$J(\mathcal{C}^1_S - \mathcal{C}^2_S) = d \left\| \left(\mathcal{I}_I \otimes (\mathcal{C}^1_S - \mathcal{C}^2_S)\right) \left[B^d_{IS}\right] \right\|_1. \tag{2.18}$$

The (dual) semidefinite programming problem for finding the diamond norm is:

$$\begin{aligned} \text{minimise}: \quad & \frac{1}{2}\|\text{Tr}_{\mathcal{Y}}(Y_0)\|_\infty + \frac{1}{2}\|\text{Tr}_{\mathcal{Y}}(Y_1)\|_\infty \\ \text{subject to}: \quad & \begin{pmatrix} Y_0 & -J(\mathcal{C}^1_S - \mathcal{C}^2_S) \\ -J(\mathcal{C}^1_S - \mathcal{C}^2_S) & Y_1 \end{pmatrix} \geq 0, \end{aligned}$$

where $Y_0$ and $Y_1$ are positive operators on $\mathcal{X} \otimes \mathcal{Y}$. Nechita et al. [1] used the semidefinite programming problem to prove the following bounds on the diamond norm:

$$\frac{1}{d}\|J(\mathcal{C}^1_S - \mathcal{C}^2_S)\|_1 \leq \|\mathcal{C}^1_S - \mathcal{C}^2_S\|_\diamond \leq \|\text{Tr}_S\left(|J(\mathcal{C}^1_S - \mathcal{C}^2_S)|\right)\|_\infty \tag{2.19}$$

[4]Some works refer to the trace/diamond distance rather than the trace/diamond norm. We avoid these terms in this work in order to prevent confusion, because the trace distance is sometimes defined as being half of the trace norm.

Another related quantity of interest is the energy-constrained diamond norm. This is defined slightly differently by Pirandola et al. [5] than by Shirokov [65] and Winter [66]. Here we present the form used by Shirokov and Winter. The energy-constrained diamond norm is defined by

$$\left\|\mathcal{C}^1 - \mathcal{C}^2\right\|_{\diamond E} = \sup_{\sigma^{\text{in}} \in \mathcal{D}_E} \left\| \left(\mathcal{I}_I \otimes (\mathcal{C}_S^1 - \mathcal{C}_S^2)\right) \left[\sigma_{SI}^{\text{in}}\right] \right\|_1, \tag{2.20}$$

$$\mathcal{D}_E = \{\sigma^{\text{in}} \ : \ \text{Tr}(\hat{H}_S \sigma^{\text{in}}) \leq E\}, \tag{2.21}$$

where $\hat{H}_S$ is the Hamiltonian for the input system.

## 2.3 Gaussian quantum information

As mentioned in Chapter 1, the study of Gaussian states and channels is an important topic within CV quantum information. Here we will give a broad overview of Gaussian states and introduce some basic notions related to Gaussian quantum information. For a more in-depth introduction to the field, see the reviews by Weedbrook et al. [55], Adesso et al. [59], and Olivares [60].

A bosonic system is a collection of modes of a quantised field. The most important example in quantum information science is the quantised electromagnetic field, whose modes represent radiation modes of light (characterised by a frequency and a direction). A state of the system can be described by the number of particles (photons) in each of the modes; this is the number state or Fock basis representation. Since there is no upper limit on the number of particles in a mode, this is an infinite-dimensional basis.

Each mode can be modelled as a quantum harmonic oscillator with its own annihilation and creation operators, $\hat{a}^\dagger$ and $\hat{a}$, defined by

$$\hat{a}_i^\dagger \left|n\right\rangle_i = \sqrt{n+1}\left|n+1\right\rangle_i, \ \hat{a}_i \left|n\right\rangle_i = \sqrt{n}\left|n-1\right\rangle_i, \ \hat{a}_i \left|0\right\rangle_i = 0, \tag{2.22}$$

where $i$ labels the mode on which the operators act. These definitions explain the names of the two types of operator: a creation operator acts on a mode by adding a particle to it, whilst an annihilation operator removes one. The operators obey the commutation relation

$$[\hat{a}^\dagger, \hat{a}] = -1. \tag{2.23}$$

Note that $\hat{a}$ and $\hat{a}^\dagger$ are not Hermitian operators and so do not represent observables of the system.

By combining the annihilation and creation operators for a mode, we get the number operator for that mode,

$$\hat{a}^\dagger \hat{a} = \hat{n}. \tag{2.24}$$

The eigenstates of the number operator are the Fock states, which make up the Fock basis. The number operator acts on Fock states to give

$$\hat{n}\left|n\right\rangle = n\left|n\right\rangle, \tag{2.25}$$

where $n$ is the number of particles in the mode (hence the name of the operator).

We can also define the quadrature field operators as

$$\hat{q} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2\kappa}}, \qquad\qquad \hat{p} = -i\frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2\kappa}}, \tag{2.26}$$

where $\kappa$ is a constant that defines the vacuum (or shot) noise of the system (the variance of the quadratures when there are no particles in a mode). Common values of $\kappa$ are $\frac{1}{2}$, corresponding to a vacuum noise of 1, and 1, corresponding to a vacuum noise of $\frac{1}{2}$. The vacuum noise will be specified whenever we consider Gaussian states; in this chapter we will work in a $\kappa$-independent setting (i.e. without setting the vacuum noise). Many of the $\kappa$-independent formulae given here are presented in Ref. [20]. $\hat{q}$ and $\hat{p}$ are Hermitian operators and so they represent real observables of the system. $\hat{q}$ and $\hat{p}$ are often referred to as the position and momentum operators respectively, since these are what they represent in the quantum harmonic oscillator model, although, in the optical case, they represent orthogonal components of the electromagnetic field.

We would now like to switch to an alternative representation of the bosonic system that is easier for us to work with. This is done by mapping the density matrix of a state to an equivalent quasi-probability distribution - called a Wigner function - on a real symplectic space. We start by defining the vector $\hat{x}$ as

$$\hat{x} = (\hat{q}_1, \hat{p}_1, \ldots \hat{q}_N, \hat{p}_N)^T, \tag{2.27}$$

where $N$ is the number of modes in the system. We then define the Weyl operator,

$$D(\xi) = e^{i\hat{x}^T\Omega\xi}, \tag{2.28}$$

$$\Omega = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.29}$$

where $\Omega$ is called the symplectic form and $\xi \in \mathbb{R}^{2N}$ is a vector. We then define the characteristic form of a density matrix $\rho$ as

$$\chi(\xi) = \mathrm{Tr}[\rho D(\xi)]. \tag{2.30}$$

The Fourier transform of the characteristic function then gives us the Wigner function of the state represented by $\rho$:

$$W(x) = \int_{\mathbb{R}^{2N}} \frac{d^{2N}\xi}{(2\pi)^{2N}} e^{-ix^T\Omega\xi}\chi(\xi), \tag{2.31}$$

where $x \in \mathbb{R}^{2N}$ is a vector of eigenvalues of the quadrature operators (i.e. of $\hat{x}$). Hence, $W(x)$ is a quasi-probability distribution over the possible quadrature values that a state can take.

Gaussian states are defined by having a Gaussian Wigner function,

$$W^{\text{Gaussian}}(x) = \frac{1}{(2\pi)^N \sqrt{\det[V]}} e^{-\frac{1}{2}(x-\bar{x})^T V^{-1}(x-\bar{x})}, \tag{2.32}$$

where $\bar{x}$ is the first moments vector (made up of the expectation values of the quadratures), defined as

$$\bar{x} = \text{Tr}(\hat{x}\rho), \tag{2.33}$$

and $V$ is the covariance matrix. The elements of the covariance matrix (also known as the matrix of second moments) are defined by

$$V_{ij} = \frac{1}{2}\text{Tr}(\{\hat{x}_i - x_i, \hat{x}_j - x_j\}\rho), \tag{2.34}$$

where $\{\cdot,\cdot\}$ denotes the anticommutator. Gaussian states are therefore completely characterised by their first moments vector and their covariance matrix. This means that we only have to deal with a $2N$-dimensional vector and a $2N$ by $2N$ covariance matrix rather than an infinite-dimensional system.

As an example, we can consider one of the most important single-mode Gaussian states: the vacuum state. This is the state of the system when there are no particles in the mode. It has the first moments vector

$$x = (0, 0)^T \tag{2.35}$$

and its covariance matrix is

$$V = \frac{1}{2\kappa} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.36}$$

Thermal states are similar to vacuum states: they also have no non-zero components in their first moments vectors and their covariance matrices are also obtained by multiplying the identity matrix by a multiplicative factor. They have covariance matrices of the form

$$V = \frac{2\bar{n} + 1}{2\kappa} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.37}$$

where $\bar{n}$ is the mean photon number of the state.

### 2.3.1 Gaussian unitaries, channels and measurements

A unitary that maps Gaussian states to other Gaussian states is called a Gaussian unitary. A Gaussian unitary, $U$, which enacts the transformation $\rho \to U\rho U^\dagger$, can be represented in the phase

space by a symplectic matrix, $S$, and a vector, $d$. $S$ transforms the first and second moments of $\rho$ according to

$$\bar{x} \to S\bar{x} + d,\ V \to SVS^T. \tag{2.38}$$

Some of the most important Gaussian unitaries are the displacement operator, the beamsplitter operator, and the two-mode squeezing operator.

The displacement operator displaces states in phase space. Its symplectic matrix is the identity matrix and so a displacement operator is defined by its vector $d$, which describes how it translates the quadratures of a state. When applied to a vacuum state, the displacement operator generates coherent states, which are eigenstates of the annihilation operator.

The beamsplitter operation mixes two modes via the symplectic matrix

$$S(\tau) = \begin{pmatrix} \sqrt{\tau}\mathbb{I} & \sqrt{1-\tau}\mathbb{I} \\ -\sqrt{1-\tau}\mathbb{I} & \sqrt{\tau}\mathbb{I} \end{pmatrix}, \tag{2.39}$$

where $\tau$ is the transmissivity of the beamsplitter, which determines the degree to which the two modes are mixed with each other. It ranges from $0$ to $1$. The components of its vector $d$ are all $0$.

The two-mode squeezing operator again acts on two modes and can be used to generate entanglement between them. Its symplectic matrix is

$$S(r) = \begin{pmatrix} \cosh r\mathbb{I} & \sinh r\mathbb{Z} \\ \sinh r\mathbb{Z} & \cosh r\mathbb{I} \end{pmatrix}, \tag{2.40}$$

where $\mathbb{Z}$ is the Pauli Z-matrix. The components of its vector $d$ are all $0$. By applying the two-mode squeezing operator to a two-mode vacuum state, one can generate a two-mode squeezed vacuum (TMSV), which is one of the most common types of entangled state found in CV quantum information. The parameter $r$ determines the degree of the squeezing and hence the amount of entanglement generated. As $r \to \infty$, we get an unphysical state with infinite entanglement and infinite energy.

Similarly to Gaussian unitaries, we define Gaussian channels as those channels that map Gaussian states to other Gaussian states. A Gaussian channel can be represented by a displacement vector, $d$, and real matrices, $N$ and $T$, that obey the complete positivity condition

$$N + i\Omega - iT\Omega T^T \geq 0. \tag{2.41}$$

Its action on the first and second moments of a Gaussian state is given by

$$\bar{x} \to T\bar{x} + d,\ V \to TVT^T + N. \tag{2.42}$$

Some of the most important Gaussian channels in quantum information are the one-mode Gaussian channels and particularly the phase-insensitive one-mode channels [67]. For this class of channels,

$$T = \sqrt{\tau}\mathbb{I}, \tag{2.43}$$

$$N = \nu\mathbb{I}, \tag{2.44}$$

where $\tau$ and $\nu$ are both positive real numbers. $\tau$ is called the transmissivity and $\nu$ is called the induced noise. By setting both of the components of $d$ to 0, we get three important types of channel. When $0 \leq \tau < 1$, we have a lossy channel. When $\tau > 1$, we have an amplifier channel. For both of these channels (i.e. for any $\tau \neq 1$), we can write

$$\nu = |1 - \tau|\frac{2\bar{n} + 1}{2\kappa},\ \bar{n} \geq 0. \tag{2.45}$$

A lossy channel with $\bar{n} = 0$ is called a pure loss channel and an amplifier channel with $\bar{n} = 0$ is called a quantum-limited amplifier. We have an additive noise Gaussian channel for $\tau = 1$.

A Gaussian measurement, analogously to Gaussian unitaries and channels, is one that, when applied to a Gaussian state, both produces a Gaussian distributed outcome and leaves the unmeasured modes of the system in a Gaussian state. We will look at two very useful Gaussian measurements: homodyne detection and heterodyne detection.

Homodyne detection measures one quadrature of a mode, whilst heterodyne detection measures both, albeit with an extra vacuum noise unit added to the variance of each quadrature. The noise is added because heterodyne measurements are experimentally realised by mixing the mode that is to be measured with a vacuum state using a balanced beamsplitter (i.e. one with a transmissivity of $\frac{1}{2}$) and then homodyning each output state (obtaining the value of one quadrature from one output state and the value of the other quadrature from the other output state).

Suppose we want to measure the first mode of an $N$-mode state. Let the covariance matrix of the state prior to the measurement be

$$V = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix},\ A = \begin{pmatrix} A_{qq} & A_{pq} \\ A_{pq} & A_{pp} \end{pmatrix}, \tag{2.46}$$

where $A$, $B$, and $C$ are all block matrices. $A$ would be the covariance matrix of the first mode if we discarded the remaining modes, $B$ would be the covariance matrix of the remaining $N - 1$ modes if we discarded the first mode, and $C$ describes the correlations between the first mode and the remaining modes. Let the first moments vector of the state be

$$\bar{x} = (q_A, p_A, \bar{x}_B^T)^T, \tag{2.47}$$

where $q_A$ and $p_A$ are real numbers and $\bar{x}_B$ is a $2(N-1)$-dimensional vector. A homodyne or heterodyne measurement will have an outcome that is Gaussianly distributed. A homodyne measurement of the $\hat{q}_A$-quadrature will have a mean of $q_A$ and a variance of $A_{qq}$ and a measurement of the $\hat{p}_A$-quadrature will have a mean of $p_A$ and a variance of $A_{pp}$. A heterodyne measurement will have an outcome drawn from a multivariate Gaussian distribution with a mean value of $(q_A, p_A)^T$ and a covariance matrix of $A + \frac{1}{2\kappa}\mathbb{I}$. If the measurement is carried out on a subset of modes of a multimode system, we may also be interested in how the measurement affects the state of the unmeasured modes. The covariance matrix of the remaining modes will become

$$\tilde{V} = B - C^T \tilde{A} C, \tag{2.48}$$

where $\tilde{A}$ is given by $\mathrm{diag}(A_{qq}^{-1}, 0)$ for a homodyne measurement of the $\hat{q}_A$-quadrature, $\mathrm{diag}(0, A_{pp}^{-1})$ for a homodyne measurement of the $\hat{p}_A$-quadrature, and $(A + \frac{1}{2\kappa}\mathbb{I})^{-1}$ for a heterodyne measurement.

One more important measurement that is worth mentioning, despite being non-Gaussian, is the photon counting measurement. The probability distribution of the measurement outcomes can be obtained using a Fock basis representation of the measured state. In particular, when applied to a thermal state, with its covariance matrix as given in Eq. (2.37), the expected value of the measurement is $\bar{n}$ and the probability of a measurement outcome of $n$ photons is

$$p(n) = \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}}. \tag{2.49}$$

Note that the outcomes of the photon counting measurement take discrete (integer) values.

Many CV protocols have been developed for which the signal states are Gaussian. Depending on the protocol, the signal states may be coherent states [68] or one-mode squeezed states [69,70]. The detection may be either homodyne [71] or heterodyne [7]. Since the signal states involved, the operations applied to the states and the measurements used are all Gaussian, we call such protocols fully Gaussian. It has been shown that collective Gaussian attacks are optimal for fully Gaussian protocols [72] and, as a result, the security analysis of fully Gaussian protocols can be reduced to a much simpler form.

The bosonic channels that signals are sent along for these protocols are generally modelled as single-mode, phase-invariant channels. Thermal loss channels (lossy channels with $\bar{n} > 0$) can model long communication channels with environmental noise (such as the channels experienced by states sent down long optical fibres or by states transmitted through the atmosphere to or from a satellite). They are also applicable to quantum metrology scenarios in which light is being used to probe a sample. Additive Gaussian noise channels are good models for low loss scenarios, such

as some quantum reading, short-range quantum sensing or short-range quantum communications applications.

Pirandola et al. found bounds on the generic two-way capacity of thermal loss channels, using channel simulation and teleportation stretching (see Section 2.4), that are tight for high transmissivities [5]. Their bounds on additive Gaussian noise channels are tight for low induced noises. They also found bounds for thermal amplifier channels (which are tight for low gains), which are less important in quantum information but complete the set of phase-insensitive Gaussian channels.

### 2.3.2 Calculating the quantities of interest

We want to be able to calculate some of the quantities of interest mentioned in Section 2.2 for Gaussian states.

We start by describing a useful tool for Gaussian quantum information: the symplectic decomposition. Any covariance matrix can be diagonalised via a symplectic transformation. Specifically, for any valid covariance matrix, $V$, there exists a symplectic matrix, $S$, such that

$$V = SDS^T,\ D = \bigoplus_{i=1}^{N} \nu_i \mathbb{I}_2, \tag{2.50}$$

where $\mathbb{I}_2$ is the 2-dimensional identity matrix. The numbers $\nu_i$ are called the symplectic eigenvalues of $V$ and are equal to the absolute values of the eigenvalues of the matrix $i\Omega V$. Recall that every symplectic matrix represents a Gaussian unitary transformation; the existence of $S$ for every $V$ means that every Gaussian state can be formed by applying a Gaussian unitary to a direct sum of thermal states. This also tells us that any one-mode Gaussian state can be purified into a TMSV with $S$ applied to the system mode (since each mode of a TMSV is a thermal state if the other mode is discarded). Note that, whilst the symplectic eigenvalues are simple to find, the diagonalising symplectic, $S$, may not be. See [73] for an algorithm for constructing $S$.

The symplectic decomposition also gives us a quick way to check the validity of a covariance matrix. A positive matrix $V$ is the covariance matrix of a valid Gaussian state iff all of its symplectic eigenvalues are greater than or equal to $(2\kappa)^{-1}$. This is the uncertainty principle applied to Gaussian states. It is also obvious from the symplectic decomposition, since, if a symplectic eigenvalue were less than $(2\kappa)^{-1}$, the corresponding thermal state (with its covariance matrix written according to Eq. (2.37)) would have to have a mean photon number $\bar{n} < 0$, which is impossible.

This leads on to an easy way to check whether a two-mode state is separable [74]. Two-mode Gaussian states are separable iff they have a positive partial transpose. Let $V$ be the covariance

matrix of a valid Gaussian state and let $\tilde{V}$ be the covariance matrix of that state after partial transposition has been applied. The original state was only separable if the symplectic eigenvalues of the transformed state are all greater than or equal to $(2\kappa)^{-1}$. In other words, a state is only separable if the covariance matrix of the partially transposed state is also the covariance matrix of a Gaussian state. In the two-mode case, this condition is also sufficient for separability. $V$ is related to $\tilde{V}$ by

$$\tilde{V} = (\mathbb{I}_1 \oplus T_2)V(\mathbb{I}_1 \oplus T_2),\ T_2 = \bigoplus_{i=1}^{N_2} \mathbb{Z}, \tag{2.51}$$

where the subscripts 1 and 2 label the two subsystems (with respect to which we want to assess the separability) and $N_2$ is the dimension of the second subsystem. Note that despite the similar form of the transformation, $T$ does not represent a unitary operation, since it transforms valid states into invalid states.

Since enacting a unitary on a quantum state cannot change its entropy, the entropy of a Gaussian state can be calculated using its symplectic decomposition [75]. We can use the formula

$$S(\rho) = \sum_{i=1}^{N} g(\bar{n}_i), \tag{2.52}$$

$$g(x) = (x+1)\log_2(x+1) - x\log_2(x),\ \bar{n}_i = \kappa\nu_i - \frac{1}{2}, \tag{2.53}$$

where the $\bar{n}_i$ give the mean photon numbers of the thermal states in the symplectic decomposition and where $S$ is given in bits.

As discussed in Section 2.2, the trace norm between two states is important for tasks involving quantum state and channel discrimination. Whilst the trace norm itself is complicated to calculate for Gaussian states, the fidelity between any two Gaussian states (which bounds the trace norm from above and below and also can be used to calculate the QFI) has an analytical form that can be given in terms of the first moments vector and covariance matrix. Banchi et al. [76] showed that the fidelity of two Gaussian states, $F(\rho_1, \rho_2)$, with first moments vectors $\bar{x}_1$ and $\bar{x}_2$ and covariance matrices $V_1$ and $V_2$ (respectively), is given by

$$F(\rho_1, \rho_2) = \frac{F_{\text{tot}}}{\sqrt[4]{\det[\kappa(V_1+V_2)]}} e^{\frac{1}{4}(\bar{x}_1-\bar{x}_2)^T(V_1+V_2)^{-1}(\bar{x}_1-\bar{x}_2)}, \tag{2.54}$$

$$F_{\text{tot}} = \sqrt[4]{\det\left[2\kappa\left(\sqrt{\mathbb{I} + \frac{(V_{\text{aux}}\Omega)^{-2}}{4\kappa^2}} + \mathbb{I}\right)V_{\text{aux}}\right]}, \tag{2.55}$$

$$V_{\text{aux}} = \Omega^T(V_1+V_2)^{-1}\left(\frac{\Omega}{4\kappa^2} + V_2\Omega V_1\right), \tag{2.56}$$

where $\det$ is the determinant function and $\Omega$, the symplectic form, is defined as in Eq. (2.29).

## 2.4 Channel simulation and protocol stretching

Channel simulation is a highly useful technique for calculating the properties of quantum channels, with applications in quantum communications [5], quantum metrology, and quantum channel discrimination [6, 22, 25, 30]. It will be used extensively in this work. It involves the replacement, in a protocol (which may be any task involving quantum channels, such as parameter estimation or QKD), of a quantum channel with a similar or equivalent quantum operation, in order to simplify a variety of calculations involving the protocol.

Suppose, for example, we are presented with the task of distributing entanglement between two remote locations over a specific quantum channel, $\mathcal{C}$. The method by which we do so is called our protocol. We wish to distribute the maximum possible amount of entanglement for a fixed number of uses, and our protocol is constrained only by the laws of physics and the fact that we cannot act globally. In other words, our sender, who we will call Alice, can prepare and send any quantum states that she wishes over the channel $\mathcal{C}$ and can freely send and receive classical communications over a classical communications channel. Our receiver, who we will call Bob, can perform any quantum operations that he wishes on the states that he receives (or any ancillary states that he prepares himself), including measurements, and can freely send and receive classical communications to Alice between transmissions of states down the quantum channel (which we will call rounds of the protocol). Alice can even decide which state to send in a given round based on the classical communications she received from Bob in previous rounds. The only thing that Alice and Bob cannot do is perform a joint quantum operation on their combined quantum state; this is because they are in remote locations. Thus, they are restricted to local operations and classical communications (LOCCs).

In order to assess the performance of a protocol, we consider the relative entropy of entanglement (REE) of the protocol's output; this is the amount of entanglement that the protocol has distributed between the parties (assuming that Alice and Bob originally started with no shared entanglement; if this is not the case, we can consider the increase in the REE compared to the initial state). If we divide the amount of distributed entanglement by the number of rounds required to achieve it, we get the rate at which a given protocol distributes entanglement (for a fixed number of transmissions). If we take the number of rounds of the protocol to infinity, we get the asymptotic rate of entanglement distribution (note that this can never be less than the rate for a fixed number of rounds, since we can simply repeat a protocol with a finite number of rounds infinite times). Suppose we are now tasked with finding (or upper bounding) the maximum asymptotic rate of entanglement distribution for any possible protocol carried out over $\mathcal{C}$. This is called the two-way

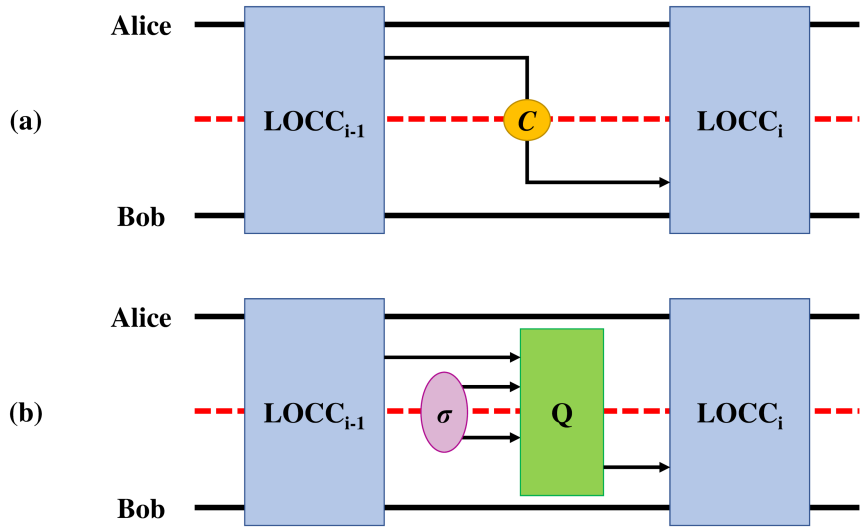


Figure 2.1: An example of quantum channel simulation. In panel (a), we show the $i$-th round of a protocol in which Alice and Bob carry out arbitrary LOCCs on their states before and after each channel use. The red dashed line demarcates the division between Alice's systems and Bob's systems and the locality of the LOCCs is defined with respect to this division. $C$ denotes the channel, and each channel use consists of Alice sending Bob a state via the channel. Note that this is a global (non-local) operation. In panel (b), we have replaced the channel $C$ with the quantum operation $Q$, which acts on the input state sent by Alice and on a resource state (denoted $\sigma$) to send the same output state to Bob as is sent by the channel in panel (a). We have therefore used $Q$ (and $\sigma$) to simulate $C$, and so Alice and Bob share the same state after the $i$-th channel use in panel (a) as in panel (b) (hence the REE between their states is the same in both cases). If $Q$ is also an LOCC (with respect to the division between Alice's systems and Bob's systems), then the channel use cannot have increased the REE between Alice and Bob by more than the REE of $\sigma$.

entanglement distribution capacity of $\mathcal{C}$ (it is called the two-way capacity because we allow classical communications in both directions). The fact that Alice and Bob can freely communicate between rounds and adapt their strategy accordingly means that we cannot simply assume that the REE of the output of the best possible $N$-round protocol is $N$ times the REE of the output of the best possible 1-round protocol. How, then, might we go about calculating this quantity?

We now apply the techniques of channel simulation and protocol stretching, as developed in Ref. [5], to this calculation, in order to demonstrate how they can be used. As a starting point, let us consider the fact that Alice and Bob are restricted to LOCCs. LOCCs cannot increase the REE between two remote parties. Therefore, none of the quantum operations between transmissions can have increased the REE between Alice's state and Bob's state (we will henceforth refer to this

simply as the REE between Alice and Bob). However, sending states down the quantum channel can increase the REE between Alice and Bob (obviously, or no protocol could ever have a non-zero rate).

Let $A$ denote the system that $\mathcal{C}$ takes as an input. Now, suppose there exists some quantum operation, $\mathcal{Q}$, that acts on the $A$ system of any input state, $\rho^{\text{in}}_{CA}$, which may be entangled with an idler system, $C$, and some ancillary state, $\sigma_B$, to produce an output state such that

$$\text{Tr}_B \left[ (\mathcal{I}_C \otimes \mathcal{Q}_{AB}) \left[ \rho^{\text{in}}_{CA} \otimes \sigma_B \right] \right] = \mathcal{I}_C \otimes \mathcal{C}_A(\rho^{\text{in}}_{CA}), \tag{2.57}$$

where $\mathcal{I}_C$ represents the identity channel on system $C$ and $\text{Tr}_B$ denotes the partial trace over system $B$. In other words, the output of $\mathcal{I}_C \otimes \mathcal{Q}_{AB}$, after tracing over the ancillary output system, $B$, is exactly the same as the output of $\mathcal{I}_C \otimes \mathcal{C}_A$, for every possible input state $\rho^{\text{in}}_{CA}$. We then say that $\mathcal{Q}$ simulates the channel $\mathcal{C}$. Note that, in general, the input and output states can have different dimensions, and that the output system need not be the "original" system $A$ (e.g. the operation $\mathcal{Q}$ could select a subsystem from $\sigma$ to be the output and then output the original input as part of the ancillary output). The idler system, $C$, is required for the same reason that an idler is required when calculating the diamond norm: two channels can affect the input system in the same way, but each can apply a different phase change to the system (relative to the idler system).

Suppose that the channel $\mathcal{C}$ were replaced by the quantum operation $\mathcal{Q}$. This replacement is depicted in Fig. 2.1, for the $i$-th channel use. In other words, suppose that each round, after Alice has prepared and transmitted her state, some third party, who we will call Charlie, implements the quantum operation $\mathcal{Q}$ on the transmitted state and an ancillary state $\sigma$ (that he prepared before the protocol started and which may be entangled) and then transmit the output state (after tracing over the ancillary system) through an identity channel to Bob. We call the state $\sigma$ a resource state, since it is a pre-prepared resource that is consumed by the operation $\mathcal{Q}$. Since the states held by Alice and Bob are exactly the same as they would be if Charlie had instead allowed Alice's signal state to simply pass through the channel $\mathcal{C}$, there is no possible way for Alice and Bob to know whether this is the case. Consequently, the output state of the protocol (and hence the final REE between Alice and Bob) must be the same in both cases. Thus, this new protocol, with every use of $\mathcal{C}$ replaced by a use of $\mathcal{Q}$, is equivalent to the old protocol in terms of every possible physical quantity that could be calculated for the output state.

Now let us suppose that $\mathcal{Q}$ is an LOCC operation with respect to some bipartition in the system $B$. By this we mean: suppose we can split Charlie's ancillary system into systems $B_1$ and $B_2$ (which may be entangled) and then perform the operation $\mathcal{Q}$ by carrying out some LOCCs on the systems $AB_1$ and $B_2$. The output that is sent on to Bob is drawn from the system $B_2$ (after the

LOCCs) and all of the remaining systems can be traced over (including the original input system, after the LOCCs have been applied). Then, the system $B_1$ can be considered to be part of Alice's system and the system $B_2$ can be considered to be part of Bob's system. Since all operations involved in the protocol are now LOCCs with respect to the bipartition between Alice's system and Bob's system, the REE of the output state of the protocol cannot be greater than the REE of the initial state (which must now include the REE of all of the copies of $\sigma_{B_1,B_2}$ used in the protocol). We have combined all of Alice and Bob's operations into a single round of LOCCs on an initially entangled state, and we say that the resource state has been "stretched" back in time to before all of the quantum operations. The new protocol is said to be in block form. Since each channel use has been replaced by one use of $\mathcal{Q}$ (with each use requiring one copy of $\sigma$), the REE between Alice and Bob after $N$ rounds of any protocol must be upper bounded by the REE of $\sigma_{B_1,B_2}^{\otimes N}$ (again, with respect to the bipartition between $B_1$ and $B_2$). Using the subadditivity of the REE, one can find the further condition that the REE of the output state is less than or equal to $N$ multiplied by the REE of $\sigma_{B_1,B_2}$, and hence that the two-way entanglement distribution capacity is upper bounded by the REE of $\sigma_{B_1,B_2}$, giving us the upper bound that we wanted.

The tightness of the upper bound on the two-way entanglement distribution capacity depends on the choice of the quantum operation, $\mathcal{Q}$, and the resource state, $\sigma$. Pirandola et al. [5] used quantum teleportation as the operation and found simple analytic bounds for a wide class of quantum channels, called Choi-stretchable channels. Specifically, these are the channels that commute with (standard) quantum teleportation and so can be simulated using their Choi matrices as the resource state.

Our replacement of the channel $\mathcal{C}$ with the quantum operation $Q$ is an example of quantum channel simulation. We then stretched the adaptive $N$-round protocol into a 1-round protocol enacted on an $N$-copy resource state. If the operation $\mathcal{Q}$ is a teleportation protocol, we call this process teleportation stretching.

For more information about the use of teleportation simulation for calculating channel capacities, see the review by Pirandola et al. [6].

### 2.4.1 Applications of the techniques

Although protocol stretching was first used to bound various channel capacities (such as the two-way entanglement distribution capacity, as described above) [5], channel simulation and the technique of stretching an initial resource state back in time to reduce an adaptive protocol to a block protocol can also be applied to a variety of other quantum information tasks.

Examples include finding the QFI of a channel parametrised by a variable, $\theta$, (and thereby bounding the performance of parameter estimation protocols) and finding channel discrimination bounds for a pair of channels, $\mathcal{C}_A$ and $\mathcal{C}_B$ [30]. For these tasks, we want to bound the performance of the most general protocols possible. We therefore allow the protocols to be adaptive, with unlimited quantum operations between channel uses; these general protocols can therefore be represented as quantum combs[5] [22, 23].

In both of these cases, the requirements for the simulating quantum operation are different from the case in which we want to bound the two-way entanglement distribution capacity. Specifically, we do not require that $\mathcal{Q}$ be an LOCC operation, since the sender is also the receiver and there is no requirement that operations on the states be in any way local. Instead, we must choose $\mathcal{Q}$ such that the channels are jointly programmable. In the parameter estimation case, this means that all of the channels parametrised by $\theta$ must be simulated by the same quantum operation, $\mathcal{Q}$, but with different resource states, $\sigma_\theta$. In the quantum channel discrimination case, this means that both of the channels must again be simulated by the same quantum operation, $\mathcal{Q}$, but with different resource states, $\sigma_A$ and $\sigma_B$. In this setting, we often refer to the resource states as program states and the quantum operation as the quantum processor, due to the correspondence with a programmable quantum gate array [77].

For greater detail about the applications of channel simulation and protocol stretching in quantum metrology see the reviews by Laurenza et al. [22] and Pirandola et al. [25].

### 2.4.2 Further considerations

It is worth noting that the condition in Eq. (2.57) is more restrictive than it needs to be. It defines a perfect simulation. In fact, we can consider an imperfect simulation, which meets the condition

$$\left\| \mathrm{Tr}_B[\mathcal{I}_C \otimes \mathcal{Q}_{AB}(\rho^{\mathrm{in}}_{CA} \otimes \sigma_B)] - \mathcal{I}_C \otimes \mathcal{C}_A(\rho^{\mathrm{in}}_{CA}) \right\|_1 \leq \epsilon, \tag{2.58}$$

where $\epsilon$ is some small, positive, real number that defines how close the simulation is to the actual channel $\mathcal{C}$. In other words, $\epsilon$ is the diamond norm between the original channel and its simulation. Then, the output of our $N$-round protocol has a trace norm from the output of some approximate,

[5] Just as quantum channels transform quantum states into other quantum states, quantum combs transform quantum channels into other quantum channels or transform quantum combs into other quantum combs. A quantum comb can be represented as a series of quantum operations with slots for quantum channels to fit in. The set of quantum combs is therefore the set of the most general maps transforming an input sequence of channels (with fixed causal order) into an output state. The concept of quantum combs is explained in greater detail in Ref. [23].

stretchable protocol that is upper bounded by $N\epsilon$ (using the triangle inequality and the fact that the quantum operations between channel uses are the same for both protocols and so cannot increase the trace norm between the outputs). We are still able to use this approximate simulation to calculate some channel properties, because it is often possible to add some $\epsilon$-dependent cost function to account for the imperfect simulation (the form of the function depends on the channel property that we are trying to bound).

As mentioned previously, Pirandola et al. [5] were able to simulate Choi-stretchable channels using the Choi matrices of the channels as their program states. This was possible because Choi-stretchable channels commute with the teleportation unitaries, meaning that for any teleportation unitary $U$,

$$\mathcal{C}(U\rho^{\text{in}}U^\dagger) = V\mathcal{C}(\rho^{\text{in}})V^\dagger, \tag{2.59}$$

for all input states $\rho^{\text{in}}$, where $V$ is some other unitary (which does not depend on $\rho^{\text{in}}$). If this is the case, the channel, $\mathcal{C}$, can be applied to Bob's half of a Bell state to produce a Choi state. Alice then carries out standard teleportation on the input state, $\rho^{\text{in}}$, using her half of the Choi state as the resource. Bob applies the appropriate correction unitary to his state (but with the teleportation unitaries $U_i$ swapped for the corresponding unitaries $V_i$), which has already had $\mathcal{C}$ applied to it, and hence teleportation with the Choi resource applies the channel $\mathcal{C}$ to the input state $\rho^{\text{in}}$.

Standard teleportation is only able to perfectly simulate certain quantum channels, even when using the most general resource states. In the DV case, standard teleportation can only perfectly simulate the Pauli channels [78]. Cope et al. generalised the standard teleportation protocol by introducing a noisy classical communication channel and thereby expanded the set of simulable channels [26]. Pirandola et al. [27] introduced conditional channel simulation, allowing some other classes of channels, such as the dephrasure channels (formed by the pointwise application of a Pauli-Z channel and an erasure channel) [79], to be simulated. The set of simulable channels can be expanded still further by considering port-based teleportation [80, 81]. This is a variant of quantum teleportation that is able to simulate any quantum channel in the asymptotic limit of infinite ports. Even for a finite number of ports, it can give a good enough simulation of many types of channels to be used as a tool for bounding the error of channel discrimination tasks [2].

Finally, we note that it is possible to use channel simulation and protocol stretching techniques on CV systems. In many cases, however, we will need to consider the asymptotic limit of a sequence of finite-energy simulations. This is because perfect quantum teleportation of a CV state requires infinite energy [82]; this is an unphysical situation and so we cannot use this type of teleportation as our simulating quantum operation. Instead, we can use the finite-energy

Braunstein-Kimble teleportation protocol [83], which can be parametrised by an energy constraint, $\mu$. For any finite $\mu$, we do not have perfect teleportation, and so our channel simulation will also be imperfect. We can find the energy-constrained diamond norm between the actual channel and its simulation for any energy constraint, $\tilde{\mu}$ [5, 65, 66]. We can then find a bound on whatever channel property we are trying to calculate based on the resource state for this finite-energy simulation. This bound will potentially be a function of $\mu$ and $\tilde{\mu}$. We must then take the limit as $\mu \to \infty$ and then $\tilde{\mu} \to \infty$ (note that the order of the limits can matter).

# Chapter 3

# Channel Position Finding

The work in Section 3.2 forms the basis of a paper that has been submitted to Physical Review A, whose authors are (in order) Jason Pereira, Leonardo Banchi, Quntao Zhuang, and Stefano Pirandola. The idea for the reduction of the fidelity calculation to a calculation between three-mode states came from Leonardo Banchi. The preprint of this work is available on the arXiv [12].

The work in Section 3.3 forms the basis of a paper published in Physical Review Research, whose authors are (in order) Jason Pereira, Quntao Zhuang, and Stefano Pirandola [9]. Quntao Zhuang proposed using photon counting and the maximum-likelihood estimation to give bounds for specific protocols and calculated the success probability of the measurement (Eqs. (3.100) to (3.107)).

The first section of this chapter will introduce the task of channel position finding (CPF) and describe some of its possible applications. The next section will give bounds on the error probability for an idler-free protocol applied to a sequence of pure loss channels. The third section will give ultimate bounds for an environment localisation task, which hold for all adaptive protocols, and will apply them to some physical scenarios. The final section summarises the presented work.

## 3.1 Introduction

CPF is a little-investigated but important subcategory of quantum channel discrimination. In channel discrimination, we know that an unknown channel is drawn from a set of possible channels and our goal is to determine which element of the set it is. In CPF, we have a sequence of channels, all but one of which are identical. The dissimilar channel is the target channel, the remaining channels are background channels, and our goal is to determine the label of the target channel (i.e. find its position in the sequence). This can be expressed as a special case of quantum channel dis-

crimination by considering the entire sequence of channels to be a single multi-mode channel and the channel sequences given by the different label options to be the elements in the set of possible multi-mode channels.

CPF is a less well-studied task than binary channel discrimination. Discrimination between multiple possible quantum states has been investigated, resulting in, for instance, the development of the pretty good measurement (PGM) [84, 85]. However, little research has been conducted on the error probability for discriminating between multiple possible quantum channels.

Recently, Zhuang and Pirandola [86] formulated a sequence of lower bounds on the error probability of identifying one channel from a set of possible channels that hold for any set of possible qudit channels and for the most general adaptive protocols. The bounds are based on channel simulation using PBT. They found that the error probability for any discrimination protocol for a set of $m$ possible channels, involving no more than $M$ channel uses is bounded by

$$p_{\text{err}} \geq \sum_{k,k' \,:\, k'>k} p_{k'} p_k F(\rho_{\mathcal{E}_{k'}}, \rho_{\mathcal{E}_k})^{2MN} - \frac{Md(d-1)}{N}, \tag{3.1}$$

where $p_i$ is the prior probability of the channel $\mathcal{E}_i$, $\rho_{\mathcal{E}_i}$ is the Choi matrix of $\mathcal{E}_i$, $d$ is the dimension of the channels in the set, $F$ is the Bures fidelity and $N$ is any positive integer. Thus, we have a sequence of lower bounds (one for each value of $N$) and must optimise over $N$ to find the tightest lower bound in the sequence. Here, $N$ represents the number of ports in the PBT simulation. Tighter bounds are given for cases in which the simulation error is known or the Helstrom limit between Choi matrices is easily calculable. Zhuang and Pirandola also simplified the bounds further for sets of channels that are jointly teleportation covariant and hence showed that, for such channel sets, there exists a non-adaptive discrimination protocol that is optimal (has the minimum possible error probability). This is a result that was previously only known to hold for binary discrimination [30]. They then applied these bounds to the task of CPF, presenting bounds for the discrimination of sets of erasure channels, depolarising channels and AD channels.

### 3.1.1 Channel position finding on lossy channels

An important case of CPF is locating a (bosonic) thermal loss channel with a different transmissivity or induced noise amongst a sequence of background lossy channels. This is a task with applications in quantum illumination [32,40,41], spectroscopy [87], and quantum reading [25,45]. In quantum illumination, one may know that a target is present in one of several locations but not know where. A discrimination protocol could involve probing the possible locations with light then collecting and carrying a measurement out on the return states. The different losses and in-

duced noises experienced by the probes, depending on whether they encountered the target or not, could be modelled as different lossy channels. A similar situation could arise in spectroscopy. In this scenario, the different channels could represent the optical absorbance of an unknown substance at different frequencies. Since different substances have different absorption spectra, finding the position of an absorption line could be equivalent to identifying the substance. In quantum reading, the reflectivity of a memory cell takes one of two possible values - encoding one of two possible bit values - and so readout is performed by probing the cell with signal states and discriminating between the possible channels. However, one could also consider a formulation in which bits are instead encoded in the position of a cell with a higher or lower transmissivity than the others [88].

Zhuang and Pirandola [88] upper bounded the performance of classical CPF protocols (i.e. non-adaptive protocols that only use signal states with positive semidefinite P-representations[1]). They then calculated the performance of a specific non-classical protocol and thereby showed a quantum advantage for the task. This protocol involves sending two-mode squeezed vacuum (TMSV) states through the channels and then measuring them with a proposed new type of receiver called the generalised conditional nulling receiver. Their bounds are applied to quantum reading - modelled as a scenario of CPF between pure loss channels - and quantum target finding, a task of quantum illumination.

Zhuang and Pirandola upper bounded the optimal error probability by proposing a specific protocol and calculating its error probability, but they did not lower bound the error probability. This highlights a difficulty with bounding the performance of CPF on lossy channels over all adaptive protocols. Any two lossy channels with different transmissivities are not jointly teleportation covariant. As such, the technique of teleportation stretching [5, 22, 30], for reducing an adaptive protocol to a block protocol, cannot be easily applied. One route for research is to find approximations of the channels that are jointly simulable by some teleportation protocol (for instance, by using continuous variable PBT [91]; little research has been carried out on this topic). Another is to find bounds for specific protocols, as Zhuang and Pirandola did. The non-classical protocol that they investigated required the retention of idler modes. If we do not have a good quantum memory (with low decoherence over a long storage time), the correlations between the signal and the idler

[1]Whilst entangled states constitute one type of quantum state with no classical analogue, they are not the only such states. In fact, there exist states with no classical analogue that are completely unentangled. One way of determining whether a state is classical is to use the Glauber–Sudarshan P representation [89, 90]. If the P-representation of a state is not positive semidefinite then that state has no classical analogue.

modes will quickly degrade and the benefits of using an initially entangled state will be lost. The idler-free case is therefore worth investigating, since this will tell us whether we can still have a quantum advantage even in the technologically limited case in which we cannot store an idler.

### 3.1.2 Environment localisation

A related case to discriminating between lossy channels with different transmissivities is CPF on a sequence of channels that all have the same transmissivity but for which the target channel has a different induced noise. The channels could be thermal loss channels, additive Gaussian noise channels, or thermal amplifier channels, depending on the transmissivity. Since the action of a phase-insensitive Gaussian channel is equivalent to mixing the signal state with some environmental thermal state at a beamsplitter [55, 92], CPF on channels with a fixed transmissivity can be regarded as environment localisation: finding the target environmental thermal state. This task has applications to thermal imaging, since the mean number of photons in an environmental mode can depend on the temperature of the environment, and quantum communications, where it can be applied to tasks such as eavesdropper localisation (attempting to find the communications line or section of line with a higher induced noise, potentially due to the presence of an eavesdropper).

An important feature of a sequence of phase-insensitive Gaussian channels with fixed transmissivity is that the channels are jointly teleportation covariant. This means that, unlike the scenario in which the target and background channels have different transmissivities, ultimate bounds on the error probability can be established using channel simulation and teleportation stretching. This is why Pirandola and Lupo could bound the minimum variance for an estimation of the noise of a thermal loss channel using teleportation stretching [30]. Bounds established in this way will hold for the most general adaptive protocols.

## 3.2 Idler-free channel position finding

We consider CPF between a sequence of pure-loss channels for a specific type of one-shot protocol. The protocols we consider send fully symmetric Gaussian states through the sequence of channels. They are non-adaptive and idler-free, meaning that the output state has tensor product form and we do not retain any modes that are entangled with the signal modes before they are sent through the channels. The advantage of idler-free protocols is that they can be easier to implement. In order to benefit from the use of an idler, the idler must be stored in a quantum memory, potentially for a long time (if the signal states take a significant time to pass though the channels).

Building quantum memories that simultaneously have a long storage time and a high memory efficiency is still a challenging area of research [93–95]. We allow entanglement between the signal states for each channel, but constrain the total mean number of photons sent through the channel sequence. Note that we set the vacuum noise equal to 1 in this section.

### 3.2.1 Finding the covariance matrices of the possible outputs

Consider a sequence of $m$ one-mode, pure-loss channels, where $m-1$ of the channels are identical "background" channels and one of the channels is a target channel. The target channel has transmissivity $\eta_T$, whilst the background channels all have transmissivity $\eta_B$. The task is to locate the target channel using an idler-free protocol. If we are allowed to send unlimited energy into the channels, the error probability trivially goes to 0, so we impose an energy constraint on our CPF protocol, allowing no more than $N_S$ photons to be sent through each channel. The $m$-partite channel input that we consider has no first moments and the CM[2]

$$V_{\text{in}} = \begin{pmatrix} \mu\mathbb{I} & \Gamma & \dots & \Gamma \\ \Gamma & \mu\mathbb{I} & \ddots & \Gamma \\ \vdots & \ddots & \ddots & \vdots \\ \Gamma & \Gamma & \dots & \mu\mathbb{I} \end{pmatrix}, \ \Gamma := \text{diag}(c_1, c_2), \tag{3.2}$$

where $\mu$ is specified by the energy constraint (via $\mu = 2N_S + 1$) and $c_1$ and $c_2$ determine the level of entanglement between the modes. We want to calculate the error probability for CPF using a signal state of this form.

The problem of CPF can be reduced to state discrimination between the $m$ possible outputs of the adaptive protocol used (with each outcome corresponding to a different target channel position). By bounding the fidelity between the different output states, we can find both upper and lower bounds for the minimum error probability $p_{\text{err}}$ (optimised over all adaptive protocols) of state discrimination. The lower bound on the discrimination error between a sequence of $m$ states $\{\rho_i\}$, with probabilities $\{p_i\}$, is [96]

$$p_{\text{err}} \geq \sum_{i,j \,:\, i>j}^{m} p_i p_j F(\rho_i, \rho_j)^2, \tag{3.3}$$

[2]This is a fully symmetric (invariant under permutation of modes) Gaussian state. In the case of maximal correlations, this state can be regarded as a Gaussian analogue of the GHZ state (a particular type of multipartite entangled DV state), since it is a maximally entangled (for a given energy constraint), fully symmetric state.

and the upper bound, based on the PGM, is [97]

$$p_{\text{err}} \leq 2 \sum_{i,j\,:\,i>j}^{m} \sqrt{p_i p_j} F(\rho_i, \rho_j), \tag{3.4}$$

where $F$ is the Bures fidelity, defined as

$$F(\rho_i, \rho_j) = \text{Tr}\sqrt{\sqrt{\rho_i}\rho_j\sqrt{\rho_i}}. \tag{3.5}$$

As a result, finding the fidelity between the possible output states of this protocol will allow us to bound its error probability.

$V_{\text{in}}$ has only two distinct symplectic eigenvalues [98]:

$$\nu_- = \sqrt{(\mu - c_1)(\mu - c_2)},\ \nu_+ = \sqrt{(\mu + (m-1)c_1)(\mu + (m-1)c_2)}. \tag{3.6}$$

$\nu_+$ is $m-1$ times degenerate. We set $c_1 = -c_2 = c$ and assume maximal correlations (meaning that we maximise $c$). By requiring $V_{\text{in}} > 0$ and $\nu_\pm \geq 1$ (the bona fide condition), we find that the maximum value of $c$ is

$$c_{\max} = \frac{\sqrt{\mu^2 - 1}}{m-1}. \tag{3.7}$$

Let us now define $\rho_i$ as the output state for the case in which the $i$-th channel is the target channel. Since both our input and the channels involved are Gaussian, the output states $\rho_i$ are also Gaussian and hence (since, like the input, they will have no first moments) can be described entirely by their CMs, $V_i$. Therefore, the fidelity between any pair of possible output states $\rho_i$ and $\rho_j$ can be expressed as $F(V_i, V_j)$. The CM of output state $\rho_i$ is

$$V_i = \begin{pmatrix} \Delta_B & \cdots & \Gamma_B & \Gamma_T & \Gamma_B & \cdots & \Gamma_B \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ \Gamma_B & & \Delta_B & \Gamma_T & \Gamma_B & \cdots & \Gamma_B \\ \Gamma_T & \cdots & \Gamma_T & \Delta_T & \Gamma_T & \cdots & \Gamma_T \\ \Gamma_B & \cdots & \Gamma_B & \Gamma_T & \Delta_B & & \Gamma_B \\ \vdots & & & \vdots & & \ddots & \vdots \\ \Gamma_B & \cdots & \Gamma_B & \Gamma_T & \Gamma_B & \cdots & \Delta_B \end{pmatrix}, \tag{3.8}$$

where we have defined

$$\Delta_B := (\eta_B\mu + (1-\eta_B))\mathbb{I},\ \Delta_T := (\eta_T\mu + (1-\eta_T))\mathbb{I}, \tag{3.9}$$

$$\Gamma_B = \eta_B c_{\max}\mathbb{Z},\ \Gamma_T = \sqrt{\eta_B\eta_T}c_{\max}\mathbb{Z}. \tag{3.10}$$

$\mathbb{Z}$ is the Pauli Z matrix.

We now need to find the fidelity between pairs of $m$-mode CMs. However, we can greatly simplify our calculations by reducing the problem to the fidelity between two three-mode CMs, via a unitary transformation of our output states.

### 3.2.2 Reduction to the fidelity between three-mode systems

Due to the symmetry, $F(V_i, V_j) = F(V_1, V_2)$ for all $i \neq j$, so it suffices to calculate $F(V_1, V_2)$. Let $\{\hat{a}_i\}$ be the set of annihilation operators for all of the modes. We can transform $\{\hat{a}_i\}$ via the unitary

$$U = I_{1,2} \otimes U' \tag{3.11}$$

where $I_{1,2}$ is the identity on modes 1 and 2 and where $U'$ has elements

$$U'_{jk} = e^{ijk\phi},\ \phi = \frac{2\pi}{m-2}. \tag{3.12}$$

We can verify that $U'$ is a valid unitary by writing

$$(U'U'^{\dagger})_{jk} = \sum_{l=1}^{m-2} e^{i(k-j)l\phi} \tag{3.13}$$

$$= \delta_{jk}. \tag{3.14}$$

$U$ transforms $\{\hat{a}_i\}$ into $\{\hat{a}'_i\}$, where

$$\hat{a}'_1 = \hat{a}_1,\ \hat{a}'_2 = \hat{a}_2, \tag{3.15}$$

$$\hat{a}'_{3+j} = \frac{1}{\sqrt{m-2}} \sum_{k=0}^{m-3} e^{ik\phi} \hat{a}_{3+k}. \tag{3.16}$$

This means that the quadrature operators of the modes, $\{q_i\}$ and $\{\hat{p}_i\}$, are transformed into $\{q'_i\}$ and $\{\hat{p}'_i\}$, where

$$\hat{q}'_1 = \hat{q}_1,\ \hat{q}'_2 = \hat{q}_2, \tag{3.17}$$

$$\hat{p}'_1 = \hat{p}_1,\ \hat{p}'_2 = \hat{p}_2, \tag{3.18}$$

$$\hat{q}'_{3+j} = \frac{1}{\sqrt{m-2}} \sum_{k=0}^{m-3} \left[\cos(jk\phi)\hat{q}_{3+k} - \sin(jk\phi)\hat{p}_{3+k}\right], \tag{3.19}$$

$$\hat{p}'_{3+j} = \frac{1}{\sqrt{m-2}} \sum_{k=0}^{m-3} \left[\sin(jk\phi)\hat{q}_{3+k} + \cos(jk\phi)\hat{p}_{3+k}\right]. \tag{3.20}$$

These are calculated using the relations $\hat{q} = \hat{a} + \hat{a}^{\dagger}$ and $\hat{p} = i(\hat{a}^{\dagger} - \hat{a})$.

This transformation puts both $V_1$ and $V_2$ in block diagonal form, such that the resulting CM has a 6 by 6 block and a $2m-6$ by $2m-6$ block, the latter of which is the same in both cases. We can verify this by calculating the components of the transformed CMs, $V_1'$ and $V_2'$. In order to demonstrate how this is done, let us explicitly calculate the value of $\left\langle \hat{p}_1' \hat{p}_{3+j}' \right\rangle$ for $V_1'$. Using the expression in Eq. (3.20), we get

$$\left\langle \hat{p}_1' \hat{p}_{3+j}' \right\rangle = \frac{1}{\sqrt{m-2}} \sum_{k=0}^{m-3} \left[ \sin(jk\phi) \left\langle \hat{p}_1 \hat{q}_{3+k} \right\rangle + \cos(jk\phi) \left\langle \hat{p}_1 \hat{p}_{3+k} \right\rangle \right]. \tag{3.21}$$

The covariances $\langle \hat{p}_1 \hat{q}_{3+k} \rangle$ and $\langle \hat{p}_1 \hat{p}_{3+k} \rangle$ are components of the original covariance matrix, $V_1$, and are given in Eqs. (3.8) to (3.10). First, note that $\langle \hat{p}_i \hat{q}_j \rangle = 0$ for all $i$ and $j$ (since every 2 by 2 submatrix of $V_1$ is diagonal). Defining

$$d_B = (\eta_B \mu + (1-\eta_B)),\ d_T = (\eta_T \mu + (1-\eta_T)), \tag{3.22}$$

$$\gamma_B = \eta_B c_{\max},\ \gamma_T = \sqrt{\eta_B \eta_T} c_{\max}, \tag{3.23}$$

we obtain

$$\begin{aligned} \left\langle \hat{p}_1' \hat{p}_{3+j}' \right\rangle &= -\frac{\gamma_T}{\sqrt{m-2}} \sum_{k=0}^{m-3} \cos(jk\phi) \\ &= -\sqrt{m-2}\gamma_T \delta_{0,j}, \end{aligned} \tag{3.24}$$

where $\delta$ is the Kronecker delta symbol, and where we have used the result

$$\sum_{k=1}^{l} \cos\left(jk\frac{2\pi}{l}\right) = l\delta_{0,j} \tag{3.25}$$

(for integer $l$). Note that this is 0 for $j > 0$, i.e. for all modes with labels greater than 3. $\left\langle \hat{q}_1' \hat{q}_{3+j}' \right\rangle$ is simply $-\left\langle \hat{p}_1' \hat{p}_{3+j}' \right\rangle$, and $\left\langle \hat{p}_2' \hat{p}_{3+j}' \right\rangle$ can be obtained simply by substituting $\eta_B c_{\max}$ for $\sqrt{\eta_B \eta_T} c_{\max}$, giving

$$\left\langle \hat{p}_2' \hat{p}_{3+j}' \right\rangle = -\sqrt{m-2}\gamma_B \delta_{0,j}. \tag{3.26}$$

Note that for $V_2$, we simply swap $\left\langle \hat{p}(\hat{q})_1' \hat{p}(\hat{q})_{3+j}' \right\rangle$ and $\left\langle \hat{p}(\hat{q})_2' \hat{p}(\hat{q})_{3+j}' \right\rangle$.

We have now shown that no correlations exist between modes 1 and 2 and modes 4 to $m$. In order to show that the transformation puts the CM in block diagonal form, we must also show that no correlations exist between mode 3 and modes 4 to $m$. To do this, we must calculate $\left\langle \hat{p}(\hat{q})_3' \hat{p}(\hat{q})_{3+j}' \right\rangle$. Again using Eq. (3.20), we obtain

$$\left\langle \hat{p}_3' \hat{p}_{3+j}' \right\rangle = \frac{1}{m-2} \sum_{k,l=0}^{m-3} \left[ \sin(jk\phi) \left\langle \hat{p}_{3+l} \hat{q}_{3+k} \right\rangle + \cos(jk\phi) \left\langle \hat{p}_{3+l} \hat{p}_{3+k} \right\rangle \right]. \tag{3.27}$$

Substituting in Eqs. (3.22) and (3.23), we derive

$$\left\langle \hat{p}'_3 \hat{p}'_{3+j} \right\rangle = -\gamma_B \sum_{k=0}^{m-3} \cos(jk\phi) + \frac{d_B + \gamma_B}{m-2} \sum_{k=0}^{m-3} \cos(jk\phi), \tag{3.28}$$

where we have split the expression into contributions from the on and off-diagonal components of the original CMs. Simplifying, we get

$$\left\langle \hat{p}'_3 \hat{p}'_{3+j} \right\rangle = (d_B - (m-3)\gamma_B)\delta_{0,j}, \tag{3.29}$$

thus there are no correlations between mode 3 and modes 4 to $m$. We have therefore carried out a unitary transform on $V_1$ and $V_2$ such that they are in block diagonal form, with a 6 by 6 block and a $2m-6$ by $2m-6$ block. Since the $2m-6$ by $2m-6$ block is the same for both $V_1$ and $V_2$, we can ignore this block (trace over the remaining $m-3$ modes) when calculating the fidelity of the two CMs. This reduces the problem to the analytically solvable case of finding the fidelity of a pair of three-mode Gaussian states.

Let $V_1'$ be the CM of $\rho_1$ after the unitary $U$ has been enacted on it, transforming it into block diagonal form. Then, let $V_1^{3-\mathrm{mode}}$ be the CM after the trace has been taken over the last $m-3$ modes. $V_1^{3-\mathrm{mode}}$ takes the form

$$V_1^{3-\mathrm{mode}} = \begin{pmatrix} \Delta_T & \Gamma_T & \sqrt{m-2}\Gamma_T \\ \Gamma_T & \Delta_B & \sqrt{m-2}\Gamma_B \\ \sqrt{m-2}\Gamma_T & \sqrt{m-2}\Gamma_B & \Delta_B + (m-3)\Gamma_B \end{pmatrix}. \tag{3.30}$$

To obtain $V_2^{3-\mathrm{mode}}$, we simply swap modes 1 and 2.

We can also calculate the structure of the traced over modes, although this does not affect the fidelity calculation, since it is the same for both $V_1'$ and $V_2'$. Let us calculate $\left\langle \hat{p}'_{3+j} \hat{p}'_{3+k} \right\rangle$ for $j, k > 0$. Considering only the non-zero components, we get

$$\begin{aligned} \left\langle \hat{p}'_{3+j} \hat{p}'_{3+k} \right\rangle = \frac{1}{m-2} \sum_{x,y=0}^{m-3} & \left[ \sin(jx\phi) \sin(ky\phi) \left\langle \hat{q}_{3+x} \hat{q}_{3+y} \right\rangle \right. \\ & \left. + \cos(jx\phi) \cos(ky\phi) \left\langle \hat{p}_{3+x} \hat{p}_{3+y} \right\rangle \right]. \end{aligned} \tag{3.31}$$

We now split this into three terms, by writing

$$\left\langle \hat{p}'_{3+j}\hat{p}'_{3+k}\right\rangle = \frac{t_1+t_2+t_3}{m-2}, \tag{3.32}$$

$$t_1 = \gamma_B \sum_{x,y=0}^{m-3} \left[\sin\left(jx\phi\right)\sin\left(ky\phi\right) - \cos\left(jx\phi\right)\cos\left(ky\phi\right)\right], \tag{3.33}$$

$$t_2 = d_B \sum_{x=0}^{m-3} \left[\sin\left(jx\phi\right)\sin\left(kx\phi\right) + \cos\left(jx\phi\right)\cos\left(kx\phi\right)\right], \tag{3.34}$$

$$t_3 = -\gamma_B \sum_{x=0}^{m-3} \left[\sin\left(jx\phi\right)\sin\left(kx\phi\right) - \cos\left(jx\phi\right)\cos\left(ky\phi\right)\right]. \tag{3.35}$$

We can then write

$$t_1 = -\gamma_B \sum_{x,y=0}^{m-3} \cos\left((jx+ky)\phi\right), \tag{3.36}$$

$$t_2 = d_B \sum_{x=0}^{m-3} \cos\left((j-k)x\phi\right), \tag{3.37}$$

$$t_3 = \gamma_B \sum_{x=0}^{m-3} \cos\left((j+k)x\phi\right), \tag{3.38}$$

where we have used

$$\cos\left(a+b\right) = \cos\left(a\right)\cos\left(b\right) - \sin\left(a\right)\sin\left(b\right). \tag{3.39}$$

Since $j,k>0$, $t_1=0$. $t_2$ is non-zero iff $j=k$ and $t_3$ is non-zero iff $j+k=m-2$. We therefore derive

$$\left\langle \hat{p}'_{3+j}\hat{p}'_{3+k}\right\rangle = d_B\delta_{j,k} + \gamma_B\delta_{j+k,m-2}. \tag{3.40}$$

Via a similar derivation, we find

$$\left\langle \hat{q}'_{3+j}\hat{q}'_{3+k}\right\rangle = d_B\delta_{j,k} - \gamma_B\delta_{j+k,m-2}. \tag{3.41}$$

We now have all of the components of the CM of the traced over modes.

The fidelity $F(V_1,V_2) = F(V_1^{3-\mathrm{mode}}, V_2^{3-\mathrm{mode}})$ can now be easily found using the formula from Ref. [76].

### 3.2.3 Numerical investigations

We investigate the behaviour of the idler-free fidelity function. The output fidelity of a classical protocol, as calculated in Ref. [88], and the output fidelity of a protocol in which each channel is individually probed by one mode of a bipartite entangled state with the idler retained (which we

refer to as the bipartite entangled protocol) are natural points of comparison. If the output fidelity of the idler-free protocol is lower than that of the classical protocol over some parameter range, it would be an indication that there is a benefit to using the input state described by Eq. (3.2), rather than using the classical protocol. If the output fidelity for the idler-free protocol is close to that of the entangled state protocol with idlers, this would indicate that the cost to performance of using an idler-free protocol is small. Note that this is only an indication, as the fidelity is a measure of the distinguishability of states, but this does not necessarily mean that the error probability in discriminating between states is completely determined by the fidelity. In order to prove an advantage of one protocol over another, we would have to bound the error probabilities based on the output fidelities. This was done in Ref. [88] to prove that a protocol involving bipartite entangled states has a quantum advantage over classical protocols.

Note that for both the classical protocol and the bipartite entangled protocol, the possible output states, $\rho_i$ (where the label $i$ indicates that the $i$-th channel is the target channel), are all in the tensor product form

$$\rho_i = \rho_i^1 \otimes \rho_i^2 \otimes \ldots \rho_i^i \otimes \ldots \rho_i^m, \tag{3.42}$$

where the state $\rho_i^j$ is the output of the $j$-th channel (conditioned on the $i$-th channel being the target channel). Now let $\rho_B$ be the output state from a background channel and let $\rho_T$ be the output state from a target channel. We can then write the output fidelity as

$$\begin{aligned} F(\rho_i, \rho_j) &= F(\rho_i^1 \otimes \ldots \rho_i^i \otimes \ldots \rho_i^j \otimes \ldots \rho_i^m, \rho_j^1 \otimes \ldots \rho_j^i \otimes \rho_j^j \otimes \ldots\ldots \rho_j^m) \\ &= F(\rho_i^i \otimes \rho_i^j, \rho_j^i \otimes \rho_j^j) \\ &= F(\rho_B, \rho_T)^2, \end{aligned} \tag{3.43}$$

where we have used the fact that the fidelity is multiplicative with respect to tensor products. This means that the number of channels in the sequence, $m$, has no effect on the fidelity between any pair of possible outputs.

The classical protocol involves sending coherent (displaced vacuum) states through the channels. The displacement of the states is the maximum allowed by the average photon number constraint on the signal states, and the energy of the input states is evenly distributed amongst the $m$ probes. The fidelity between output states is given by

$$F^{\text{class}} = e^{-N_S(\sqrt{\eta_B}-\sqrt{\eta_T})^2}. \tag{3.44}$$

The bipartite entangled protocol individually probes each channel with one mode of a bipartite entangled state (with no correlation between the signal states for each channel). Each bipartite entangled state is a TMSV with the maximum squeezing parameter allowed by the energy constraint.

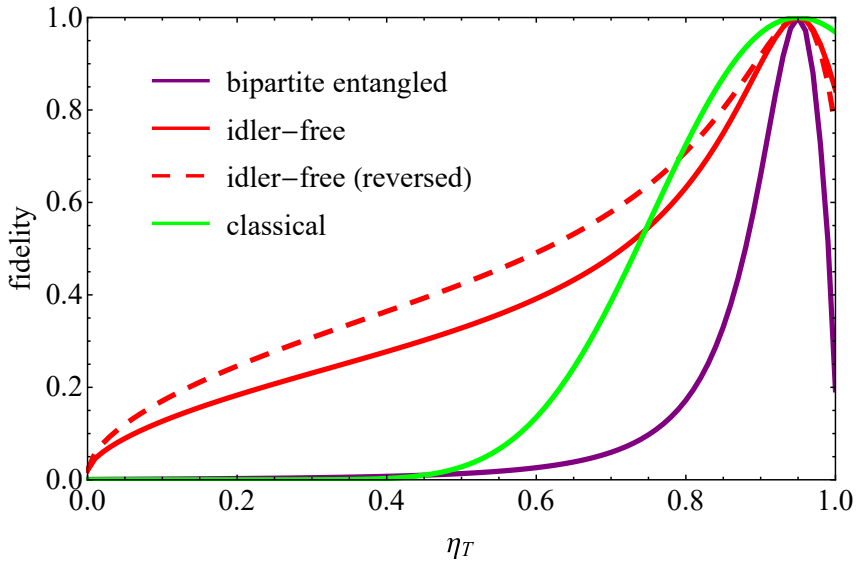


Figure 3.1: The output fidelity of the classical, bipartite entangled, and idler-free protocols as a function of the transmissivity of the target channel, $\eta_T$. We set the transmissivity of the background channels, $\eta_B$, to $0.95$ and impose an energy constraint so that the average number of photons per channel use is no more than $50$. We also set $m = 3$, so that there are two identical background channels and one target channel. The output fidelity for the idler-free protocol with $\eta_B$ and $\eta_T$ swapped is also shown. Unlike for the classical and bipartite entangled protocols, this swap affects the output fidelity for the idler-free protocol (since, in the classical and bipartite entangled cases, the output states are in tensor product form). The output fidelities are highest when $\eta_T$ is close to $\eta_B$ and decrease as the difference between the two transmissivities increases. The idler-free protocol gives a lower output fidelity than the classical protocol for $\eta_T \gtrapprox 0.75$.

The bipartite entangled protocol has an output fidelity of

$$F^{\text{bipartite}} = (1 + N_S(1 - \sqrt{(1-\eta_B)(1-\eta_T)} - \sqrt{\eta_B \eta_T}))^{-2}. \tag{3.45}$$

In Fig. 3.1, we plot the output fidelities for the various protocols against the transmissivity of the target channel, $\eta_T$. We fix the background transmissivity, $\eta_B = 0.95$, the number of channels in the sequence, $m = 3$, and the average number of photons sent through each channel per channel use, $N_S = 50$. We see that there is a region ($\eta_T \gtrapprox 0.75$) for which the idler-free protocol has a lower fidelity than the classical protocol. This indicates that the idler-free protocol could have a use as an intermediate between the easily implemented classical protocol, based on the sending of coherent states, and the bipartite entangled protocol, which gives a lower output fidelity in this range but could be harder to implement, due to the need for a quantum memory to preserve the idlers. The idler-free protocol could be easier to implement, despite the fact it still requires the

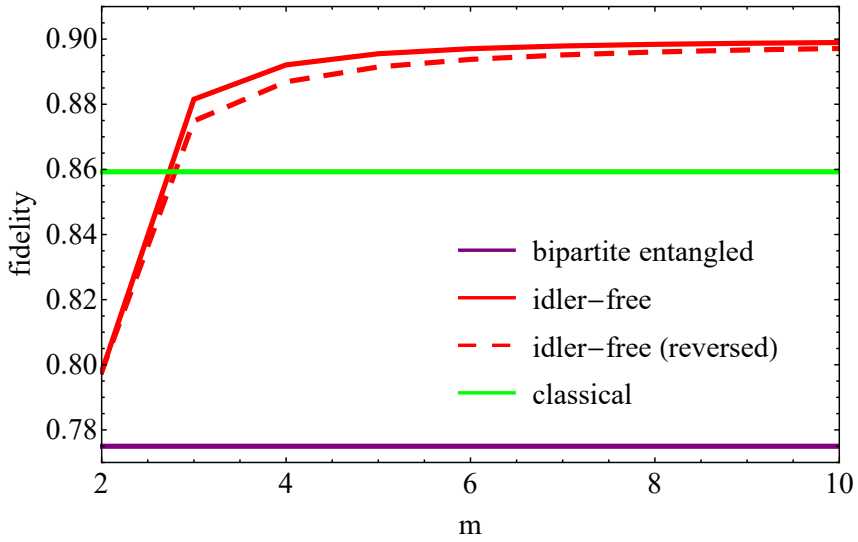


Figure 3.2: The output fidelity of the classical, bipartite entangled, and idler-free protocols as a function of the total number of channels in the sequence, $m$. We set the background transmissivity, $\eta_B$, to 0.2, the transmissivity of the target channel, $\eta_T$, to 0.7, and the average number of photons per channel use, $N_S$, to 1. Only the idler-free protocol is affected by changing $m$. We see that the output transmissivity increases as $m$ increases, but levels off for large $m$. As $m$ increases, the effect on the output fidelity of swapping $\eta_B$ and $\eta_T$ decreases.

generation of a non-classical state, because it does not require a quantum memory.

Fig. 3.1 also has a curve labelled "idler-free (reversed)". This gives the fidelity for the idler-free protocol when the values of $\eta_B$ and $\eta_T$ are swapped. It is immediate from Eq. (3.43) that neither the fidelity of the classical protocol nor that of the bipartite entangled protocol are affected by swapping $\eta_B$ and $\eta_T$, however this is not the case for the idler-free protocol (for $m > 2$). In fact, Fig. 3.1 shows that there can be a significant difference between the two fidelities.

Fig. 3.2 plots the various output fidelities against the number of channels in the sequence. In this plot, $\eta_B = 0.2$, $\eta_T = 0.7$, and $N_S = 1$. As previously mentioned, the output fidelities of the classical and the bipartite entangled protocols do not depend on $m$. Fig. 3.2 shows that the output fidelity for the idler-free protocol increases as $m$ increases, but levels off for large $m$.

Since the output fidelity for the idler-free protocol increases with $m$, it makes sense to study the $m = 2$ case when comparing the protocols. It is possible to analytically find the output fidelity for this case. We calculate

$$F^{\text{idler-free,2-mode}} = (1 + N_S(\eta_B + \eta_T - 2\eta_B\eta_T - 2\sqrt{\eta_B\eta_T(1-\eta_B)(1-\eta_T)}))^{-1}. \quad (3.46)$$

Fig. 3.3 plots the output fidelities against the average number of photons sent into each channel.

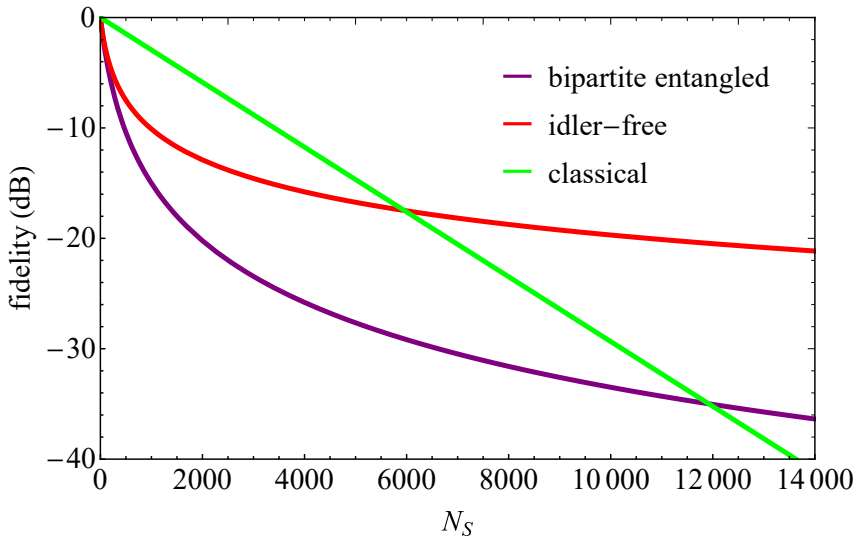


Figure 3.3: The output fidelity of the classical, bipartite entangled, and idler-free protocols as a function of the average number of photons in the signal states, $N_S$. We set the background transmissivity, $\eta_B = 0.9$, the transmissivity of the target channel, $\eta_T = 0.95$, and the number of channels in the sequence, $m = 2$. Fidelity is given in decibels. The output fidelity of the classical protocol gives a straight line because the scale is logarithmic and the classical output fidelity scales exponentially. This line crosses the curves representing the output fidelities for both the idler-free and the bipartite entangled protocols, showing that the classical protocol gives a lower output fidelity than either of the other protocols over some parameter ranges.

We have set $\eta_B = 0.9$, $\eta_T = 0.95$, and $m = 2$; since there are only two channels in the sequence, switching $\eta_B$ and $\eta_T$ does not result in a different task, and so we do not plot the case with $\eta_B$ and $\eta_T$ switched. The fidelity is given in decibels; this allows it to be clearly seen that the output fidelity of the classical protocol scales exponentially with $N_S$, since the curve is linear in a log scale. In fact, this is evident from the form of the expression in Eq. (3.44). On the other hand, $F^{\text{bipartite}}$ is inversely proportional to a polynomial in $N_S$. Considering the expression in Eq. (3.45) for large $N_S$, we see that it scales as roughly $N_S^{-2}$. We can see from Fig. 3.3 that the scaling of the idler-free output fidelity is also less than exponential. From Eq. (3.46), it can be seen that the output fidelity in the $m = 2$ case scales as approximately $N_S^{-1}$ for large $N_S$. Since the output fidelity is lowest in the $m = 2$ case, the classical protocol will always beat the idler-free protocol (and the bipartite entangled protocol) for sufficiently high $N_S$, due to the different scalings.

## 3.3 Optimal environment localisation

We now consider a case of CPF in which the target channel has the same transmissivity as the background channel but a different induced noise at the output. In this scenario, the channel outputs are not identical even in the case of a vacuum input (i.e. when no signal states are sent into the channels). We refer to this as the channel sequence having a passive signature. The task can be regarded as environment localisation: we are finding the channel whose Stinespring dilation [55, 92] has a different environmental noise from that of the other channels. We consider all phase-insensitive, Gaussian channels: these comprise the thermal loss channels, the thermal amplifier channels, and the additive noise channels. A key property of a sequence of one-mode, phase-insensitive, Gaussian channels with the same transmissivity is that - unlike a sequence of channels with different transmissivities - they are jointly teleportation covariant. This is important because it means that it is possible to use channel simulation in order to establish lower bounds on the error probability for discriminating between the channels in the sequence, even for the most general, adaptive protocols [22]. Note that we set the vacuum noise equal to $\frac{1}{2}$ in this section.

### 3.3.1 Channel simulation

Consider a sequence of $m$ one-mode, phase-insensitive, Gaussian channels, where $m-1$ of the channels are identical "background" channels and one of the channels is a target channel. The target channel has the same transmissivity, $\tau$, as the background channels, but a different induced noise, $\nu$ (note that we consider a generalised transmissivity which may take values between zero and infinity). Suppose we want to identify the target channel and can do so by probing the sequence of channels using some adaptive protocol that involves sending $M$ transmissions through the sequence of channels (each transmission consists of sending a one-mode state through every channel in the sequence). We do not impose any energy bound on the transmissions. We would like to bound the minimum probability of error in identifying the target channel, with the minimisation carried out over all possible adaptive protocols. The structure of the most general adaptive protocol can be considered to be a quantum comb [22, 23].

A schematic of a possible setup is given in Fig. 3.4, which shows a sequence of three thermal loss channels with the same transmissivity, $\tau$. Two of these channels are background channels (with environmental noise $\bar{n}_B$) and one of the channels is the target channel (with environmental noise $\bar{n}_T$). At each channel use, we are allowed to send an input state through the sequence of channels, and this input state may be dependent on the previous channel outputs. Each channel is

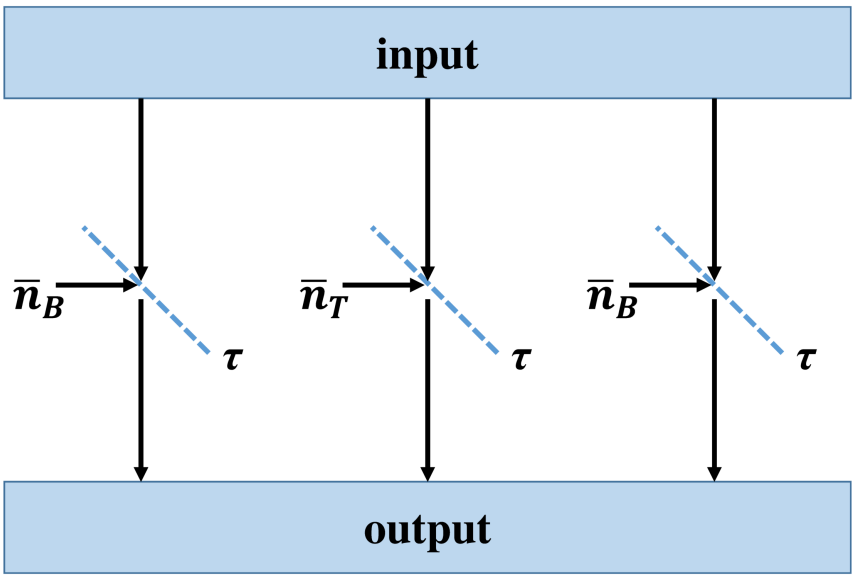


Figure 3.4: An example of the setup in the thermal loss case. Each thermal loss channel can be represented by a beamsplitter that mixes the input mode with an environmental thermal state. Thermal loss channels are parametrised by the transmissivity of the beamsplitter and the average photon number, $\bar{n}$, of the thermal state. We consider a sequence of thermal loss channels for which the beamsplitters all have the same transmissivity, $\tau$. One of the channels has a thermal state with a different average number of photons from the others; this is the target channel. The average number of photons in the thermal state of the target channel is denoted $\bar{n}_T$, whilst the average number of photons in the thermal state of the background channel is denoted $\bar{n}_B$. The task is to locate the target channel; in the case of this setup, it is the middle channel.

represented by a beamsplitter interaction with a thermal mode, and all of the beamsplitters have the same transmissivity, but the thermal mode with which the input modes interact is different for the target and background channels.

Any pair of one-mode, phase-insensitive, Gaussian channels with the same transmissivity is jointly teleportation covariant, using the Braunstein-Kimble (BK) protocol [83]. This means that both channels can be simulated using the same teleportation protocol, but with different resource states. In fact, using the BK protocol, a valid resource state for channel simulation is the asymptotic Choi matrix of the channel [99–101]. The Choi matrix of a channel is the output state when part of a maximally entangled state is passed through the channel. For bosonic systems, the maximally entangled state $\Phi$ is the limit for infinite squeezing of a sequence of TMSV states [55] $\Phi^a$,

i.e. $\Phi = \lim_a \Phi^a$, where $a$ is the level of squeezing and each $\Phi^a$ has covariance matrix (CM)

$$V_{\text{in}}^a = \begin{pmatrix} a\mathbb{I} & \sqrt{a^2 - \frac{1}{4}}\mathbb{Z} \\ \sqrt{a^2 - \frac{1}{4}}\mathbb{Z} & a\mathbb{I} \end{pmatrix}. \tag{3.47}$$

Therefore, the Choi matrix $\sigma_{\mathcal{E}}$ of a bosonic channel $\mathcal{E}$ is defined as the infinite-squeezing limit of a sequence of states $\{\sigma_{\mathcal{E}}^a\}$ where the generic element is given by a TMSV state partially propagated through the channel, i.e. $\sigma_{\mathcal{E}}^a := \mathcal{I} \otimes \mathcal{E}(\Phi^a)$. In the following, when we work with an asymptotic Choi matrix $\sigma_{\mathcal{E}}$ we implicitly mean that this is the limit of an underlying 'Choi sequence' $\{\sigma_{\mathcal{E}}^a\}$. Correspondingly, the teleportation simulation over $\sigma_{\mathcal{E}}$ is meant to be an asymptotic operation, where the simulation is defined over the Choi sequence $\{\sigma_{\mathcal{E}}^a\}$ after which the limit for infinite squeezing is taken [5]. Note that Gaussian states, which all elements of the sequence are, are completely described by their CM and their first moments vector. For states in the Choi sequence, all elements of the first moments vector are 0.

As previously mentioned (in Subsection 3.2.1), the error probability of any CPF protocol can be bounded using the fidelity between its possible outputs.

Since we can use the same teleportation protocol for both the target and the background channels, the entire discrimination protocol can be reduced, via stretching [2, 5, 30], to a single processor applied to different resource states (with the resource state depending on the position of the target channel). This adaptive-to-block reduction is shown in Fig. 3.5.

Since no trace preserving quantum operation can increase the distance between two quantum states (the fidelity of any two input states will be less than or equal to the fidelity of the resulting output states), the fidelity between the possible output states is lower bounded by the fidelity between the possible resource states. Let $\sigma_M^i$ be the resource state composed of $M(m-1)$ copies of the asymptotic Choi matrix of the background channel, $\sigma_B$, and $M$ copies of the asymptotic Choi matrix of the target channel, $\sigma_T$, arranged such that the $M$ copies of the asymptotic Choi matrix of the target channel is the $i$-th $2M$-mode subsystem. Note that each asymptotic Choi matrix consists of two modes. We can write

$$\sigma_M^i = P_{1i}\left[\sigma_T^{\otimes M} \otimes \sigma_B^{\otimes M(m-1)}\right], \tag{3.48}$$

where the operator $P_{1i}$ swaps the first $2M$-mode subsystem with the $i$-th $2M$-mode subsystem. We can then lower bound the fidelity of any pair of output states of a discrimination protocol with $M$ channel uses using

$$F(\rho_M^i, \rho_M^j) \geq F(\sigma_M^i, \sigma_M^j). \tag{3.49}$$

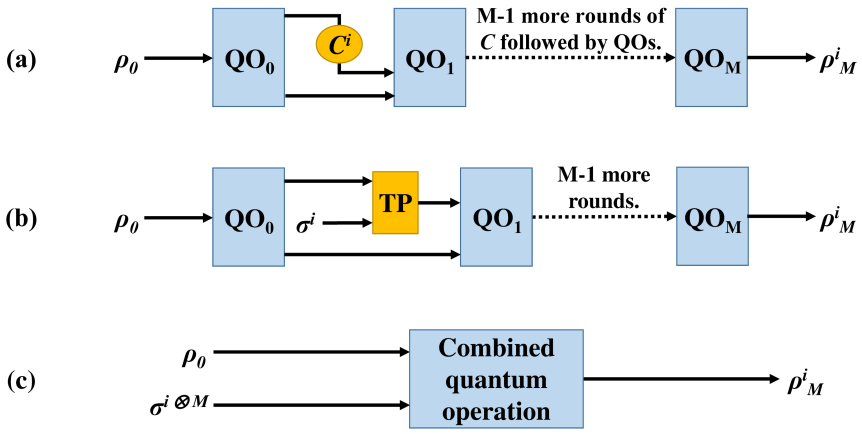


Figure 3.5: The reduction of a general adaptive discrimination protocol to a single round of quantum operations on a resource state. In panel (a), we have the most general discrimination protocol using $M$ uses of the sequence of channels. $\rho_0$ is some initial quantum state. We then apply some sequence of quantum operations (denoted by QO) interspersed with uses of the sequence of channels (denoted by $C^i$, where the label $i$ depends on the channel position). At each channel use, we may send a one-mode state through each of the channels in the sequence (and these modes are generally correlated with auxiliary modes that do not pass through the channels). Each round of quantum operations is allowed to be adaptive. This means that (i) entanglement can be present between ancillary modes of different quantum operations and (ii) measurements can be done on some subset of the modes and used to optimise following quantum operations. These measurements can always be delayed to the end of the protocol, by using controlled operations, so as to make all the QOs trace preserving. The final output of the adaptive protocol is denoted $\rho_0^i$; there are $m$ possible outputs depending on the channel position. Channel discrimination is then the task of discriminating between these $m$ different possible outputs, by means of an optimal collective quantum measurement (which may include all the delayed measurements). In panel (b), we simulate the channel with teleportation, using some teleportation protocol (TP) and a resource state ($\sigma^i$). Note that $\sigma^i$ is the resource state for the entire sequence of channels and is the tensor product of the resource states for teleportation of the $m-1$ background channels and the target channel, with the order of the subsystems determined by the label $i$. Note that neither the teleportation protocol nor the quantum operations depend on the label $i$ and so the entire discrimination protocol can be represented as some single fixed quantum operation on $\rho_0$ and $M$ copies of the resource state, $\sigma^i$. This representation is shown in panel (c).

Using the fact that each asymptotic Choi matrix in the resource is independent (i.e. using the tensor product structure of the resource states), we can write

$$F(\sigma_M^i, \sigma_M^j) = F(\sigma_T, \sigma_B)^{2M}, \tag{3.50}$$

for all $i \neq j$.

More precisely, since the asymptotic Choi matrices, $\sigma_T$ and $\sigma_B$, are defined by the infinite-squeezing limit of two sequences of output states, $\{\sigma_T^a\}$ and $\{\sigma_B^a\}$, the fidelity functional is computed over the elements of the sequences and then the limit is taken, i.e. $F(\sigma_T, \sigma_B) := \lim_a F(\sigma_T^a, \sigma_B^a)$. It is important to notice that the bound $F(\rho_M^i, \rho_M^j) \geq F(\sigma_T, \sigma_B)^{2M}$ holds for any generally adaptive protocol $\mathcal{P}$. Therefore, we may write

$$F_{i,j} := \inf_{\mathcal{P}} F(\rho_M^i, \rho_M^j) \geq F(\sigma_T, \sigma_B)^{2M}. \tag{3.51}$$

At the same time, we note that this lower bound is achievable by a block protocol $\mathcal{P}_{\text{block}}^a$ where $m$ copies of the tensor product state $\Phi^{a\otimes M}$ are prepared and each TMSV state $\Phi^a$ is used for the single-probing of $\mathcal{I} \otimes \mathcal{E}_{B/T}$, so that the quasi-Choi matrix $\sigma_{B/T}^a$ is generated at the output for measurement. It is easy to see that, in the limit of infinite squeezing $a \to \infty$, this protocol achieves the performance at the right hand side of Eq. (3.51), so that we may write

$$F_{i,j} = F(\sigma_T, \sigma_B)^{2M}, \text{ for any } i, j. \tag{3.52}$$

Let us optimise the error probability over all possible (generally adaptive) protocols $\mathcal{P}$. We define this optimal error probability as

$$p_{\text{err}}^{\text{opt}} = \inf_{\mathcal{P}} p_{\text{err}}; \tag{3.53}$$

it is the smallest achievable error probability for any discrimination protocol. As a consequence of the reasoning above and the inequalities in Eqs. (3.3) and (3.4), we can write

$$p_{\text{err}}^{\text{opt}} \geq \sum_{i>j}^{m} p_i p_j F(\sigma_T, \sigma_B)^{4M}, \tag{3.54}$$

$$p_{\text{err}}^{\text{opt}} \leq 2 \sum_{i>j}^{m} \sqrt{p_i p_j} F(\sigma_T, \sigma_B)^{2M}. \tag{3.55}$$

Let us now assume that each channel position is equally likely, and so $p_i = \frac{1}{m}$ for every value of $i$. We can then carry out the sums in Eqs. (3.54) and (3.55) and write

$$p_{\text{err}}^{\text{opt}} \geq \frac{m-1}{2m} F(\sigma_T, \sigma_B)^{4M}, \tag{3.56}$$

$$p_{\text{err}}^{\text{opt}} \leq (m-1) F(\sigma_T, \sigma_B)^{2M}. \tag{3.57}$$

### 3.3.2 Calculating the fidelity between Choi matrices

We now must calculate the fidelity between the (asymptotic) Choi matrices of the target and the background channels. A phase-insensitive, one-mode, Gaussian channel [55] can be parametrised by two parameters: its transmissivity, $\tau$, and its induced noise, $\nu$. It transforms the CM of an input two-mode state, $V_{in}$, with the transformation

$$V_{\text{in}} \to \left(\mathbb{I} \oplus \sqrt{\tau}\mathbb{I}\right) V_{\text{in}} \left(\mathbb{I} \oplus \sqrt{\tau}\mathbb{I}\right)^T + (0 \oplus \nu\mathbb{I}), \tag{3.58}$$

where $\mathbb{I}$ is the 2 by 2 identity matrix. There are three main classes of phase-insensitive, Gaussian channels that we must consider: thermal loss channels, thermal amplifier channels and additive noise channels. Loss and amplifier channels both have $\nu \geq \frac{|1-\tau|}{2}$ (recall that the vacuum noise is set to $\frac{1}{2}$), but loss channels have $0 \leq \tau < 1$, whilst amplifier channels have $1 < \tau$. Additive noise channels have $\nu \geq 0$ and $\tau = 1$.

Passing the second mode of a TMSV state $\Phi^a$ with an average photon number per mode of $\bar{n} = a - \frac{1}{2}$ through a phase-insensitive, Gaussian channel results in the state with CM

$$V_{\text{out}} = \begin{pmatrix} a\mathbb{I} & \sqrt{\tau\left(a^2 - \frac{1}{4}\right)}\mathbb{Z} \\ \sqrt{\tau\left(a^2 - \frac{1}{4}\right)}\mathbb{Z} & (a\tau + \nu)\mathbb{I} \end{pmatrix}, \tag{3.59}$$

where $\mathbb{Z}$ is the Pauli Z matrix.

The Bures fidelity of a pair of two-mode Gaussian states $\rho_i$ and $\rho_j$, with zero first moments and CM $V_i$ and $V_j$ is given by [76, 102]

$$F(\rho_i, \rho_j) = \frac{\sqrt{\chi} + \sqrt{\chi - 1}}{\sqrt[4]{\det\left(V_i + V_j\right)}}, \tag{3.60}$$

$$\chi = 2\sqrt{A} + 2\sqrt{B} + \frac{1}{2}, \tag{3.61}$$

$$A = \frac{\det\left(\Omega V_i \Omega V_j - \frac{1}{4}\mathbb{I}\right)}{\det\left(V_i + V_j\right)}, \tag{3.62}$$

$$B = \frac{\det\left(V_i + \frac{i}{2}\Omega\right)\det\left(V_j + \frac{i}{2}\Omega\right)}{\det\left(V_i + V_j\right)}, \tag{3.63}$$

$$\Omega = \mathbb{I} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.64}$$

Using this expression, we can calculate the fidelity of a pair of output states of phase-insensitive, Gaussian channels (when the input state is a TMSV) with the same transmissivity.

In the case of thermal loss and amplifier channels, we define $\epsilon_T = \frac{\nu_T}{|1-\tau|}$ and $\epsilon_B = \frac{\nu_B}{|1-\tau|}$, where $\nu_T$ is the induced noise of the target channel, $\nu_B$ is the induced noise of the background

channels, and $\tau$ is the transmissivity of all of the channels in the sequence. In fact, $\epsilon_T$ and $\epsilon_B$ give us the mean photon number of the environment for each channel, via the equation

$$\bar{n}_{T(B)} = \epsilon_{T(B)} - \frac{1}{2}. \tag{3.65}$$

We find that the fidelity of the outputs of two such thermal loss or amplifier channels is analytically given by

$$F_{\text{loss/amp}}(\tau, \epsilon_T, \epsilon_B, a) = \frac{\sqrt{2}\left(\sqrt{\alpha+\beta}+\sqrt{\alpha-\beta}\right)}{\beta}, \tag{3.66}$$

where we define

$$\begin{aligned}\alpha =& \Big(4\epsilon_T\epsilon_B + 4a^2(4\epsilon_T\epsilon_B+1) \\ &+(4a^2-1)\sqrt{(4\epsilon_T^2-1)(4\epsilon_B^2-1)}\Big)|1-\tau|^2 \\ &+ 8a(\epsilon_T+\epsilon_B)\tau|1-\tau| + (1+\tau)^2,\end{aligned} \tag{3.67}$$

$$\beta = 4\left(\tau + 2a(\epsilon_T+\epsilon_B)|1-\tau|\right). \tag{3.68}$$

Taking the limit of this expression as $a \to \infty$, in order to obtain the fidelity between the Choi matrices, we get

$$F^{\infty}_{\text{loss/amp}}(\epsilon_T, \epsilon_B) = \frac{\sqrt{4\epsilon_T\epsilon_B + 1 + \sqrt{(4\epsilon_T^2-1)(4\epsilon_B^2-1)}}}{\sqrt{2}(\epsilon_T+\epsilon_B)}. \tag{3.69}$$

Note that we no longer have any explicit dependence on $\tau$.

Thus, our discrimination bounds for thermal loss or amplifier channels become

$$p^{\text{opt}}_{\text{err}} \geq \frac{m-1}{2m}(F^{\infty}_{\text{loss/amp}}(\epsilon_T, \epsilon_B))^{4M}, \tag{3.70}$$

$$p^{\text{opt}}_{\text{err}} \leq (m-1)(F^{\infty}_{\text{loss/amp}}(\epsilon_T, \epsilon_B))^{2M}. \tag{3.71}$$

The upper bound in Eq. (3.71) can become larger than the error probability for randomly guessing the position of the target channel, which is given by $\frac{m-1}{m}$. We can combine these two upper bounds to get

$$p^{\text{opt}}_{\text{err}} \leq (m-1)\min\{m^{-1}, (F^{\infty}_{\text{loss/amp}}(\epsilon_T, \epsilon_B))\}. \tag{3.72}$$

In order to investigate the behaviour of $F^{\infty}_{\text{loss/amp}}$, we re-parametrise Eq. (3.69) in terms of the mean of $\epsilon_T$ and $\epsilon_B$, i.e.

$$\epsilon_{\text{av}} = \frac{\epsilon_T + \epsilon_B}{2}, \tag{3.73}$$

and the absolute value of their difference, i.e.

$$\epsilon_{\text{dif}} = |\epsilon_T - \epsilon_B|. \tag{3.74}$$

Differentiating with regard to $\epsilon_{\text{dif}}$, we get a negative semi-definite function and differentiating with regard to $\epsilon_{\text{av}}$, we get a positive semi-definite function. This means that either increasing the difference in the average number of photons between the target and background channels (whilst keeping the mean fixed) or decreasing the mean of the $\epsilon$-values, whilst keeping the difference fixed, will decrease the minimum fidelity of the output states.

We now consider the case of additive noise channels. We find that the fidelity of the outputs of two such channels becomes

$$F_{\text{add}}(\nu_T, \nu_B, a) = \frac{2a\sqrt{\nu_T \nu_B} + \sqrt{(2a\nu_T + 1)(2a\nu_B + 1)}}{(2a(\nu_T + \nu_B) + 1)}. \tag{3.75}$$

Taking the limit of this expression as $a \to \infty$, we get

$$F^{\infty}_{\text{add}}(\nu_T, \nu_B) = \frac{2\sqrt{\nu_T \nu_B}}{\nu_T + \nu_B}. \tag{3.76}$$

We can again substitute this expression into Eqs. (3.56) and (3.57). Our discrimination bounds for additive noise channels become

$$p^{\text{opt}}_{\text{err}} \geq \frac{m-1}{2m}(F^{\infty}_{\text{add}}(\nu_T, \nu_B))^{4M}, \tag{3.77}$$

$$p^{\text{opt}}_{\text{err}} \leq (m-1)(F^{\infty}_{\text{add}}(\nu_T, \nu_B))^{2M}. \tag{3.78}$$

We now investigate the behaviour of $F^{\infty}_{\text{add}}$ by re-parametrising Eq. (3.76) in terms of $\nu_{\text{av}}$ and $\nu_{\text{dif}}$, where $\nu_{\text{av}}$ is the mean of $\nu_T$ and $\nu_B$ and $\nu_{\text{dif}}$ is the absolute value of the difference between them. Note that $\nu_{\text{dif}} \leq 2\nu_{\text{av}}$. We can then rewrite Eq. (3.76) as

$$F^{\infty}_{\text{add}}(r) = \sqrt{1 - \frac{r^2}{4}}, \;\; r = \frac{\nu_{\text{dif}}}{\nu_{\text{av}}}. \tag{3.79}$$

Thus, we can see that the fidelity between the Choi matrices of two additive noise channels depends only on the ratio of $\nu_{\text{dif}}$ to $\nu_{\text{av}}$. Differentiating with regard to $r$, we see that the fidelity decays as $r$ increases.

### 3.3.3 Classical limits

Let us define a classical protocol as a non-adaptive protocol that restricts the states sent through the sequence of channels to an arbitrary mixture of coherent states. Since the Gaussian channels we are considering are phase-insensitive and since both the target and the background channels have the same transmissivity, enacting a phase-shift or displacement on the input states sent through the channels cannot affect the fidelity of the output states (since these unitary operations commute with the channels). The joint concavity of the Bures fidelity and the linearity of the channels means

that the optimal classical input state (to minimise the fidelity between output states) is a single coherent state (not a mixture). As a result, the classical discrimination protocol that minimises the lower bound on the error probability sends vacuum states through the channel at each channel use. This means that such protocols use only the passive signature of the channels.

We can obtain expressions for the minimum fidelity between output states for classical protocols by using our expressions for the fidelity between the output states using TMSV inputs in Eqs. (3.66) and (3.75) and setting $a = \frac{1}{2}$. This gives us the fidelity between the output states of the channels when the input state is a vacuum state.

In the case of thermal loss and amplifier channels, the minimum classical fidelity between output states is

$$F^{\text{class}}_{\text{loss/amp}}(\tau, \epsilon_T, \epsilon_B) = \frac{\sqrt{\gamma+\delta}+\sqrt{\gamma-\delta}}{\delta}, \tag{3.80}$$

where we define

$$\gamma = 4\epsilon_T\epsilon_B|1-\tau|^2 + 2(\epsilon_T+\epsilon_T)\tau|1-\tau| + (1+\tau^2), \tag{3.81}$$

$$\delta = 2\left(\tau + (\epsilon_T+\epsilon_T)|1-\tau|\right). \tag{3.82}$$

In the case of additive noise channels, the minimum classical fidelity between output states is

$$F^{\text{class}}_{\text{add}}(\nu_T, \nu_B) = \frac{1}{\sqrt{(\nu_T+1)(\nu_B+1)}-\sqrt{\nu_T\nu_B}}. \tag{3.83}$$

We can now give upper and lower bounds on the error of classical discrimination protocols. We write

$$p^{\text{class}}_{\text{err}} \geq \frac{m-1}{2m}(F^{\text{class}})^{4M}, \tag{3.84}$$

$$p^{\text{class}}_{\text{err}} \leq (m-1)(F^{\text{class}})^{2M}, \tag{3.85}$$

where the fidelity function is given by either Eq. (3.80) or Eq. (3.83), depending on the class of channel.

### 3.3.4 Quantum advantage

We say that there is a quantum advantage if we can show that there exists some quantum discrimination protocol that gives a lower probability of error than any classical protocol. In order to prove a quantum advantage for CPF, we need to show that the lower bound on the error of classical protocols is larger than the upper bound on the error of all protocols. In other words, we must show that

$$\frac{m-1}{2m}(F^{\text{class}})^{4M} \geq (m-1)(F^{\infty})^{2M}. \tag{3.86}$$

This is equivalent to showing

$$2M \ln\left(\frac{(F^{\text{class}})^2}{F^{\infty}}\right) \geq \ln(2m). \tag{3.87}$$

Noting that $\ln(2m) > 0$, since $m \geq 2$, we can see that the condition in Eq. (3.87) will always be met for sufficiently large $M$ (number of probes) as long as the condition

$$(F^{\text{class}})^2 > F^{\infty} \tag{3.88}$$

holds. Whether this condition is met depends only on the parameters of the target and background channels. Note that even if this condition is not met, it does not mean there is no quantum advantage; it could be the case that the bounds are not tight. In fact, in Subsection 3.3.5, we provide alternative bounds which can potentially show quantum advantage even in cases in which the condition in Eq. (3.88) is not met.

Unlike $F^{\infty}_{\text{loss/amp}}$, the fidelity $F^{\text{class}}_{\text{loss/amp}}$ depends on the transmissivity $\tau$. In fact, differentiating, we find that $\frac{dF}{d\tau} \geq 0$ for $0 \leq \tau < 1$ and that $\frac{dF}{d\tau} \leq 0$ for $\tau > 1$. Further, as $\tau \to 0$, we have $F^{\text{class}}_{\text{loss/amp}} \to F^{\infty}_{\text{loss/amp}}$. This can be intuitively understood, since the entire channel discrimination process, including the coupling of the signal mode with the environment, can be regarded as a (generalised) measurement on the environmental modes. Thus, no matter how much entanglement the interacting modes have, the possible output states that the final measurement distinguishes between cannot have a lower (pairwise) fidelity than the possible configurations of environmental modes that are being discriminated between. In other words, the infinite squeezing case is equivalent to a direct measurement on the environmental modes before they are mixed with the signal states, whilst, in any finite energy scenario, we send signal states to interact with the environmental modes and then measure the signal states. Since the $\tau = 0$ case corresponds to the signal states being completely replaced by the environmental modes, the classical protocol, in this case, is also a direct measurement on the environmental modes. Consequently, in the case of thermal loss channels, for all values of $\epsilon_T$ and $\epsilon_B$, there is some threshold value of $\tau$ such that channels with $\tau$ below the threshold do not meet the condition in Eq. (3.88). Setting $\tau = \frac{1}{2}$, we find that $\frac{(F^{\text{class}})^2}{F^{\infty}} \leq 1$, and hence the inequality in Eq. (3.88) does not hold for any channel ensemble with $\tau \leq \frac{1}{2}$. See Appendix A and the supplementary Mathematica files of Ref. [9] for more details.

Fig. 3.6 illustrates the region in which we meet the condition in Eq. (3.88) (and so can prove a quantum advantage for some number of probes), in the case of thermal loss channels, for a few choices of transmissivity, $\tau$. The plot is in terms of $\epsilon_{\text{dif}}$ and $\epsilon_{\text{av}}$, as defined in Eqs. (3.73)-(3.74). We see that higher transmissivities result in a larger region in which we can prove a quantum

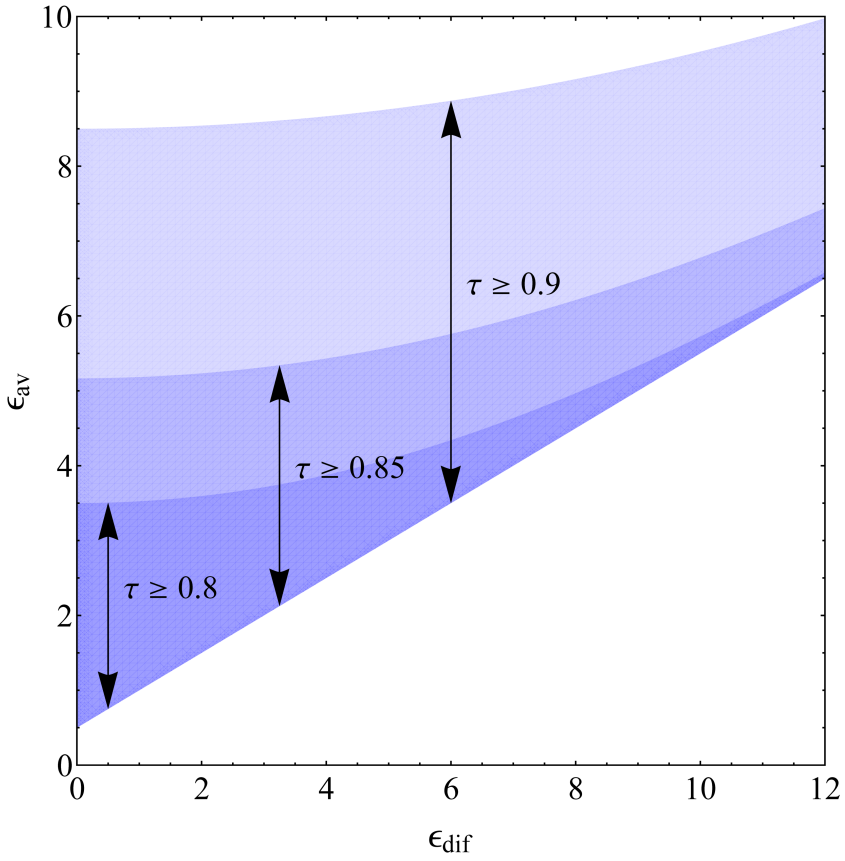


Figure 3.6: Regions in which we can prove a quantum advantage for thermal loss channels, as a function of their noise difference $\epsilon_{\text{dif}}$ and mean noise $\epsilon_{\text{av}}$, for different values of the transmissivity $\tau$. Note that the region for a higher value of $\tau$ completely contains the region for any lower value of $\tau$. The minimum value of $\epsilon_{\text{av}}$ for fixed $\epsilon_{\text{dif}}$ is $\frac{\epsilon_{\text{dif}}+1}{2}$, since neither $\epsilon_T$ nor $\epsilon_B$ can be less than $\frac{1}{2}$.

advantage. Further, as $\epsilon_{\text{dif}}$ increases, the region in which we can prove quantum advantage narrows (in terms of the allowed values of $\epsilon_{\text{av}}$).

The condition for the inequality in Eq. (3.88) to hold takes a simple form for additive noise channels. We again re-parametrise in terms of $\nu_{\text{av}}$ and $\nu_{\text{dif}}$. We can then write the condition purely in terms of $\nu_{\text{av}}$. Thus, we find that for a sequence of additive noise channels, we will always have a quantum advantage for some number of probes as long as

$$\nu_{\text{dif}} > \frac{\sqrt{32\nu_{\text{av}}^4 - 8\nu_{\text{av}}^2 - 8\nu_{\text{av}} - 1 - (4\nu_{\text{av}}+1)\sqrt{8\nu_{\text{av}}+1}}}{2\sqrt{2}\nu_{\text{av}}}. \tag{3.89}$$

### 3.3.5 Bounds from specific protocols

We can consider specific discrimination protocols; these can provide benchmarks for both the classical (entanglement-free) and entangled cases. In the classical case, we have vacuum input. In

this case, the return state is thermal, therefore a photon counting measurement coupled with the maximum-likelihood estimation (MLE) gives the Helstrom performance [103]. In this protocol, we carry out photon counting on each of the return states, and simple derivation shows that the MLE decision rule reduces to choosing the channel with the maximum/minimum photon count, i.e. we estimate the target channel to be

$$\arg\max_s N_s, \text{if } \bar{n}_T > \bar{n}_B, \tag{3.90}$$

and

$$\arg\min_s N_s, \text{if } \bar{n}_T < \bar{n}_B, \tag{3.91}$$

where $s$ is an index labelling the channels in the sequence and $N_s$ denotes the total number of photons counted from the return states of channel $s$ (cumulatively, over all M channel uses).

We can consider a similar protocol involving entanglement, in the cases of thermal loss and amplifier channels. In these cases, we can get thermal return states by sending TMSV states through the channels, carrying out anti-squeezing operations on the return states and then tracing over one of the two modes. For each probe sent through one of the channels, we start by carrying out two-mode squeezing on a pair of vacuum modes, with squeezing parameter

$$r_0 = \frac{1}{2}\ln\left(2a + \sqrt{4a^2 - 1}\right). \tag{3.92}$$

This results in the TMSV state $\Phi^a$, which has an average photon number per mode of $\bar{n} = a - \frac{1}{2}$ and the CM given by Eq. (3.47). The first mode is kept as an idler, whilst the second mode is passed through the channel. Each individual channel output state will then have a CM of the form in Eq. (3.59); we then carry out two-mode squeezing on the state, with squeezing parameter

$$r_1 = \frac{1}{2}\ln\left(\frac{|1 - \sqrt{\tau}|}{1 + \sqrt{\tau}}\right). \tag{3.93}$$

For a thermal loss channel, we discard the idler mode; the resulting state has the CM

$$V^a_{\text{ret,loss}} = \text{Disc}_1\left[S(r_1)V^a_{\text{out,loss}}S^T(r_1)\right] \tag{3.94}$$

$$= \frac{\nu + 2a\tau - \tau\sqrt{4a^2 - 1}}{|1 - \tau|}\mathbb{I}, \tag{3.95}$$

where $S$ is the two-mode squeezing matrix, given by

$$S(r) = \begin{pmatrix} \cosh(r)\mathbb{I} & \sinh(r)\mathbb{Z} \\ \sinh(r)\mathbb{Z} & \cosh(r)\mathbb{I} \end{pmatrix}, \tag{3.96}$$

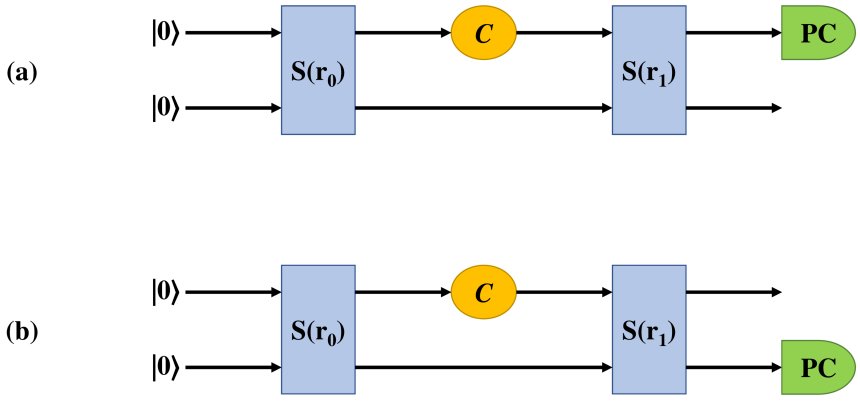


Figure 3.7: The setup for a CPF protocol that provides a benchmark for the general quantum case. In panel (a), we have the protocol for the thermal loss case and in panel (b), we have the protocol for the thermal amplifier case. In both cases, we begin by carrying out two-mode squeezing on a vacuum state, with squeezing parameter $r_0$, as given in Eq. (3.92). This is denoted $\mathrm{S}(\mathrm{r}_0)$. We then pass one of the modes through the channel, denoted $C$, and then carry out two-mode squeezing again, this time with squeezing parameter $r_1$. Finally, we carry out a photon counting measurement (denoted PC) on one of the modes and trace over the other mode. This process is repeated $M$ times (where $M$ is the number of probes used) for every channel in the sequence. Note that in the thermal loss case, the measurement is carried out on the channel mode, whilst in the thermal amplifier case, the measurement is carried out on the idler mode.

and where $\mathrm{Disc}_1$ indicates that we discard the first (idler) mode. We can get a return state with the same form for an amplifier channel by carrying out the same process, but tracing over the other mode (the mode which passed through the channel). In other words, we have

$$V^a_{\mathrm{ret,amp}} = \mathrm{Disc}_2\left[S(r_1)V^a_{\mathrm{out,amp}}S^T(r_1)\right] \tag{3.97}$$

$$= \frac{\nu + 2a\tau - \tau\sqrt{4a^2-1}}{|1-\tau|}\mathbb{I}. \tag{3.98}$$

This protocol is illustrated in Fig. 3.7.

We now note that the CM in Eq. (3.95) has finite energy, even in the limit of infinite squeezing ($a \to \infty$). Letting $V^\infty_{\mathrm{ret},T(B)}$ be the asymptotic return state from the target (background) channel (for either a thermal loss or a thermal amplifier channel), we find that

$$V^\infty_{\mathrm{ret},T(B)} = \frac{\nu_{T(B)}}{|1-\tau|}\mathbb{I} = \epsilon_{T(B)}\mathbb{I}. \tag{3.99}$$

Hence, we can get thermal return states even in the case of infinite entanglement. Note that these are the same return states we would get in the classical case if the channels had a transmissivity of 0. Note too that we cannot enact this protocol in the additive noise case, since our expression

in Eq. (3.93) for the squeezing parameter $r_1$ diverges as $\tau \to 1$. We can then carry out photon counting measurements on the return states and estimate the target channel using the MLE.

We now calculate the success probability of the MLE. The probability that a thermal mode with average photon number $\bar{n}$ is measured to have $k$ photons is given by

$$P_{\bar{n}}(k) = \frac{\bar{n}^k}{(\bar{n}+1)^{k+1}}. \tag{3.100}$$

We then calculate the probability that $M$ thermal modes, with the same average photon number of $\bar{n}$, are measured to have a total of $k$ photons, by replacing the thermal distribution with a sum of independent and identically distributed (iid) thermal distributions. We find that this probability is given by

$$P_{\bar{n},M}(k) = \binom{k+M-1}{k} \left(\frac{\bar{n}}{1+\bar{n}}\right)^k \left(\frac{1}{1+\bar{n}}\right)^M, \tag{3.101}$$

where the binomial coefficient accounts for the different ways in which the photons can be distributed across the measured modes. From this we can calculate the probability that the $M$ modes are measured to have fewer than $n_c$ photons in total:

$$\mathrm{pr}_{\bar{n},M}(\mathrm{count} < n_c) = \sum_{k=0}^{n_c-1} P_{\bar{n},M}(k). \tag{3.102}$$

Let us first consider the case in which $\bar{n}_T > \bar{n}_B$. In this case the MLE gives the correct answer when all of the background channels have return states that are measured to have fewer photons than those of the target channel. We must also consider the possibility that the return states of one or more of the background channels are measured to have the same number of photons as the return states of the target channel (but not more). In this case, we choose randomly between the channels that gave the highest photon counts. This gives a total success probability (for the entangled case) of

$$\begin{aligned} p^{\mathrm{MLE}}_{\mathrm{succ},\bar{n}_T>\bar{n}_B} = &\sum_{c=1}^{m} \frac{1}{c} \sum_{n_c=0}^{\infty} \left[\mathrm{pr}_{\bar{n}_B,M}(\mathrm{count} < n_c)\right]^{m-c} \\ &\times P_{\bar{n}_T,M}(n_c) \binom{m-1}{c-1} (P_{\bar{n}_B,M}(n_c))^{c-1}. \end{aligned} \tag{3.103}$$

Here, the index $c$ is the number of channels with the same photon count (hence, $c = 1$ is the case in which all of the background channels give a lower photon count than the target channel). The factor of $\frac{1}{c}$ comes from the random choice when multiple channels give the same photon count. Note that in the case of $n_c = 0$, the only non-zero contribution is in the case $c = m$, corresponding to a photon count of 0 for the target and all of the background channels. If this occurs, there is a

$\frac{1}{m}$ chance of the target channel being randomly guessed correctly. In this case, we define

$$\mathrm{pr}_{\bar{n}_B,M}(\mathrm{count} < 0)^0 = 1. \tag{3.104}$$

Extension to the case in which $\bar{n}_T < \bar{n}_B$ can be done trivially, by writing

$$\mathrm{pr}_{\bar{n},M}(\mathrm{count} > n_c) = 1 - \mathrm{pr}_{\bar{n},M}(\mathrm{count} < n_c + 1). \tag{3.105}$$

Then we have a success probability of

$$\begin{aligned} p^{\mathrm{MLE}}_{\mathrm{succ},\bar{n}_T<\bar{n}_B} = &\sum_{c=1}^{m} \frac{1}{c} \sum_{n_c=0}^{\infty} \left[\mathrm{pr}_{\bar{n}_B,M}(\mathrm{count} > n_c)\right]^{m-c} \\ &\times P_{\bar{n}_T,M}(n_c) \binom{m-1}{c-1} (P_{\bar{n}_B,M}(n_c))^{c-1}. \end{aligned} \tag{3.106}$$

In both cases, the error probability is given by

$$p^{\mathrm{MLE}}_{\mathrm{err}} = 1 - p^{\mathrm{MLE}}_{\mathrm{succ}}. \tag{3.107}$$

Note that for the classical MLE error probabilities, we simply substitute $\bar{n}_{T(B)}$ with the average photon numbers of the classical return states, i.e. $\bar{n}_{T(B)}|1-\tau|$.

This quantity can be easily numerically calculated. Using this semi-analytic benchmark, we can show a quantum advantage with a lower value of $M$ than is required for the condition in Eq. (3.87) to be met. This is demonstrated in Fig. 3.8. It is also useful as it is based on a protocol that can be easily implemented.

The scaling of the MLE error with the number of subsystems is of interest. We can upper bound the error in the case of $m$ subsystems in terms of the success probability for 2 subsystems, which we will call $p^{\mathrm{MLE}}_{\mathrm{succ},2}$. The error probability for $m$ subsystems then obeys the inequality

$$p^{\mathrm{MLE}}_{\mathrm{err},m} \leq 1 - (p^{\mathrm{MLE}}_{\mathrm{succ},2})^{m-1} = 1 - \left(1 - p^{\mathrm{MLE}}_{\mathrm{err},2}\right)^{m-1}, \tag{3.108}$$

since the target channel having a higher photon count than one background channel cannot decrease the probability that it will have a higher photon count than a different background channel. In fact, this bound is an overestimate for any $m > 2$, since the conditional probability that the target channel has a higher photon count than one background channel, given that it has a higher photon count than a different background channel, is more than $p^{\mathrm{MLE}}_{\mathrm{succ},2}$. This can be understood by considering the iid outcomes of 3 (6-sided) dice rolls denoted $a$, $b$ and $c$. The probability that $a > b$ is the same as the probability that $a > c$ and is equal to $\frac{5}{12}$, however the probability that $a > c$ given that $a > b$ is more than $\frac{5}{12}$, since the condition makes it less likely that $a$ is a small

number and more likely that $a$ is a large number. Expanding the inequality in Eq. (3.108) to the first order in $p_{\mathrm{err},2}^{\mathrm{MLE}}$, we get

$$p_{\mathrm{err},m}^{\mathrm{MLE}} \leq (m-1)p_{\mathrm{err},2}^{\mathrm{MLE}}. \tag{3.109}$$

This inequality is strict for $m > 2$. This means that the MLE error scales more slowly with $m$ than the upper bound in Eq. (3.57), which is based on the PGM. However, for some sets of channel parameters, the upper bound in Eq. (3.109) can be close to the actual value of $p_{\mathrm{err},m}^{\mathrm{MLE}}$.

It is also of note that, whilst the bounds based on the fidelity are symmetric under the exchange of $\nu_T$ and $\nu_B$, the MLE bound is not (for more than two subsystems). Thus, using this protocol in one of our applications, we may achieve a different error probability for finding a single cold pixel in a hot background than for finding a single hot pixel in a cold background.

### 3.3.6 Applications of the bounds

Let us consider some physical applications of these bounds. One possible scenario in which one may need to discriminate between various channels with the same transmissivity is thermal imaging. The sequence of channels could represent a sequence of pixels that is being probed with microwave or infrared radiation, where we know that one pixel is hotter (or colder) than its surroundings and want to know its location. Alternatively, we could be imaging a surface with a microscope and want to find the frequency at which a source on the surface is emitting radiation. The different channels would then represent different frequencies. These tasks can both be modelled as a CPF task over a sequence of thermal loss channels with the same transmissivity.

In Fig. 3.8, we consider an imaging task, in which a colder pixel must be located from a sequence of $9$ pixels, each of which has an area, $A$, of $4000\ \mu\mathrm{m}^2$. We consider a case in which imaging is carried out in the microwave range (with a wavelength of $1$ mm), with high transmissivity, a background temperature of $\sim -0.39°\mathrm{C}$ and a target temperature of $\sim -25.59°\mathrm{C}$. We assume that our detectors are very close to the pixels and that our imaging pulses have a time duration, $t$, of $100$ ns. We also assume that the pulses are transform-limited (meaning that they have the minimum possible time-bandwidth product) and so set the bandwidth of detection to $2.5$ MHz. This is in line with the fact that a transform-limited pulse has a time-bandwidth product (in terms of the variances) of $\frac{1}{4}$ [104].

We find the mean photon numbers by calculating the induced noise, which is independent of the transmissivity. Planck's law states that the spectral radiance of a black body, at a frequency $f$,

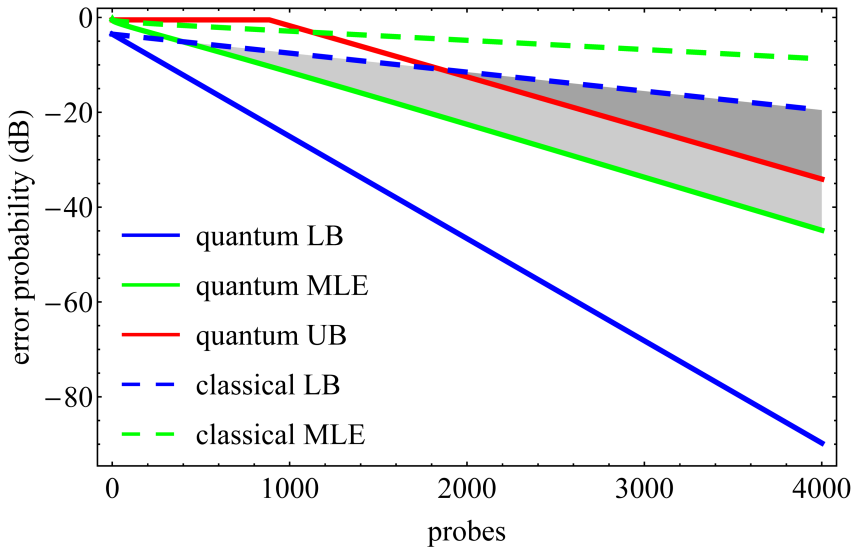


Figure 3.8: Error probability in decibels (dB), $10\log_{10}(p_{\text{err}})$, as a function of the number of the probes per pixel, for a thermal imaging task in which a sequence of $m = 9$ pixels, each of area $4000\ \mu\text{m}^2$, is probed using microwaves (with wavelength 1 mm). The transmissivity of each pixel is 0.99 and the goal is finding the one pixel at temperature 247.56 K ($-25.59$°C, $\epsilon_T = 21$) in a background of pixels at temperature 272.76 K ($-0.39$°C, $\epsilon_B = 23.2$). Lower and upper bounds on the error probability are given for general quantum protocols (labelled "quantum LB" and "quantum UB") and a lower bound on the error is given for classical protocols (labelled "classical LB"), for differing numbers of states sent through the channels (probes). Benchmarks based on the MLE are also shown for both the quantum and the classical cases (labelled "quantum MLE" and "classical MLE"). For the quantum upper bound, we use the expression in Eq. (3.72). For a large number of probes (in this case, greater than or equal to 1854), the upper bound on the error of quantum protocols is smaller than the lower bound on the error of classical protocols, proving we have a quantum advantage (in the darker shaded area). However, a much smaller number of probes (396) is required for the bound based on the MLE in the quantum case to beat the classical lower bound, and hence we are able to show a quantum advantage for any number of probes greater than 395 (in the lighter shaded area).

is given by

$$R(f,T) = \frac{2hf^3}{c^2(e^{\frac{hf}{kT}} - 1)}, \tag{3.110}$$

where $c$ is the speed of light, $h$ is Planck's constant, $k$ is the Boltzmann constant, and $T$ is the temperature of the pixel. By dividing $R$ by $hf$, we obtain the number of photons emitted per unit time, per unit area of the pixel into an infinitesimal frequency range and into a unit solid angle.

We must then integrate $\frac{R}{hf}$ over the bandwidth of the detector and multiply it by the duration of the imaging pulse, $t$, the solid angle over which the detector collects photons, $\omega$, and the area of the pixels, $A$, in order to obtain the induced noise, $\nu$. We therefore write

$$\nu_{B/T} = A\omega t \int_{f_{\min}}^{f_{\max}} \frac{2f^2}{c^2(e^{\frac{hf}{kT_{B/T}}} - 1)} df, \tag{3.111}$$

where $T_{B/T}$ is the temperature of the background/target pixel and $f_{\min/\max}$ is the minimum/maximum frequency in our frequency range. We set $\omega = 2\pi$ (i.e. we assume that the detector collects all light emitted in one hemisphere normal to the surface of the pixel). This is justified by our assumption that the detector is close to the pixels. If the detector were further away, we could adjust $\omega$ accordingly (and may have to reduce the transmissivity, $\tau$). Dividing $\nu_B$ and $\nu_T$ by $|1 - \tau|$ gives the values of $\epsilon_B$ and $\epsilon_T$ respectively.

Note that, for the bounds based on fidelity, swapping $\epsilon_T$ and $\epsilon_B$ does not affect the calculations, so these would be the same if the task were to find a target pixel at temperature $-0.39$°C in a background of pixels at $-25.59$°C. This is not the case for the benchmark based on the MLE. From Fig. 3.8, we see that we can prove a quantum advantage for a large number of channel uses (probes). We also see that the (quantum) MLE bound enables us to show a quantum advantage at a much lower value of $M$ than the fidelity-based quantum upper bound.

Before considering the next example, it is also worth noting that it is likely that the classical lower bound (blue dashed) in Fig. 3.8 is not tight, since we see a gap between it and the classical MLE performance (green dashed). Therefore quantum advantage is likely to hold for any number of probes, since we see that the quantum MLE (green solid) beats the classical MLE (green dashed) for any number of probes. A future study might be able to prove such a quantum advantage.

Another scenario in which one may wish to discriminate between thermal loss channels with different noises could arise in quantum communications. One may know that one of a sequence of communications lines has a higher excess noise than the others, perhaps due to the presence of an eavesdropper, and may wish to localise the eavesdropper by finding the channel with the higher excess noise.

This scenario is illustrated in Fig. 3.9, where we consider transmission over communication lines with a loss of 10 dB. Excess noise is expressed in dimensionless vacuum noise units and is defined in terms of the transmissivity and the thermal number of the channel as $\epsilon = \tau^{-1}(1 - \tau)\bar{n}$ [20]. We consider background excess noises of 0.01 and an excess noise for the eavesdropper of 0.1. In this case, we cannot prove a quantum advantage, although the quantum lower bound is lower than the classical lower bound. This is in accordance with the fact that we cannot meet the

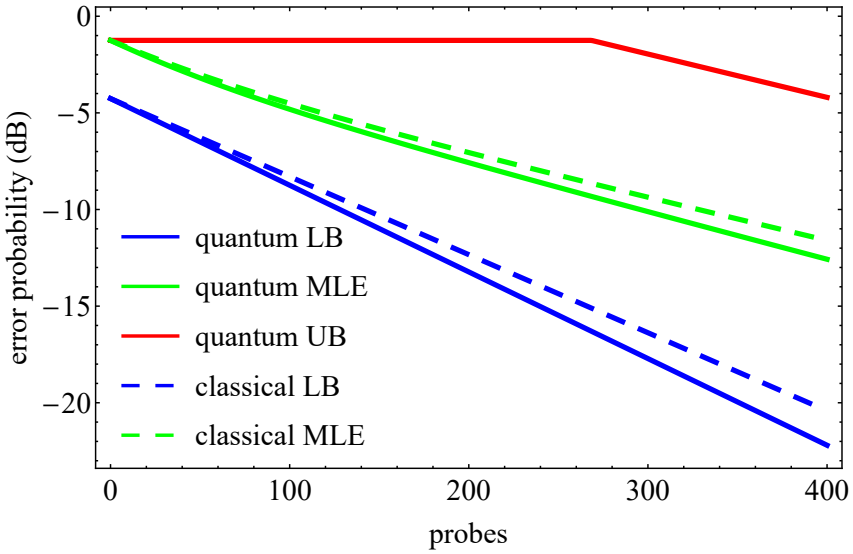


Figure 3.9: Error probability in decibels versus number of probes per communication line for the problem of eavesdropper localisation. We consider a transmissivity of 0.1, corresponding to a loss of 10 dB. The background channels have an excess noise of 0.01, whilst the channel with the eavesdropper has an excess noise of 0.1. Lower and upper bounds on the error probability are given for general quantum protocols (labelled "quantum LB" and "quantum UB") and a lower bound on the error is given for classical protocols (labelled "classical LB"). Benchmarks based on the MLE are shown for both the quantum and the classical cases (labelled "quantum MLE" and "classical MLE"). In this case, the quantum upper bound never goes below the classical upper bound, so we are not able to prove a quantum advantage.

condition in Eq. (3.88) with any channel ensemble that has $\tau \leq \frac{1}{2}$. The quantum MLE benchmark is also lower than the classical MLE benchmark, but does not go below the classical lower bound. This is again likely to be caused by the classical lower bound not being tight.

Another possibility is that we could have a multi-mode cable with multiple frequency channels and wish to find a channel with lower noise than the others. This is another case of discrimination between a sequence of thermal loss channels with different noises. If the transmissivity is high enough (for instance, for a short-range cable) we could potentially also model this scenario as a sequence of additive noise channels.

Fig. 3.10 illustrates this situation. We consider a sequence of 100 additive noise channels and want to find the channel with the lower induced noise. The background channels have an induced noise of 0.03 and the target channel has an induced noise of 0.01. We can show a quantum advantage for a number of probes greater than or equal to 20. Note that, whilst we can provide

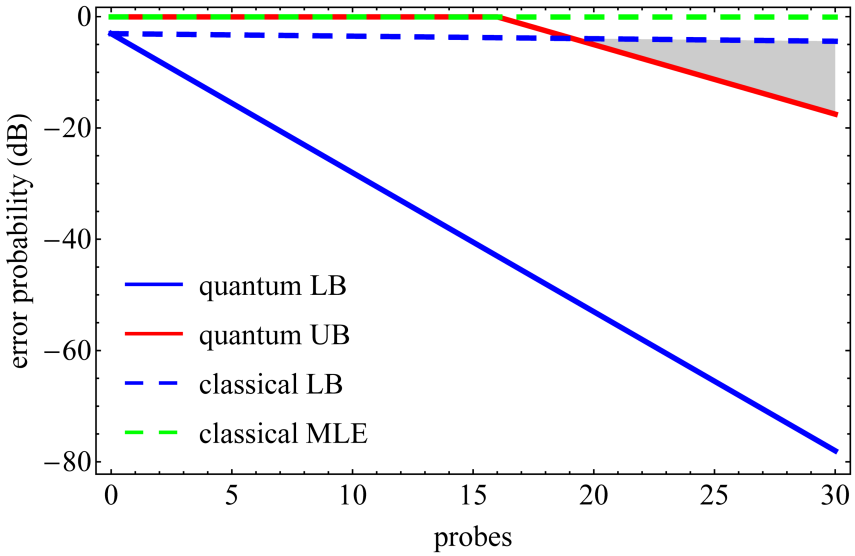


Figure 3.10: Error probability in decibels versus number of probes per channel for the problem of additive noise localisation. We want to find the channel with the lower induced noise from a sequence of 100 additive-noise channels. The background channels have an induced noise of 0.03, whilst the target channel has an induced noise of 0.01. Lower and upper bounds on the error probability are given for general quantum protocols (labelled "quantum LB" and "quantum UB") and a lower bound on the error is given for classical protocols (labelled "classical LB"). The benchmark based on the MLE is shown for the classical case (labelled "classical MLE"). For a number of probes greater than or equal to 20, the upper bound on the error of quantum protocols is smaller than the lower bound on the error of classical protocols, proving we have a quantum advantage (in the shaded area).

a classical benchmark based on the MLE, we cannot provide a quantum MLE benchmark in the additive noise case. This is due to the fact that the squeezing parameter in Eq. (3.93) diverges as $\tau \to 1$, meaning that the protocol shown in Fig. 3.7 cannot be enacted in the additive noise case.

## 3.4 Summary

In this chapter, we considered the task of CPF, both on a sequence of pure loss channels and on a sequence of phase-insensitive Gaussian channels with fixed transmissivity (environment localisation).

In the pure loss case, we found the output fidelity for an idler-free protocol. The protocol is assumed to be one-shot in this work, but could be trivially extended to the $M$-round case by taking the $M$-th power of the calculated fidelity. We showed that such a protocol has a lower output

fidelity than the classical (coherent state) protocol over some parameter ranges. This means that it could be a viable alternative to the bipartite entangled protocol for technologically limited scenarios in which we do not have access to a quantum memory. We also investigated the behaviour of the output fidelities of the three different types of protocol.

In the environment localisation case, we calculated the minimum output fidelities for thermal loss, thermal amplifier, and additive noise channels and used them to establish upper and lower bounds on the error probability of discrimination. These bounds hold for the most general adaptive protocols. We then calculated the minimum output fidelity for a classical protocol and so found a region in which we could show a quantum advantage. We also considered a specific protocol involving a photon counting measurement followed by a maximum-likelihood estimation, which allowed us to numerically tighten the upper bound on the error probability. The bounds were then applied to a range of scenarios, as a demonstration. We therefore proved that there exist quantum protocols that are advantageous over all classical protocols for a variety of environment localisation tasks, and detailed a specific quantum protocol that achieves a lower error probability than any classical CPF protocol, for certain channel ensembles.

A possible extension to this work would be formulating bounds on the related task of quantum pattern recognition [105, 106]. In this scenario, there may be multiple target channels or there could be multiple different types of background channel.

# Chapter 4

# Characterising qubit port-based teleportation

The work in this chapter forms the basis of a paper that has been accepted for publication in Journal of Physics A: Mathematical and Theoretical, whose authors are (in order) Jason Pereira, Leonardo Banchi, and Stefano Pirandola [11].

The calculation in Subsection 4.2.5 was used to strengthen a result in the paper "Fundamental limits to quantum channel discrimination", whose authors are (in order) Stefano Pirandola, Riccardo Laurenza, Cosmo Lupo, and Jason Pereira.

We start this chapter by introducing port-based teleportation (PBT) and discussing its usefulness for channel simulation. We then calculate the Choi matrix (and Kraus operators) of the quantum channel simulated by qubit PBT with a given resource state (and using the square-root measurement). We also give simplified expressions for the two port case. We use the formulae to calculate the depolarising probability of a PBT channel using maximally entangled states as a resource. After this, we characterise the PBT process itself, by finding the Kraus operators of the channel mapping a resource state to the output Choi matrix for PBT using that resource state. Next, we apply the formulae to resources that can simulate the amplitude damping (AD) channel and present new classes of resource states that can simulate it better than using multiple copies of the Choi matrix of the simulated channel. Finally, we summarise our findings.

## 4.1 Introduction

Quantum teleportation [83, 107, 108] is a powerful tool in quantum information [15, 54, 55, 109–112]. Teleportation protocols utilise entanglement between quantum states held by a sender and

a receiver to transmit a state. The resulting quantum channel, which maps the sent state to the received state, is determined by the protocol used and by the resource state held by the sender and receiver prior to the protocol being enacted. Such protocols have applications in quantum communications protocols (for example, superdense coding [15]) as well as in quantum computing (quantum gate teleportation [113]), and they can be used as a mathematical tool for the simulation of quantum channels [5, 6] (as mentioned in Chapter 2) and quantum networks [114, 115].

The standard teleportation protocol, as proposed by Bennett et al. [107], uses a shared (between the sender and the receiver) two-qubit state. A measurement is performed on the sender's qubit and the qubit to be teleported, projecting the pair of qubits onto a Bell state. Based on the result of this measurement, one of the four Pauli operators (including the identity) is applied to the receiver's state. The quantum channel resulting from teleportation using this protocol depends on the resource state used. This protocol has limitations, however, as it is only able to simulate Pauli channels [78]. This stems from the fact that the Pauli operators, which are probabilistically applied to the receiver's state, do not commute with every unitary operator. The class of simulable channels was expanded using a generalisation of the standard teleportation protocol, however this protocol is still not capable of simulating all channels [26].

One option for a universal processor that can simulate any channel (for a large enough program state) is the programmable quantum circuit (PQC) from Ref. [28]. Note, however, that this is not a teleportation protocol and does not enact LOCCs. It therefore cannot be used to stretch a key/entanglement distribution protocol (but is still an option for stretching a channel discrimination or parameter estimation protocol).

In Refs. [80, 81], Ishizaka and Hiroshima introduced a new teleportation protocol, called port-based teleportation (PBT). We consider the qubit version of this protocol. In the protocol, the sender and receiver each hold part of a resource state. Each qubit held by the receiver corresponds to a qubit held by the sender, and this shared two-qubit state is referred to as a port. In the standard case introduced by Ishizaka and Hiroshima, each port is an identical Bell pair. Then, a joint measurement is carried out on the sender's states and the qubit to be teleported; the result of this measurement is transmitted to the receiver, and based on this result, the receiver selects one of the ports and traces out the others. This measurement is chosen to be the square-root measurement, which projects the qubit to be teleported and one of the sender's resource qubits onto a Bell pair. Ishizaki and Hiroshima were able to simplify the calculation of the entanglement fidelity of the teleportation channel by representing the qubits held by the sender as a system of spins.

For a finite number of ports $N$, the input-output channel from this protocol is a depolarising

channel. The diamond norm between this channel and the identity channel decreases to zero in the limit of $N \to \infty$. Consequently, in the asymptotic limit, PBT can perfectly simulate any quantum channel, due to the fact that the only post-processing required is the selection of the correct port (which commutes with every channel). In Ref. [2], Pirandola et al. took advantage of this fact to formulate bounds on the error probability of general adaptive discrimination protocols acting on any pair of quantum channels.

In a more general setting, one can replace the original Bell pairs of the PBT protocol with any two-qubit state, and we may even allow entanglement between the ports. Doing so results in the simulation of channels other than the depolarising channel. An explicit characterisation of the qubit channel given by enacting PBT using a given resource state is of interest in quantum information science. If we know the input-output relations for the PBT protocol, we can calculate analytical expressions for the PBT output for any input state and resource state. Such expressions could be used to improve channel bounds based on channel simulation.

## 4.2 Finding the qubit PBT channel for an arbitrary resource

### 4.2.1 Calculating the elements of the channel's Choi matrix

We consider an $N$-port qubit PBT protocol. We call the sender's part of the resource state the $A$ modes and the receiver's part of the resource state the $B$ modes. In order to characterise the channel simulated by PBT using a given resource state, we calculate the Choi matrix for that channel. To do so, we consider a maximally entangled 2-mode state, $\left|\Phi^{\text{Bell}}\right\rangle_{C_0C_1} = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)$. $C_0$ denotes the idler mode and $C_1$ denotes the signal mode. The measurement consists of a POVM described by the operators $\hat{O}_i = \Pi_{i,AC_1} \otimes \mathbb{I}_{BC_0}$, where $i = 1, \ldots, N$. We consider the case in which the $\Pi_i$s describe a square-root measurement. Given a certain measurement result $i$, Bob assumes that the state is teleported to the $i$-th mode $B_i$ and discards all the other ports via a partial trace applied to all $B_j$ with $j \neq i$, all the $A$ modes and $C_1$.

We assume that each port is symmetric under permutation of labels, i.e. that a swap operation that swaps both ports $A_i$ and $A_j$ and ports $B_i$ and $B_j$ does not change the density matrix of the resource state. This does not mean that the ports have to be independent of each other; it is still possible for the $A$ modes (or the $B$ modes or both) to have some entanglement with each other. Consequently, all measurement outcomes are equally likely and all outcomes result in the same channel for the teleported state. We can therefore assume that the state is teleported to the first $B$ port without loss of generality and so only consider one operator. We can justify this assumption as

it is simple to show that, for any non-symmetric resource state $\phi$, there exists a symmetric resource state $\phi^{sym}$ that gives precisely the same channel [116].

Defining $\mathcal{P}_\pi$ as the qubit channel resulting from PBT using the program state $\pi$, we write

$$\mathcal{P}_{\pi_{AB}}(\rho_{C_1}) = \sum_{i=1}^{N} \mathrm{Tr}_{A\bar{B}_i C_1}\left[\left(\sqrt{\Pi_i}_{AC_1} \otimes \mathbb{I}_B\right)(\pi_{AB} \otimes \rho_{C_1})\left(\sqrt{\Pi_i}_{AC_1} \otimes \mathbb{I}_B\right)^\dagger\right], \tag{4.1}$$

where $B_i$ is the port to which the state is teleported, $\bar{B}_i$ denotes all ports except for $B_i$ and $\Pi_i$ is the measurement operator applied to teleport the state to port $i$. Applying the symmetry condition, each value of $i$ gives the same output state, so we can carry out the sum and write

$$\mathcal{P}_{\pi_{AB}}(\rho_{C_1}) = N\mathrm{Tr}_{A\bar{B}_1 C_1}\left[\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_B\right)(\pi_{AB} \otimes \rho_{C_1})\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_B\right)^\dagger\right]. \tag{4.2}$$

The Choi matrix of this channel is then given by

$$\mathbb{I}_{C_0} \otimes \mathcal{P}_{\pi_{AB}}\left(\left|\Phi^{\mathrm{Bell}}\right\rangle\left\langle\Phi^{\mathrm{Bell}}\right|_{C_1 C_0}\right). \tag{4.3}$$

For simplicity, let us initially consider what happens to a teleported arbitrary state $\rho_{C_1}$ (i.e. temporarily ignore the idler mode). Using the fact that the operator enacts the identity on the $B$ modes, we can take the trace on the $\bar{B}$ modes prior to the action of the operator. This allows us the simplification

$$\mathcal{P}_{\pi_{AB}}(\rho_{C_1}) = N\mathrm{Tr}_{AC_1}\left[\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_{B_1}\right)\mathrm{Tr}_{\bar{B}_1}\left[\pi_{AB} \otimes \rho_{C_1}\right]\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_{B_1}\right)^\dagger\right]. \tag{4.4}$$

We denote the matrix representation of $\mathcal{P}_{\pi_{AB}}(\rho_{C_1})$ as $V_{\mathrm{out}}$. We can then write

$$V_{\mathrm{out}} = \begin{pmatrix} V_{\mathrm{out}}^{00} & V_{\mathrm{out}}^{01} \\ V_{\mathrm{out}}^{10} & V_{\mathrm{out}}^{11} \end{pmatrix}, \tag{4.5}$$

$$\begin{aligned} V_{\mathrm{out}}^{ij} &= \langle i|\mathcal{P}_{\pi_{AB}}(\rho_{C_1})|j\rangle \\ &= N\left\langle i\middle|\mathrm{Tr}_{AC_1}\left[\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_{B_1}\right)\mathrm{Tr}_{\bar{B}_1}\left[\pi_{AB} \otimes \rho_{C_1}\right]\left(\sqrt{\Pi_1}_{AC_1} \otimes \mathbb{I}_{B_1}\right)^\dagger\right]\middle|j\right\rangle. \end{aligned} \tag{4.6}$$

Again using the fact that we enact the identity on the $B$ modes, we can take the contraction over the mode $B_1$ within the operation, arriving at

$$V_{\mathrm{out}}^{ij} = N\mathrm{Tr}\left[\sqrt{\Pi_1}_{AC_1}\left\langle i\middle|\mathrm{Tr}_{\bar{B}_1}\left[\pi_{AB} \otimes \rho_{C_1}\right]\middle|j\right\rangle\sqrt{\Pi_1}^\dagger_{AC_1}\right] \tag{4.7}$$

$$= N\mathrm{Tr}\left[\Pi_1\left\langle i\middle|\mathrm{Tr}_{\bar{B}_1}\left[\pi_{AB} \otimes \rho_{C_1}\right]\middle|j\right\rangle\right], \tag{4.8}$$

where we have used the cyclic invariance of the trace and the fact that $\Pi_1$ is a hermitian operator. In the second line and henceforth, we neglect the subscripts on $\Pi_1$. We now define $R^{i+1,j+1} =$

$\langle i|_{B_1}\mathrm{Tr}_{\bar{B}_1}[\pi_{AB}]|j\rangle_{B_1}$ (the $+1$ is so that the labels run from 1 to 2 rather than from 0 to 1). Using this, we can simplify the expression for $V_{\mathrm{out}}$ to

$$V_{\mathrm{out}} = N \begin{pmatrix} \mathrm{Tr}[\Pi_1(R^{11} \otimes \rho_{C_1})] & \mathrm{Tr}[\Pi_1(R^{12} \otimes \rho_{C_1})] \\ \mathrm{Tr}[\Pi_1(R^{21} \otimes \rho_{C_1})] & \mathrm{Tr}[\Pi_1(R^{22} \otimes \rho_{C_1})] \end{pmatrix}. \tag{4.9}$$

Returning to considering the Choi matrix, $C$, we can use this simplification to write

$$C = \frac{N}{2} \begin{pmatrix} \chi_{00}^{11} & \chi_{00}^{12} & \chi_{01}^{11} & \chi_{01}^{12} \\ \chi_{00}^{21} & \chi_{00}^{22} & \chi_{01}^{21} & \chi_{01}^{22} \\ \chi_{10}^{11} & \chi_{10}^{12} & \chi_{11}^{11} & \chi_{11}^{12} \\ \chi_{10}^{21} & \chi_{10}^{22} & \chi_{11}^{21} & \chi_{11}^{22} \end{pmatrix}, \tag{4.10}$$

$$\chi_{mn}^{ij} = \mathrm{Tr}[\Pi_1(R^{ij} \otimes |m\rangle \langle n|_{C_1})]. \tag{4.11}$$

It is worth noting that the Choi matrix is a valid density matrix, so we need only find expressions for the terms on or above the main diagonal. It is also worth noting that $R^{11}$ and $R^{22}$ are (unnormalised) density matrices, whilst $R^{12}$ and $R^{21}$ are not, in general.

### 4.2.2 Simplifying by representing the qubits as a system of spins

Let us now consider the structure of the measurement $\Pi_1$, in a similar way to the analysis in Ref. [81]. $\Pi_1$ is a square-root measurement and can be linearly decomposed as $\Pi_1 = \rho^{-\frac{1}{2}}\sigma_1\rho^{-\frac{1}{2}} + \frac{1}{N}(\mathbb{I} - \rho^{-\frac{1}{2}}\rho\rho^{-\frac{1}{2}})$, where $\sigma_i$ is the projector onto the Bell pair $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ between qubit $C$ and the $i^{th}$ qubit in the sender's resource state (note that it is a different Bell pair from $\left|\Phi^{\mathrm{Bell}}\right\rangle$, the Bell pair we used to define the Choi matrix) and $\rho = \sum_{i=1}^{N}\sigma_i$, as defined in Ref. [81]. Note that the powers of $\rho$ are taken over its support. Let us call the first term in this linear decomposition $M_1$ and call the second term $M_2$; we then have $\Pi_1 = M_1 + M_2$. Ishizaka and Hiroshima found that the eigenvalues of $\rho$ take one of two possible forms: $\lambda_j^- = \frac{1}{2}\left(\frac{N}{2} - j\right)$ or $\lambda_j^+ = \frac{1}{2}\left(\frac{N}{2} + j + 1\right)$ (these expressions differ slightly from those given in Ref. [81], using a pre-factor of $\frac{1}{2}$ rather than $\frac{1}{2^N}$; this is purely due to defining $\sigma_i$ slightly differently). The two types of eigenvalues correspond

to two types of eigenvectors:

$$\begin{aligned}\left|\Psi(\lambda_j^{\mp},m,\alpha)\right\rangle &= \Xi^{\pm-}(j,m+\frac{1}{2})\left|\Phi^{[N]}(j,m+\frac{1}{2},\alpha)\right\rangle_A|0\rangle_C \\ &+ \Xi^{\pm+}(j,m-\frac{1}{2})\left|\Phi^{[N]}(j,m-\frac{1}{2},\alpha)\right\rangle_A|1\rangle_C\,,\end{aligned} \tag{4.12}$$

$$\begin{aligned}\Xi^{++}(j,m) &= \left\langle j,m,\frac{1}{2},\frac{1}{2}\middle|j+\frac{1}{2},m+\frac{1}{2}\right\rangle, \Xi^{+-}(j,m) = \left\langle j,m,\frac{1}{2},-\frac{1}{2}\middle|j+\frac{1}{2},m-\frac{1}{2}\right\rangle, \\ \Xi^{-+}(j,m) &= \left\langle j,m,\frac{1}{2},\frac{1}{2}\middle|j-\frac{1}{2},m+\frac{1}{2}\right\rangle, \Xi^{--}(j,m) = \left\langle j,m,\frac{1}{2},-\frac{1}{2}\middle|j-\frac{1}{2},m-\frac{1}{2}\right\rangle,\end{aligned} \tag{4.13}$$

where $\Xi^{\pm\pm}(j,m)$ represents a Clebsch-Gordan coefficient, with the first superscripted sign determining whether $j$ increases or decreases by $\frac{1}{2}$ and the second superscripted sign determining whether $m$ increases or decreases by $\frac{1}{2}$. Note that $\langle j,m,\frac{1}{2},\pm\frac{1}{2}|J,M\rangle = 0$ if $|M| > J$ or $m \pm \frac{1}{2} \neq M$.

Ishizaka and Hiroshima treat the qubits as spins and hence treat the state $AC$ as a combination of an $N$-spin system and a spin singlet; $\left|\Phi^{[N]}(\lambda_j^{\mp},m,\alpha)\right\rangle$ then gives the orthogonal basis vectors of an $N$-spin system. $j$ corresponds to the magnitude of the spin of the resource state; this is a positive integer or half-integer with minimum value 0 ($\frac{1}{2}$) when $N$ is even (odd). We call the magnitude of the total spin (of the $A$ and $C$ modes) $s$; $s$ has a maximum value of $\frac{N+1}{2}$, which occurs when every spin is aligned (all qubits in $AC$ are 0 or all are 1). $m$ corresponds to the spin of the total system in the z-direction. For fixed $s$, $m$ runs from $-s$ to $s$. The eigenvectors with eigenvalues $\lambda_j^-$ correspond to those states in which the total spin magnitude of the system $AC$ is the sum of the spin magnitudes of the systems $A$ and $C$ (i.e. the $A$ qubits have total spin $j$, the $C$ qubit has total spin $\frac{1}{2}$, so the system $AC$ has total spin $j+\frac{1}{2}$) and the eigenvectors with eigenvalues $\lambda_j^+$ correspond to states in which the spins subtract (i.e. the $A$ qubits have total spin $j$, the $C$ qubit has total spin $\frac{1}{2}$, so the system $AC$ has total spin $j-\frac{1}{2}$). Consequently, for fixed $s$, we have eigenvalues $\lambda_j^-$ with $j$ taking values up to $s-\frac{1}{2}$ and eigenvalues $\lambda_j^+$ with $j$ taking values up to $s+\frac{1}{2}$ (we also cannot have $\lambda_0^+$, since this would require the $A$ qubits to have negative total spin). For some values of $j$, multiple states $\left|\Phi^{[N]}(\lambda_j^{\mp},m)\right\rangle$ exist (i.e. $j$ and $m$ do not uniquely define a basis vector); in this case, we label the different states with $\alpha$, which runs from 1 to the degeneracy of the $j$-value, $\gamma(N,j)$, (which depends only on $N$ and $j$, not on $m$). The degeneracy is given by

$$\gamma(N,j) = \frac{(2j+1)N!}{\left(\frac{N}{2}-j\right)!\left(\frac{N}{2}+j+1\right)!}. \tag{4.14}$$

Ishizaka and Hiroshima then divide the vectors in the $N$-spin basis into two types, based on how they are constructed from the $(N-1)$-spin basis; these are labelled $\left|\Phi_I^{[N]}(j,m,\alpha)\right\rangle$ and

$\left|\Phi_{II}^{[N]}(j,m,\alpha)\right\rangle$. The eigenvectors of $\rho$ constructed using these basis vectors are then labelled $\left|\Psi_I(\lambda_j^{\mp},m,\alpha)\right\rangle$ and $\left|\Psi_{II}(\lambda_j^{\mp},m,\alpha)\right\rangle$. This categorisation is useful, because we can express $\rho$ and $M_1$ in terms of these vectors. The $N$-spin vectors are constructed as

$$\begin{aligned}\left|\Phi_{I}^{[N]}(j,m,\alpha)\right\rangle &= \Xi^{--}(j+\frac{1}{2},m+\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m+\frac{1}{2},\alpha)\right\rangle_{\bar{A}}|0\rangle_{A_1} \\ &\quad+\Xi^{-+}(j+\frac{1}{2},m-\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m-\frac{1}{2},\alpha)\right\rangle_{\bar{A}}|1\rangle_{A_1}\,,\end{aligned} \tag{4.15}$$

$$\begin{aligned}\left|\Phi_{II}^{[N]}(j,m,\alpha)\right\rangle &= \Xi^{+-}(j-\frac{1}{2},m+\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m+\frac{1}{2},\alpha)\right\rangle_{\bar{A}}|0\rangle_{A_1} \\ &\quad+\Xi^{++}(j-\frac{1}{2},m-\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m-\frac{1}{2},\alpha)\right\rangle_{\bar{A}}|1\rangle_{A_1}\,,\end{aligned} \tag{4.16}$$

and the eigenvectors of $\rho$ are constructed as

$$\begin{aligned}\left|\Psi_I(\lambda_j^{\mp},m,\alpha)\right\rangle &= \Xi^{--}(j+\frac{1}{2},m+1)\Xi^{\pm-}(j,m+\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m+1,\alpha)\right\rangle_{\bar{A}}|00\rangle_{A_1C} \\ &\quad+\Xi^{-+}(j+\frac{1}{2},m)\Xi^{\pm-}(j,m+\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m,\alpha)\right\rangle_{\bar{A}}|10\rangle_{A_1C} \\ &\quad+\Xi^{--}(j+\frac{1}{2},m)\Xi^{\pm+}(j,m-\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m,\alpha)\right\rangle_{\bar{A}}|01\rangle_{A_1C} \\ &\quad+\Xi^{-+}(j+\frac{1}{2},m-1)\Xi^{\pm+}(j,m-\frac{1}{2})\left|\Phi^{[N-1]}(j+\frac{1}{2},m-1,\alpha)\right\rangle_{\bar{A}}|11\rangle_{A_1C}\,,\end{aligned} \tag{4.17}$$

$$\begin{aligned}\left|\Psi_{II}(\lambda_j^{\mp},m,\alpha)\right\rangle &= \Xi^{+-}(j-\frac{1}{2},m+1)\Xi^{\pm-}(j,m+\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m+1,\alpha)\right\rangle_{\bar{A}}|00\rangle_{A_1C} \\ &\quad+\Xi^{++}(j-\frac{1}{2},m)\Xi^{\pm-}(j,m+\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m,\alpha)\right\rangle_{\bar{A}}|10\rangle_{A_1C} \\ &\quad+\Xi^{+-}(j-\frac{1}{2},m)\Xi^{\pm+}(j,m-\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m,\alpha)\right\rangle_{\bar{A}}|01\rangle_{A_1C} \\ &\quad+\Xi^{++}(j-\frac{1}{2},m-1)\Xi^{\pm+}(j,m-\frac{1}{2})\left|\Phi^{[N-1]}(j-\frac{1}{2},m-1,\alpha)\right\rangle_{\bar{A}}|11\rangle_{A_1C}\,.\end{aligned} \tag{4.18}$$

These explicit expressions will be useful later.

First, we write $\rho$ as a sum of projectors,

$$\begin{aligned}\rho = \sum_{s=s_{\min}}^{\frac{N+1}{2}} \Bigg[ \lambda^-_{s-\frac{1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Bigg( \left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| + \\ \left|\Psi_{II}(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| \Bigg) + \\ \lambda^+_{s+\frac{1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Bigg( \left|\Psi_I(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right| + \\ \left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right| \Bigg) \Bigg].\end{aligned} \tag{4.19}$$

We then write $\rho^{-\frac{1}{2}}$ in the same way, getting

$$\begin{aligned}\rho^{-\frac{1}{2}} = \sum_{s=s_{\min}}^{\frac{N+1}{2}} \Bigg[ (\lambda^-_{s-\frac{1}{2}})^{-\frac{1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Bigg( \left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| + \\ \left|\Psi_{II}(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| \Bigg) + \\ (\lambda^+_{s+\frac{1}{2}})^{-\frac{1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Bigg( \left|\Psi_I(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right| + \\ \left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right| \Bigg) \Bigg].\end{aligned} \tag{4.20}$$

The above expression is taken only over the support of $\rho$; some of the eigenvectors have an eigenvalue of 0, and we leave these out of the sum. From the form of the eigenvalues, we can see that they are all positive definite except for in the case where $j = \frac{N}{2}$. The eigenvalue $\lambda^-_{\frac{N}{2}} = 0$. The corresponding eigenvectors, $\left|\Psi_{II}(\lambda^-_{\frac{N}{2}}, m, \alpha)\right\rangle$, define the vector space that is not part of the support of $\rho$ and hence the sum of the corresponding projectors gives us $M_2$ (since $\rho^{-\frac{1}{2}}\rho\rho^{-\frac{1}{2}}$ is the identity over the support of $\rho$). Note that there is no $\left|\Psi_I(\lambda^-_{\frac{N}{2}}, m, \alpha)\right\rangle$ vector, since this would require basis vectors of the $(N-1)$-spin subsystem with $j = \frac{N+1}{2}$ to exist. We can write the expression for $M_2$,

$$M_2 = \frac{1}{N} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} \sum_{\alpha} \left|\Psi_{II}(\lambda^-_{\frac{N}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^-_{\frac{N}{2}}, m, \alpha)\right|. \tag{4.21}$$

We now want to find the form of $M_1 = \rho^{-\frac{1}{2}}\sigma_1\rho^{-\frac{1}{2}}$. We express $\sigma_1$ as

$$\begin{aligned}\sigma_1 =& \frac{1}{2}(|01\rangle - |10\rangle)(\langle 01| - \langle 10|)_{A_1 C} \\ & \otimes \sum_{j=j_{\min}}^{\frac{N-1}{2}} \sum_{m=-j}^{j} \sum_{\alpha} \left|\Phi^{[N-1]}(j, m, \alpha)\right\rangle \left\langle\Phi^{[N-1]}(j, m, \alpha)\right|_{\bar{A}}.\end{aligned} \tag{4.22}$$

We then want to find $\frac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{A_1C} \left|\Psi_{I(II)}(\lambda^{\mp}_{s\mp\frac{1}{2}}, m, \alpha)\right\rangle_{AC}$; this will allow us to calculate $\rho^{-\frac{1}{2}}\sigma_1\rho^{-\frac{1}{2}}$. Ishizaka and Hiroshima calculated these using the expressions in Eqs. (4.17) and (4.18) (and the explicit form of the Clebsch-Gordan coefficients), finding

$$\frac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{A_1C} \left|\Psi_I(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle_{AC} = \sqrt{\frac{s}{2s+1}} \left|\Phi^{[N-1]}(s, m, \alpha)\right\rangle_{\bar{A}}, \tag{4.23}$$

$$\frac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{A_1C} \left|\Psi_I(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle_{AC} = 0, \tag{4.24}$$

$$\frac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{A_1C} \left|\Psi_{II}(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle_{AC} = 0, \tag{4.25}$$

$$\frac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{A_1C} \left|\Psi_{II}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle_{AC} = -\sqrt{\frac{s+1}{2s+1}} \left|\Phi^{[N-1]}(s, m, \alpha)\right\rangle_{\bar{A}}. \tag{4.26}$$

Combining our expressions for $\rho^{-\frac{1}{2}}$ and $\sigma_1$ and Eqs. (4.23) to (4.26), we find that $M_1$ takes the form

$$\begin{aligned} M_1 = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Bigg[ & (\lambda^{-}_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1} \left|\Psi_I(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle \Psi_I(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right| \\ & - (\lambda^{-}_{s-\frac{1}{2}}\lambda^{+}_{s+\frac{1}{2}})^{-\frac{1}{2}} \frac{\sqrt{s(s+1)}}{2s+1} \Bigg( \left|\Psi_I(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle \Psi_{II}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right| \\ & + \left|\Psi_{II}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle \Psi_I(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right| \Bigg) \\ & + (\lambda^{+}_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1} \left|\Psi_{II}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle \Psi_{II}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right| \Bigg]. \end{aligned} \tag{4.27}$$

We have summed $s$ from $s_{\min}$ to $\frac{N-1}{2}$, rather than to $\frac{N+1}{2}$, since $\lambda^{-}_{\frac{N}{2}} = 0$ and the vector $\left|\Psi(\lambda^{+}_{\frac{N}{2}+1}, m, \alpha)\right\rangle$ does not exist.

We now calculate $\left\langle 0\middle|\Psi_I(\lambda^{\mp}_{s\mp\frac{1}{2}}, m, \alpha)\right\rangle$, $\left\langle 0\middle|\Psi_{II}(\lambda^{\mp}_{s\mp\frac{1}{2}}, m, \alpha)\right\rangle$, $\left\langle 1\middle|\Psi_I(\lambda^{\mp}_{s\mp\frac{1}{2}}, m, \alpha)\right\rangle$ and $\left\langle 1\middle|\Psi_{II}(\lambda^{\mp}_{s\mp\frac{1}{2}}, m, \alpha)\right\rangle$ (where the contraction is over the $C$ qubit). Using the expressions in Eqs. (4.17) and (4.18) and substituting in the explicit form of the Clebsch-Gordan coefficients [81], we calculate

$$\left\langle 0\middle|\Psi_{I(II)}(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle = \sqrt{\frac{1}{2} - \frac{m}{2s}} \left|\Phi^{[N]}_{I(II)}(s-\frac{1}{2}, m+\frac{1}{2}, \alpha)\right\rangle_{A}, \tag{4.28}$$

$$\left\langle 1\middle|\Psi_{I(II)}(\lambda^{-}_{s-\frac{1}{2}}, m, \alpha)\right\rangle = \sqrt{\frac{1}{2} + \frac{m}{2s}} \left|\Phi^{[N]}_{I(II)}(s-\frac{1}{2}, m-\frac{1}{2}, \alpha)\right\rangle_{A}, \tag{4.29}$$

$$\left\langle 0\middle|\Psi_{I(II)}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle = \sqrt{\frac{1}{2} + \frac{m}{2(s+1)}} \left|\Phi^{[N]}_{I(II)}(s+\frac{1}{2}, m+\frac{1}{2}, \alpha)\right\rangle_{A}, \tag{4.30}$$

$$\left\langle 1\middle|\Psi_{I(II)}(\lambda^{+}_{s+\frac{1}{2}}, m, \alpha)\right\rangle = -\sqrt{\frac{1}{2} - \frac{m}{2(s+1)}} \left|\Phi^{[N]}_{I(II)}(s+\frac{1}{2}, m-\frac{1}{2}, \alpha)\right\rangle_{A}. \tag{4.31}$$

We now have enough to start calculating the components of the Choi matrix. As an example, let us consider the top-left component, $\chi_{00}^{11}$. We are given $R^{11}$, $R^{12}$ and $R^{22}$ as the specification of the resource state. Let us demand that these are given in the $N$-spin basis (the $\left|\Phi_{I(II)}^{[N]}(j,m,\alpha)\right\rangle$ basis). In order to make it clear which components of the resource state we are referring to without choosing some specific matrix representation, we define the function $f_{I,I}^{11}$ such that $f_{I,I}^{11}(j_1,m_1,\alpha_1,j_2,m_2,\alpha_2)$ is the coefficient of $\left|\Phi_I^{[N]}(j_1,m_1,\alpha_1)\right\rangle\left\langle\Phi_I^{[N]}(j_2,m_2,\alpha_2)\right|$ in $R^{11}$. We similarly define $f_{I,II}^{11}$, $f_{II,I}^{11}$ and $f_{II,II}^{11}$, and similar functions for $R^{12}$, $R^{21}$ and $R^{22}$. These functions are simply a way of specifying the resource state. Together, $R^{11}$, $R^{12}$ and $R^{22}$ give the resource state after tracing over all but one $B$ mode. With our assumption that the resource state is unchanged by a swap operation between two ports, this is sufficient to specify the resource state.

We then calculate contributions to the Choi matrix from $M_1$ and $M_2$, using the expressions in Eq. (4.27), Eq. (4.21), and Eqs. (4.28) to (4.31). Recall that $M_1$ acts on the support of $\rho$ and $M_2$ acts on the part of the resource state that is not on the support of $\rho$. The contribution to $\chi_{00}^{11}$ from $M_1$ is

$$\begin{aligned}\mathrm{Tr}[M_1(R^{11}\otimes|0\rangle\langle 0|_{C_1})] = \sum_{s=s_{\min}}^{\frac{N-1}{2}}\sum_{m=-s}^{s}\sum_{\alpha}\Big[&q_-^2 f_{I,I}^{11}(s-\tfrac{1}{2},m+\tfrac{1}{2},\alpha,s-\tfrac{1}{2},m+\tfrac{1}{2},\alpha)\\ &-q_-r_+\Big(f_{I,II}^{11}(s-\tfrac{1}{2},m+\tfrac{1}{2},\alpha,s+\tfrac{1}{2},m+\tfrac{1}{2},\alpha)\\ &+f_{II,I}^{11}(s+\tfrac{1}{2},m+\tfrac{1}{2},\alpha,s-\tfrac{1}{2},m+\tfrac{1}{2},\alpha)\Big)\\ &+r_+^2 f_{II,II}^{11}(s+\tfrac{1}{2},m+\tfrac{1}{2},\alpha,s+\tfrac{1}{2},m+\tfrac{1}{2},\alpha)\Big],\end{aligned} \tag{4.32}$$

$$q_\pm = \sqrt{\frac{2(s\pm m)}{(N+1-2s)(2s+1)}}, \tag{4.33}$$

$$r_\pm = \sqrt{\frac{2(s\pm m+1)}{(N+3+2s)(2s+1)}}, \tag{4.34}$$

where we have used the explicit form of the eigenvalues. The contribution to $\chi_{00}^{11}$ from $M_2$ is

$$\mathrm{Tr}[M_2(R^{11}\otimes|0\rangle\langle 0|_{C_1})] = \frac{1}{N}\sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}}(\frac{1}{2}-\frac{m}{N+1})f_{II,II}^{11}(\frac{N}{2},m+\frac{1}{2},1,\frac{N}{2},m+\frac{1}{2},1). \tag{4.35}$$

We do not need to sum over $\alpha$, since there is no degeneracy in the states we sum over. By adding these two contributions and multiplying by $\frac{N}{2}$ (as per Eq. (4.10)), we get the top-left component

of the Choi matrix. We call this component $C^{11}$. Then,

$$\begin{aligned} C^{11} =& \frac{N}{2} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} \sum_{\alpha} \Big[ q_-^2 f_{I,I}^{11}(s-\frac{1}{2}, m+\frac{1}{2}, \alpha, s-\frac{1}{2}, m+\frac{1}{2}, \alpha) \\ & - q_- r_+ \Big( f_{I,II}^{11}(s-\frac{1}{2}, m+\frac{1}{2}, \alpha, s+\frac{1}{2}, m+\frac{1}{2}, \alpha) \\ & + f_{II,I}^{11}(s+\frac{1}{2}, m+\frac{1}{2}, \alpha, s-\frac{1}{2}, m+\frac{1}{2}, \alpha) \Big) \\ & + r_+^2 f_{II,II}^{11}(s+\frac{1}{2}, m+\frac{1}{2}, \alpha, s+\frac{1}{2}, m+\frac{1}{2}, \alpha) \Big] \\ & + \frac{1}{2} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} (\frac{1}{2} - \frac{m}{N+1}) f_{II,II}^{11}(\frac{N}{2}, m+\frac{1}{2}, 1, \frac{N}{2}, m+\frac{1}{2}, 1). \end{aligned} \tag{4.36}$$

We can express this more succinctly by defining the functions

$$g_b^a[-+-+](s,m) = \sum_{\alpha} f_b^a(s-\frac{1}{2}, m+\frac{1}{2}, \alpha, s-\frac{1}{2}, m+\frac{1}{2}, \alpha), \tag{4.37}$$

where the index $a$ could be "11", "12", "21", or "22" and the index $b$ could be "$I, I$", "$I, II$", "$II, I$", or "$II, II$". Equally, the signs given as arguments to the $g$ function can be changed (e.g. we could have "++++" instead of "-+-+"), and in this case the signs in the $f$ function change accordingly. We can then express $C^{11}$ as

$$\begin{aligned} C^{11} =& \frac{N}{2} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} \big[ q_-^2 g_{I,I}^{11}[-+-+](s,m) + r_+^2 g_{II,II}^{11}[++++](s,m) \\ & - q_- r_+ \big( g_{I,II}^{11}[-+++](s,m) + g_{II,I}^{11}[++-+](s,m) \big) \big] \\ & + \frac{1}{2} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} (\frac{1}{2} - \frac{m}{N+1}) g_{II,II}^{11}[-+-+](\frac{N+1}{2}, m). \end{aligned} \tag{4.38}$$

To get the expressions for $C^{12}$ and $C^{22}$, we simply replace $g^{11}$ with $g^{12}$ and $g^{22}$ respectively in the expression for $C^{11}$. Equally, once we have the expression for $C^{13}$, we can get the expressions for $C^{14}$, $C^{23}$ and $C^{24}$ by replacing $g^{11}$ with $g^{12}$, $g^{21}$ and $g^{22}$ respectively in the expression for $C^{13}$. Similarly, starting from the expressions for $C^{33}$, we get the expressions for $C^{34}$ and $C^{44}$ by replacing $g^{11}$ with $g^{12}$ and $g^{22}$ respectively in the expression for $C^{33}$. Essentially, if we divide the Choi matrix into quarters, we only need one expression per block of four elements, and the other expressions only require trivial modifications. We also only need the expressions for the upper triangle of the Choi matrix, since the Choi matrix is a valid density matrix and so is hermitian. We

give the expressions for $C^{13}$ and $C^{33}$ below:

$$\begin{aligned}
C^{13} =& \frac{N}{2} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} \big[q_- q_+ g^{11}_{I,I}[-+--](s,m) - r_- r_+ g^{11}_{II,II}[+++-](s,m) \\
& + q_- r_- g^{11}_{I,II}[-++-](s,m) - q_+ r_+ g^{11}_{II,I}[++--](s,m)\big] \\
& + \frac{1}{2} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} \sqrt{\frac{1}{4} - \left(\frac{m}{N+1}\right)^2}\, g^{11}_{II,II}[-+--](\frac{N+1}{2}, m),
\end{aligned} \tag{4.39}$$

$$\begin{aligned}
C^{33} =& \frac{N}{2} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} \big[q_+^2 g^{11}_{I,I}[----](s,m) + r_-^2 g^{11}_{II,II}[+-+-](s,m) \\
& + q_+ r_- \left(g^{11}_{I,II}[--+-](s,m) + g^{11}_{II,I}[+---](s,m)\right)\big] \\
& + \frac{1}{2} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} (\frac{1}{2} + \frac{m}{N+1}) g^{11}_{II,II}[----](\frac{N+1}{2}, m).
\end{aligned} \tag{4.40}$$

These are, in fact, fairly simple expressions, although quite long when written in this form. If we impose constraints on the resource state, we can simplify the expressions.

We now have an analytical expression for the Choi matrix for any PBT qubit operation. The only assumption made is that all ports are identical. Any channel simulable via PBT can be simulated using a resource state of this type [116].

To show how the Choi matrix, $C$, is constructed from the components given, we write the following, where * denotes the complex conjugate and where $C^{ij}(g^{11} \to g^{kl})$ means the expression for $C^{ij}$ with all instances of $g^{11}$ replaced with $g^{kl}$:

$$C = \begin{pmatrix}
C^{11}(g^{11}) & C^{11}(g^{11} \to g^{12}) & C^{13}(g^{11}) & C^{13}(g^{11} \to g^{12}) \\
C^{11}(g^{11} \to g^{12})^* & C^{11}(g^{11} \to g^{22}) & C^{13}(g^{11} \to g^{21}) & C^{13}(g^{11} \to g^{22}) \\
C^{13}(g^{11})^* & C^{13}(g^{11} \to g^{21})^* & C^{33}(g^{11}) & C^{33}(g^{11} \to g^{12}) \\
C^{13}(g^{11} \to g^{12})^* & C^{13}(g^{11} \to g^{22})^* & C^{33}(g^{11} \to g^{12})^* & C^{33}(g^{11} \to g^{22})
\end{pmatrix}. \tag{4.41}$$

We may also wish to find the Kraus operators [99] of the qubit channel resulting from PBT using a given resource state. This is an alternative but equivalent channel representation to the Choi matrix. We may also wish to characterise the channel mapping from a given resource state to the output Choi matrix of the qubit channel. This channel takes a resource state as input and outputs the Choi matrix of the qubit channel resulting from PBT using that resource state. These Kraus operators are rectangular (the number of qubits in the output is less than the number in the input). They characterise the processor (i.e. the operation of carrying out a square-root measurement on the modes $AC_1$, followed by the selection of a $B$ port based on the measurement outcome).

### 4.2.3 Converting from the Choi matrix to the Kraus operators of the qubit channel

The Choi matrix holds all information about the state, but we would like to also be able to express the channel as a set of Kraus operators [99]. We can do this using the following algorithm, starting from the Choi matrix $V$.

1. Find the eigendecomposition of $V$ and write:

$$V = \sum_{i=1}^{4} \lambda_i \left|v_i'\right\rangle \left\langle v_i'\right| . \tag{4.42}$$

2. We then define $|v_i\rangle = \sqrt{\lambda_i}\,|v_i'\rangle$, so that we can write:

$$V = \sum_{i=1}^{4} |v_i\rangle \langle v_i| . \tag{4.43}$$

3. The (up to) four Kraus operators, labelled as $K_i$, are then written (in the canonical basis) as

$$K_i = \begin{pmatrix} \langle 00|v_i\rangle & \langle 10|v_i\rangle \\ \langle 01|v_i\rangle & \langle 11|v_i\rangle \end{pmatrix} \tag{4.44}$$

We can verify that, if the Kraus operators constructed in this way are applied to a Bell state, we recover the initial Choi matrix. Numerically, this algorithm is simple to implement, since we are only finding the eigendecomposition of a 4 by 4 matrix.

An intuition about why this algorithm works can be gained by calculating the output state, $\rho^{\text{out}}$, for an arbitrary input state, $\rho^{\text{in}}$, with no idler modes, using both the Kraus operators, $K_i$, and the Choi matrix, $V$. Using the Kraus operator formalism, we can write

$$\rho^{\text{out}} = \sum_{i=1}^{4} K_i \rho^{\text{in}} K_i^{\dagger}. \tag{4.45}$$

Using the link product formalism, which gives the output state of a channel directly from the Choi matrix and the input state, we can write [23]

$$\rho^{\text{out}} = \rho_A^{\text{in}} * V_{AB} = \text{Tr}_A[(\rho_A^{\text{in},T} \otimes \mathbb{I}_B) V_{AB}], \tag{4.46}$$

where $*$ denotes the link product, $T$ denotes the transpose, and the subscripts $A$ and $B$ denote the systems on which the operators are defined. Applying the decomposition in Eq. (4.43), we can write

$$\rho^{\text{out}} = \sum_{i=1}^{4} \text{Tr}_A[(\rho_A^{\text{in},T} \otimes \mathbb{I}_B) \left|v_i\right\rangle \left\langle v_i\right|_{AB}]. \tag{4.47}$$

By direct calculation, using the expression for the Kraus operators given in Eq. (4.44), we get

$$K_i\rho^{\text{in}}K_i^\dagger = \langle 0|_A \rho_A^{\text{in},T} |0\rangle_A \langle 0|_A |v_i\rangle_{AB} \langle v_i|_{AB} |0\rangle_A + \langle 0|_A \rho_A^{\text{in},T} |1\rangle_A \langle 1|_A |v_i\rangle_{AB} \langle v_i|_{AB} |0\rangle_A + \langle 1|_A \rho_A^{\text{in},T} |0\rangle_A \langle 0|_A |v_i\rangle_{AB} \langle v_i|_{AB} |1\rangle_A + \langle 1|_A \rho_A^{\text{in},T} |1\rangle_A \langle 1|_A |v_i\rangle_{AB} \langle v_i|_{AB} |1\rangle_A . \tag{4.48}$$

Finally, we can simplify, writing

$$K_i\rho^{\text{in}}K_i^\dagger = \text{Tr}_A[(\rho_A^{\text{in},T} \otimes \mathbb{I}_B) |v_i\rangle \langle v_i|_{AB}], \tag{4.49}$$

and thus showing the equivalence of Eqs. (4.45) and (4.47). Hence, we demonstrate that the Kraus operators defined by Eq. (4.44) describe the same channel as the Choi matrix $V$.

### 4.2.4 Two port PBT

As an example, suppose we only have two ports. Let us calculate the Choi matrix for this case. We again assume that the two ports are identical under exchange of labels. The reduced resource states $R^{11}$, $R^{12}$ and $R^{22}$ are then 4 by 4 matrices. We will write them in the basis: $\{\frac{1}{\sqrt{2}}(|10\rangle - |01\rangle), |00\rangle, \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle), |11\rangle\}$. These are the vectors $\{\left|\Phi_I^{[2]}(0,0)\right\rangle, \left|\Phi_{II}^{[2]}(1,-1)\right\rangle, \left|\Phi_{II}^{[2]}(1,0)\right\rangle, \left|\Phi_{II}^{[2]}(1,1)\right\rangle\}$. Note that there are no degenerate $(j,m)$ combinations for two ports, so we do not need to specify the degeneracy, $\alpha$. We can therefore immediately remove the sum over $\alpha$. Since $s = \frac{1}{2}$ is the only value of $s$ for which either $\left|\Phi_I^{[2]}(s-\frac{1}{2},m)\right\rangle$ or $\left|\Phi_{II}^{[2]}(s+\frac{1}{2},m)\right\rangle$ exist, we do not need to sum over $s$ either and simply set $s = \frac{1}{2}$. $R^{ij}$ takes the form

$$R^{ij} = \begin{pmatrix} f^{ij}_{I,I}(0,0,0,0) & f^{ij}_{I,II}(0,0,1,-1) & f^{ij}_{I,II}(0,0,1,0) & f^{ij}_{I,II}(0,0,1,1) \\ f^{ij}_{II,I}(1,-1,0,0) & f^{ij}_{II,II}(1,-1,1,-1) & f^{ij}_{II,II}(1,-1,1,0) & f^{ij}_{II,II}(1,-1,1,1) \\ f^{ij}_{II,I}(1,0,0,0) & f^{ij}_{II,II}(1,0,1,-1) & f^{ij}_{II,II}(1,0,1,0) & f^{ij}_{II,II}(1,0,1,1) \\ f^{ij}_{II,I}(1,1,0,0) & f^{ij}_{II,II}(1,1,1,-1) & f^{ij}_{II,II}(1,1,1,0) & f^{ij}_{II,II}(1,1,1,1) \end{pmatrix}, \tag{4.50}$$

where we have excluded $\alpha$ from the arguments of $f$. We again note that $R^{11}$, $R^{12}$, $R^{21}$ and $R^{22}$ are derived from the density matrix of the full resource state by taking the trace over all $B$ modes except for the first $B$ mode. In the two mode case, they can be written as

$$R^{11} = \langle 0|_{B1} \text{Tr}_{B2}(R) |0\rangle_{B1}, \tag{4.51}$$

$$R^{12} = \langle 0|_{B1} \text{Tr}_{B2}(R) |1\rangle_{B1}, \tag{4.52}$$

$$R^{21} = \langle 1|_{B1} \text{Tr}_{B2}(R) |0\rangle_{B1}, \tag{4.53}$$

$$R^{22} = \langle 1|_{B1} \text{Tr}_{B2}(R) |1\rangle_{B1}. \tag{4.54}$$

The expression for $C^{11}$ now reduces to

$$C^{11} = \frac{1}{2}\mathrm{Tr}\left[R^{11}\right] - \frac{1}{2\sqrt{3}}\left(f^{11}_{I,II}(0,0,1,0) + f^{11}_{II,I}(1,0,0,0)\right), \tag{4.55}$$

where we have used

$$\begin{aligned}\mathrm{Tr}\left[R^{11}\right] &= f^{11}_{I,I}(0,0,0,0) + f^{11}_{II,II}(1,-1,1,-1) + f^{11}_{II,II}(1,0,1,0) + f^{11}_{II,II}(1,1,1,1) \\ &= \mathrm{Tr}\left[\langle 0|_{B1}\, R\, |0\rangle_{B1}\right].\end{aligned} \tag{4.56}$$

The expressions for $C^{13}$ and $C^{33}$ reduce to

$$C^{13} = \frac{1}{\sqrt{6}}\left(f^{11}_{I,II}(0,0,1,-1) - f^{11}_{II,I}(1,1,0,0)\right), \tag{4.57}$$

$$C^{33} = \frac{1}{2}\mathrm{Tr}\left[R^{11}\right] + \frac{1}{2\sqrt{3}}\left(f^{11}_{I,II}(0,0,1,0) + f^{11}_{II,I}(1,0,0,0)\right). \tag{4.58}$$

### 4.2.5 Calculating the depolarisation probability for qubit PBT with a maximally entangled resource

PBT with a maximally entangled resource state enacts a depolarising channel [81]. Our analytical formulae for the components of the output Choi matrix give an easy way to calculate the depolarising probability of the channel simulated by $N$-port PBT.

The Choi matrix of a depolarising channel is

$$C_{\mathrm{dep}} = \begin{pmatrix} \frac{1}{2}-\frac{\xi}{4} & 0 & 0 & \frac{1}{2}-\frac{\xi}{2} \\ 0 & \frac{\xi}{4} & 0 & 0 \\ 0 & 0 & \frac{\xi}{4} & 0 \\ \frac{1}{2}-\frac{\xi}{2} & 0 & 0 & \frac{1}{2}-\frac{\xi}{4} \end{pmatrix}, \tag{4.59}$$

where $\xi$ is the depolarising probability of the channel. Since the channel has only one parameter, we only need to find one (non-zero) element of the Choi matrix in order to characterise it. We pick $C^{33}_{\mathrm{dep}}$ (the third element on the main diagonal); the expression for this component is given by Eq. (4.40).

We start by finding $R^{11}$ for the maximally entangled resource, $\left|\Phi^{\mathrm{Bell}}\right\rangle\left\langle\Phi^{\mathrm{Bell}}\right|^{\otimes N}$. We find

$$\begin{aligned} R^{11} &= \langle 0|_{B_1}\mathrm{Tr}_{\bar{B}_1}[\left|\Phi^{\mathrm{Bell}}\right\rangle\left\langle\Phi^{\mathrm{Bell}}\right|^{\otimes N}_{AB}]|0\rangle_{B_1} \\ &= \frac{1}{2^{N-1}}\langle 0|_{B_1}\left|\Phi^{\mathrm{Bell}}\right\rangle\left\langle\Phi^{\mathrm{Bell}}\right|_{A_1B_1}|0\rangle_{B_1}\otimes\mathbb{I}_{\bar{A}_1} \\ &= \frac{1}{2^N}\left|1\right\rangle\left\langle 1\right|_{A_1}\otimes\left(\sum_{j,m,\alpha}\left|\Phi^{[N-1]}(j,m,\alpha)\right\rangle\left\langle\Phi^{[N-1]}(j,m,\alpha)\right|_{\bar{A}_1}\right), \end{aligned} \tag{4.60}$$

where the sum is over all valid values of $j$, $m$, and $\alpha$. We can express $R^{11}$ in the $N$-spin basis using Eqs. (4.15) and (4.16). This allows us to write the functions

$$f^{11}_{I,I}(s-\frac{1}{2},m-\frac{1}{2},s-\frac{1}{2},m-\frac{1}{2}) = \frac{1}{2^N}\left[\Xi^{-+}(s,m-1)\right]^2 = \frac{s-m+1}{2^N(2s+1)} \tag{4.61}$$

$$\begin{aligned} f^{11}_{I,II}(s-\frac{1}{2},m-\frac{1}{2},s+\frac{1}{2},m-\frac{1}{2}) &= \frac{1}{2^N}\left[\Xi^{-+}(s,m-1)\Xi^{++}(s,m-1)\right] \\ &= -\frac{\sqrt{(s-m+1)(s+m)}}{2^N(2s+1)} \end{aligned} \tag{4.62}$$

$$f^{11}_{II,II}(s+\frac{1}{2},m-\frac{1}{2},s+\frac{1}{2},m-\frac{1}{2}) = \frac{1}{2^N}\left[\Xi^{++}(s,m-1)\right]^2 = \frac{s+m}{2^N(2s+1)}, \tag{4.63}$$

noting also that $f^{11}_{I,II} = f^{11}_{II,I}$, since $R^{11}$ is a conditional density matrix (and therefore must be hermitian).

We can express the degeneracy for the $N-1$-spin basis as

$$\gamma(N-1,s) = \frac{(2s+1)(N-1)!}{\left(\frac{N-1}{2}-s\right)!\left(\frac{N+1}{2}+s\right)!} = \frac{2s+1}{N}\binom{N}{\frac{N-1}{2}-s}, \tag{4.64}$$

where the expression on the right hand side uses a binomial coefficient. We can therefore write

$$\begin{aligned} q_+^2 g_{I,I} + r_-^2 g_{II,II} + q_+ r_- \left(g_{I,II} + g_{II,I}\right) = &\frac{(s+m)(s-m+1)}{2^{N+1}N(2s+1)}\binom{N}{\frac{N-1}{2}-s} \\ &\times \left[(\lambda^-_{s-\frac{1}{2}})^{-\frac{1}{2}} - (\lambda^+_{s+\frac{1}{2}})^{-\frac{1}{2}}\right]^2, \end{aligned} \tag{4.65}$$

where it is implicit that the indices for the $g$-functions are those found in the first sum in Eq. (4.40). We then carry out the sum

$$\sum_{m=-s}^{s}(s+m)(s-m+1) = \frac{2}{3}s(s+1)(2s+1). \tag{4.66}$$

We expand the last term in Eq. (4.65), getting

$$\left[(\lambda^-_{s-\frac{1}{2}})^{-\frac{1}{2}} - (\lambda^+_{s+\frac{1}{2}})^{-\frac{1}{2}}\right]^2 = 8\frac{(N+2)-\sqrt{(N+2)^2-(2s+1)^2}}{(N+2)^2-(2s+1)^2}. \tag{4.67}$$

We then calculate

$$\left(\frac{1}{2}+\frac{m}{N+1}\right)g_{II,II} = \frac{1}{2^N N(N+1)}\left(\frac{N-1}{2}+m\right)\left(\frac{N+1}{2}+m\right), \tag{4.68}$$

where it is implicit that the indices for the $g$-function are those found in the second sum in Eq. (4.40). We perform the sum

$$\sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}}\left(\frac{N-1}{2}+m\right)\left(\frac{N+1}{2}+m\right) = \frac{1}{3}N(N+1)(N+2). \tag{4.69}$$

Substituting these expressions into Eq. (4.40), we get

$$C_{\text{dep}}^{33} = \frac{1}{3\times 2^{N-2}} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \left[ s(s+1) \binom{N}{\frac{N-1}{2}-s} \frac{(N+2)-\sqrt{(N+2)^2-(2s+1)^2}}{(N+2)^2-(2s+1)^2} \right] + \frac{N+2}{3\times 2^{N+1}}, \tag{4.70}$$

which immediately gives

$$\xi_N = \frac{1}{3\times 2^{N-4}} \sum_{s=s_{\min}}^{\frac{N-1}{2}} \left[ s(s+1) \binom{N}{\frac{N-1}{2}-s} \frac{(N+2)-\sqrt{(N+2)^2-(2s+1)^2}}{(N+2)^2-(2s+1)^2} \right] + \frac{N+2}{3\times 2^{N-1}}, \tag{4.71}$$

where $\xi_N$ is the depolarising probability of the $N$-port qubit PBT channel with a maximally entangled resource. We numerically observe that $\xi_N$ scales approximately with $\frac{1}{N}$ for large $N$. This probability is calculated in a similar way in Ref. [2], but without using the explicit formulae presented here.

## 4.3 Characterising the qubit PBT protocol

We want to characterise the channel mapping from the (input) program state (with $2N$ qubits) to the (output) Choi matrix of the PBT channel (with 2 qubits). This is a characterisation of the PBT protocol itself (with the square-root measurement and a permutation-symmetric resource state). An implicit expression for this map is derived in Ref. [28], however here we derive explicit expressions for the Kraus operators.

Defining $\Lambda$ as the channel from the program state to the Choi matrix of the qubit channel, we can write

$$\Lambda(\pi) = \sum_{i=1}^{N} \text{Tr}_{A\bar{B}_i C_1} \left[ \left( \sqrt{\Pi_i}_{AC_1} \otimes \mathbb{I}_{BC_0} \right) \left( \pi_{AB} \otimes \left| \Phi_{C_0C_1}^{\text{Bell}} \right\rangle \left\langle \Phi_{C_0C_1}^{\text{Bell}} \right| \right) \left( \sqrt{\Pi_i}_{AC_1} \otimes \mathbb{I}_{BC_0} \right)^\dagger \right] \tag{4.72}$$

$$= \sum_{ik} K_{ik} \pi K_{ik}^\dagger, \tag{4.73}$$

where $B_i$ is the port to which the state is teleported, $\Pi_i$ is the measurement operator applied to teleport the state to port i and

$$K_{ik} = \left\langle e_k^{(i)} \right| \sqrt{\Pi_i}_{AC_1} \otimes \mathbb{I}_{BC_0} \left| \Phi_{C_0C_1}^{\text{Bell}} \right\rangle. \tag{4.74}$$

The $\left| e_k^{(i)} \right\rangle$ are basis vectors on the systems $A\bar{B}_i C_1$ (the traced over systems).

First, let us apply the assumption of symmetry under exchange of labels. We can therefore replace $K_{ik}$ with $K_k = \sqrt{N}K_{1k}$. We can now calculate $\sqrt{\Pi_1}$, using the expressions in Eqs. (4.27) and (4.21). From the fact that $M_1$ and $M_2$ have orthogonal supports, we can take the square roots of each separately. In fact, due to $M_1$ having no mixing between basis vectors with different $s$, $m$, or $\alpha$ values, we can treat each set of values $\{s, m, \alpha\}$ separately and hence can write

$$\sqrt{\Pi_1} = \sum_{sm\alpha} \sqrt{M_1^{sm\alpha}} + \sqrt{M_2}, \tag{4.75}$$

where $M_1^{sm\alpha}$ is the contribution to $M_1$ from the two eigenvectors $\left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle$ and $\left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle$. Since $M_2$, as expressed in Eq. (4.21), is already diagonal, it is trivial to write

$$\sqrt{M_2} = \frac{1}{\sqrt{N}} \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} \left|\Psi_{II}(\lambda^-_{\frac{N}{2}}, m)\right\rangle \left\langle\Psi_{II}(\lambda^-_{\frac{N}{2}}, m)\right|, \tag{4.76}$$

where we have removed the sum over $\alpha$, due to there being no degeneracy in the component eigenvectors.

We now want to find $\sqrt{M_1^{sm\alpha}}$, starting from

$$\begin{aligned} M_1^{sm\alpha} = {} & (\lambda^-_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1} \left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| \\ & - (\lambda^-_{s-\frac{1}{2}}\lambda^+_{s+\frac{1}{2}})^{-\frac{1}{2}} \frac{\sqrt{s(s+1)}}{2s+1} \Bigg( \left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right| \\ & + \left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right| \Bigg) \\ & + (\lambda^+_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1} \left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle \left\langle\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right|. \end{aligned} \tag{4.77}$$

From the form of Eq. (4.77), we can see that $M_1^{sm\alpha}$ can be written as

$$M_1^{sm\alpha} = |\text{vec}^{sm\alpha}\rangle \langle\text{vec}^{sm\alpha}|, \tag{4.78}$$

$$|\text{vec}^{sm\alpha}\rangle = \sqrt{(\lambda^-_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1}} \left|\Psi_I(\lambda^-_{s-\frac{1}{2}}, m, \alpha)\right\rangle - \sqrt{(\lambda^+_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1}} \left|\Psi_{II}(\lambda^+_{s+\frac{1}{2}}, m, \alpha)\right\rangle, \tag{4.79}$$

where it must be noted that $|\text{vec}^{sm\alpha}\rangle$ is unnormalised. This means that $M_1^{sm\alpha}$ has only one non-zero eigenvalue, given by

$$\begin{aligned} \text{eig}^{sm\alpha} &= (\lambda^-_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1} + (\lambda^+_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1} \\ &= \frac{4(N+1)}{(N+1-2s)(N+3+2s)}. \end{aligned} \tag{4.80}$$

Consequently, we can write

$$\sqrt{M_1^{sm\alpha}} = (\text{eig}^{sm\alpha})^{-\frac{1}{2}} \left|\text{vec}^{sm\alpha}\right\rangle \left\langle\text{vec}^{sm\alpha}\right| . \tag{4.81}$$

Combining our expressions for $M_1$ and $M_2$, we have

$$\begin{aligned}\sqrt{\Pi_1} = \frac{1}{\sqrt{N}} \sum_m \left|\Psi_{II}(\lambda^-_{\frac{N}{2}}, m)\right\rangle \left\langle\Psi_{II}(\lambda^-_{\frac{N}{2}}, m)\right| \\ + \sum_{sm\alpha} \sqrt{\frac{(N+1-2s)(N+3+2s)}{4(N+1)}} \left|\text{vec}^{sm\alpha}\right\rangle \left\langle\text{vec}^{sm\alpha}\right| .\end{aligned} \tag{4.82}$$

We now express the basis vectors $\left|e_k^{(1)}\right\rangle$ as

$$|e_k\rangle = |e_{k_1}\rangle_{AC_1} |e_{k_2}\rangle_{\bar{B}} , \tag{4.83}$$

where $\bar{B}$ refers to the $B$ modes except for $B_1$. $|e_{k_1}\rangle_{AC_1}$ are the $|\text{vec}_2^{sm\alpha}\rangle$ basis vectors (on the system $AC_1$) and the $|e_{k_2}\rangle_{\bar{B}}$ are any choice of orthonormal basis vectors on the system $\bar{B}$. There are two types of Kraus operator, depending on whether $|e_{k_1}\rangle_{AC_1}$ lies in the support of $M_1$ or of $M_2$. We will label these Kraus operators $K_k^1$ and $K_k^2$ respectively. Using Eqs. (4.28) to (4.31), we find that the Kraus operators $K_k^2$ take the form

$$\begin{aligned}K_k^2 = \frac{1}{\sqrt{2}} \Bigg( \sqrt{\frac{1}{2} - \frac{m}{N+1}} \left|0\right\rangle_{C_0} \left\langle \Phi_{II}(\frac{N}{2}, m+\frac{1}{2}) \right|_{AC_1} \\ + \sqrt{\frac{1}{2} + \frac{m}{N+1}} \left|1\right\rangle_{C_0} \left\langle \Phi_{II}(\frac{N}{2}, m-\frac{1}{2}) \right|_{AC_1} \Bigg) \langle e_{k_2}|_{\bar{B}} \otimes \mathbb{I}_{B_1},\end{aligned} \tag{4.84}$$

where the label $k$ determines the $m$ value and the choice of basis vector $|e_{k_2}\rangle_{\bar{B}}$. We find that the Kraus operators $K_k^1$ take the form

$$\begin{aligned}K_k^1 = \sqrt{\frac{N}{2}} \Bigg[ \left|0\right\rangle_{C_0} \Bigg( \sqrt{(\lambda^-_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1} \left(\frac{1}{2} - \frac{m}{2s}\right)} \left\langle \Phi_I(s-\frac{1}{2}, m+\frac{1}{2}, \alpha) \right| \\ - \sqrt{(\lambda^+_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1} \left(\frac{1}{2} + \frac{m}{2(s+1)}\right)} \left\langle \Phi_{II}(s+\frac{1}{2}, m+\frac{1}{2}, \alpha) \right| \Bigg)_{AC_1} \\ + \left|1\right\rangle_{C_0} \Bigg( \sqrt{(\lambda^-_{s-\frac{1}{2}})^{-1} \frac{s}{2s+1} \left(\frac{1}{2} + \frac{m}{2s}\right)} \left\langle \Phi_I(s-\frac{1}{2}, m-\frac{1}{2}, \alpha) \right| \\ + \sqrt{(\lambda^+_{s+\frac{1}{2}})^{-1} \frac{s+1}{2s+1} \left(\frac{1}{2} - \frac{m}{2(s+1)}\right)} \left\langle \Phi_{II}(s+\frac{1}{2}, m-\frac{1}{2}, \alpha) \right| \Bigg)_{AC_1} \Bigg] \langle e_{k_2}|_{\bar{B}} \otimes \mathbb{I}_{B_1},\end{aligned} \tag{4.85}$$

where the label $k$ determines the values of $s$, $m$ and $\alpha$, and the choice of basis vector $|e_{k_2}\rangle_{\bar{B}}$. We can simplify this expression, and so can write

$$K_k^1 = \sqrt{\frac{N}{2}} \Bigg[ |0\rangle_{C_0} \left( q_- \left\langle \Phi_I(s-\frac{1}{2}, m+\frac{1}{2}, \alpha) \right| - r_+ \left\langle \Phi_{II}(s+\frac{1}{2}, m+\frac{1}{2}, \alpha) \right| \right)_{AC_1} + |1\rangle_{C_0} \left( q_+ \left\langle \Phi_I(s-\frac{1}{2}, m-\frac{1}{2}, \alpha) \right| + r_- \left\langle \Phi_{II}(s+\frac{1}{2}, m-\frac{1}{2}, \alpha) \right| \right)_{AC_1} \Bigg] \langle e_{k_2}|_{\bar{B}} \otimes \mathbb{I}_{B_1}, \tag{4.86}$$

where $q_\pm$ and $r_\pm$ are defined as per Eqs. (4.33) and (4.34).

Note that the basis vectors $|e_{k_2}\rangle_{\bar{B}}$ simply trace over the $\bar{B}$ system, i.e. for each Kraus operator, there are $2^N - 1$ other Kraus operators that are identical up to a change in $k_2$. Hence, we can trace over the $\bar{B}$ modes of the resource state; in this case the Kraus operators of the channel from $\mathrm{Tr}_{\bar{B}}[\pi_{AB}]$ to the output Choi matrix are $K_k^1$ and $K_k^2$ without the vectors $|e_{k_2}\rangle_{\bar{B}}$ (i.e. the labels $k$ determine only the values of $s$, $m$ and $\alpha$).

## 4.4 Simulating the amplitude damping channel

We know that in the limit of $N \to \infty$, a resource state comprised of $N$ copies of the Choi matrix of a given channel perfectly simulates that channel. This is because PBT over such a resource state is equivalent to passing the transmitted state through an identity channel followed by the desired channel. However, for finite $N$, it may be the case that there is a resource state that simulates a given channel better than $N$ copies of the Choi matrix. Our metric for judging which of two channels is a better simulation of a given channel is the diamond norm, $D_\diamond$, between the simulated channel and the channel simulating it. The diamond norm between channels $\mathcal{E}_1$ and $\mathcal{E}_2$ is defined by

$$D_\diamond = \sup_\phi \mathrm{Tr} \left| \mathbb{I} \otimes \mathcal{E}_1(\phi) - \mathbb{I} \otimes \mathcal{E}_2(\phi) \right|, \tag{4.87}$$

where the supremum is taken over all input states $\phi$ (and where the identity is enacted on idler modes of $\phi$). Of particular interest are resource states with tensor-product structure (i.e. $N$ identical copies of a two-qubit state). The simple structure of such states makes it easier to carry out calculations on them for channel simulation. For instance, [5] found that the achievable secret key rate of a quantum channel can be upper bounded by the relative entropy of entanglement (REE) of a resource state that can be used to simulate that channel. If a state has tensor-product structure, the calculation of its REE can be simplified: the REE of such a state is $N$ times the REE of a

single copy of the two-qubit state. Let us refer to all resource states with tensor-product structure as tensor-product resources.

One channel of interest is the AD channel. This channel is characterised by the Choi matrix (for the input state $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$)

$$R(p) = \begin{pmatrix} \frac{p}{2} & 0 & 0 & 0 \\ 0 & \frac{1-p}{2} & -\frac{\sqrt{1-p}}{2} & 0 \\ 0 & -\frac{\sqrt{1-p}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4.88}$$

where $p$ is the probability of a qubit with value one being flipped to a zero. One possible type of resource state is comprised of $N$ copies of this state, $R(p_1)^{\otimes N}$, where $p_1$ is the damping probability of the AD channel used to generate the resource state, i.e. the resource state is $N$ copies of the output Choi matrix of an AD channel with damping probability $p_1$. Note that this is not the same Bell state that we have been using to define the Choi matrix previously; we have previously used the input state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. We have chosen the state $\left|\Phi^{\text{Bell}}\right\rangle$ in this case because it is the resource state $\left|\Phi^{\text{Bell}}\right\rangle \left\langle\Phi^{\text{Bell}}\right|^{\otimes N}$ that simulates the identity channel (due to the structure of the measurement). Consequently, it is the resource $R(p_1)^{\otimes N}$ that asymptotically gives a perfect simulation of the AD channel. Let $p_0$ be the damping probability of the AD channel that we are trying to simulate; this need not necessarily be equal to $p_1$. We denote the Choi matrix of the PBT channel with resource state $\phi$, $PBT[\phi]$. Applying the explicit expressions that we have derived, we find

$$PBT[R(p_1)^{\otimes N}] = \begin{pmatrix} \frac{1}{2} - \frac{\xi_N(1-p_1)}{4} & 0 & 0 & \left(\frac{1}{2} - \frac{\xi_N}{2}\right)\sqrt{1-p_1} \\ 0 & \frac{\xi_N(1-p_1)}{4} & 0 & 0 \\ 0 & 0 & \frac{p_1}{2} - \frac{\xi_N(1-p_1)}{4} & 0 \\ \left(\frac{1}{2} - \frac{\xi_N}{2}\right)\sqrt{1-p_1} & 0 & 0 & (1-p_1)\left(\frac{1}{2} - \frac{\xi_N}{4}\right) \end{pmatrix}, \tag{4.89}$$

where $\xi_N$ is again the depolarisation probability of the channel given by carrying out $N$-port PBT with a maximally entangled resource state (as calculated in Subsection 4.2.5). We will refer to such a resource state ($N$ copies of the Choi matrix of an AD channel, with damping probability generally different from that of the simulated channel) as a Choi resource.

Consider the special case of $p_1 = p_0$ (simulating an AD channel with $N$ copies of its own output Choi matrix); it has been shown that in this case, the diamond norm of the simulated channel from the simulating channel is the same as the trace norm between the Choi matrices [2].

We will denote the diamond norm using this resource as $D_\diamond^0$; it is given by

$$D_\diamond^0 = \xi_N \left( \frac{1-p_0}{2} + \sqrt{1-p_0} \right). \tag{4.90}$$

$\xi_N \leq \frac{6-\sqrt{3}}{6} \simeq 0.71$, since this is the value for 2 ports. $D_\diamond^0$ provides a useful benchmark, since we know it converges to 0 in the limit of infinite ports, and hence $R(p_0)^{\otimes N}$ is a common choice of resource state for calculations involving channel simulation. For instance, in Ref. [2], resource states composed of $N$ copies of the Choi matrix of the simulated channel were used to obtain a general bound on channel discrimination, and this bound was specifically applied to the AD channel.

In the asymptotic limit, in the case of $p_1 = p_0$, the output Choi matrix in Eq. (4.89) tends to the Choi matrix of the simulated channel, as expected. However, for finite $N$, a lower $D_\diamond$ can be achieved by choosing a value of $p_1$ for the resource state different from $p_0$ (the damping probability of the channel we are simulating).

Let us consider for which values of $p_1$ we can know the diamond norm exactly. We have upper and lower bounds on the diamond norm between (qubit) channels with Choi matrices $X$ and $Y$ given by Ref. [1]:

$$\operatorname{Tr}|X-Y| \leq D_\diamond \leq 2 \left\| \operatorname{Tr}_2 |X-Y| \right\|_\infty, \tag{4.91}$$

where the trace is taken over the mode which passed through the channel. These two bounds are equal (and therefore give the exact diamond norm) if the matrix $\operatorname{Tr}_2|X-Y|$ is scalar (proportional to the identity matrix). The difference between the Choi matrices of the simulated and simulating channels, in this case, is

$$PBT[R(p_1)^{\otimes N}] - R'(p_0) = \begin{pmatrix} -e_1 & 0 & 0 & -c \\ 0 & e_1 & 0 & 0 \\ 0 & 0 & e_2 & 0 \\ -c & 0 & 0 & -e_2 \end{pmatrix}, \tag{4.92}$$

$$e_1 = \frac{\xi_N}{4}(1-p_1),\ e_2 = e_1 - \frac{p_0-p_1}{2},\ c = \frac{1}{2}\left( \sqrt{1-p_0} - (1-\xi_N)\sqrt{1-p_1} \right), \tag{4.93}$$

where $R'$ is the Choi matrix for the input state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. If $e_1 = \pm e_2$, the modulus of the matrix, with the trace taken over the second mode, will be scalar. This is true in two cases:

$$p_1 = p_0, \tag{4.94}$$

$$p_1 = \frac{p_0 - \xi_N}{1-\xi_N}. \tag{4.95}$$

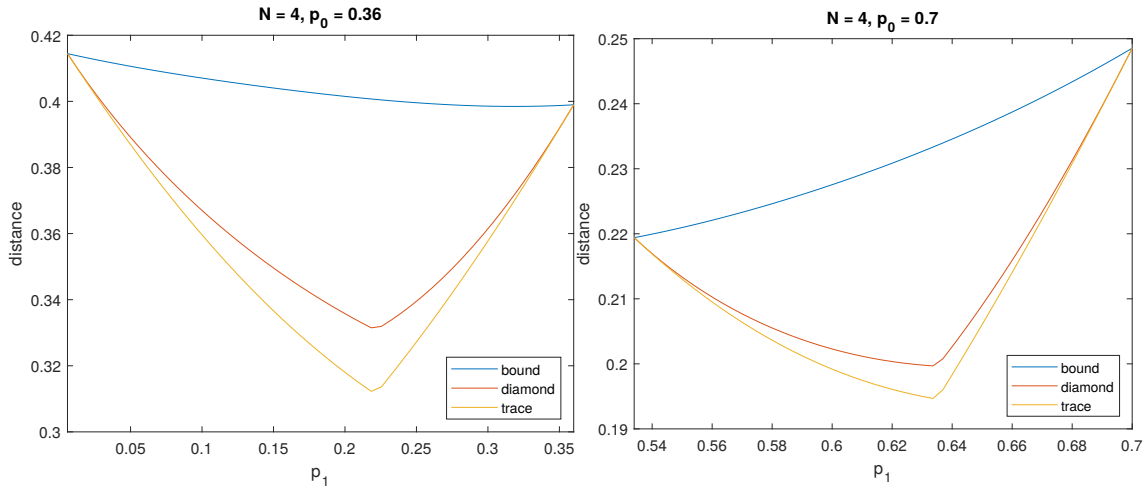


Figure 4.1: The trace norm, the numerically found diamond norm and the analytical upper bound on the diamond norm from Ref. [1] are plotted against $p_1$, the damping value of the AD channel used to produce the resource state, for the resource given in Eq. (4.88). The plot with $p_0 = 0.36$ lies in the regime where $p_1 = p_0$ gives a better simulation than $p_1 = \frac{p_0 - \xi_N}{1-\xi_N}$, and the plot with $p_0 = 0.7$ lies in the regime where the opposite is true. In both cases, the actual minimum of the diamond norm lies between these points and lies near the minimum of the trace norm. In both cases, this minimum of the trace norm lies at exactly $p_1 = \frac{2p_0 - \xi_N}{2-\xi_N}$.

The first case is the known case of $N$ copies of the Choi matrix of the simulated channel. In the second case, we find that the diamond norm, $D^1_\diamond$, is given by

$$D^1_\diamond = \frac{1}{2}\left(\frac{(1-p_0)\xi_N}{1-\xi_N} + \sqrt{4(1-p_0)\left(1-\sqrt{1-\xi_N}\right)^2 + \frac{(1-p_0)^2\xi_N^2}{(1-\xi_N)^2}}\right). \tag{4.96}$$

For sufficiently low values of $\xi_N$ and sufficiently high values of $p_0$, this second expression for the diamond norm, $D^1_\diamond$ is lower than $D^0_\diamond$. Specifically, we find that there is a function in $\xi_N$ separating the two regimes. This function crosses $p_0 = 0$ at a $\xi_N$ value of about 0.237 and for values of $\xi_N < 0.237$, the second expression is always lower (except in the trivial case of $p_0 = 1$). $\xi_N < 0.237$ for a number of ports equal to or greater than 6, so for $N \geq 6$, $D^1_\diamond \leq D^0_\diamond$. Note that if $p_0 < \xi_N$, this second point does not exist, since that would require a negative value of $p_1$. The plots in Fig. 4.1 illustrate these two regimes in the case of 4 ports. We therefore have a resource that simulates a given AD channel better than $N$ copies of the Choi matrix of that channel, for any finite number of ports, with an analytical expression for the diamond norm between the channels.

Asymptotically (in $N$), the right hand side of Eq. (4.95) tends to the right hand side of Eq. (4.94), since $\xi_N$ tends to 0. This is as expected, since we know that the Choi resource with $p_1 = p_0$ simulates the AD channel perfectly in the asymptotic limit of $N$.

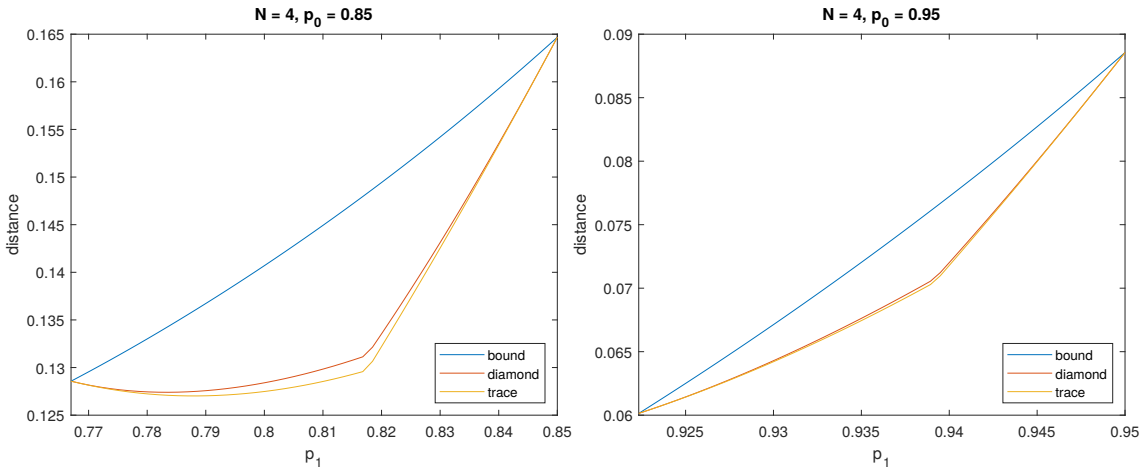


Figure 4.2: The trace norm, the numerically found diamond norm and the analytical upper bound on the diamond norm from Ref. [1] are plotted against $p_1$, the damping value of the AD channel used to produce the resource state, for the resource given in Eq. (4.88). In both of the cases shown, the minimum of the trace norm no longer lies at $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$, but rather at a lower value of $p_1$. In the case of $p_0 = 0.85$, the minimum of the trace norm (and therefore of the diamond norm) still lies between the two points for which the diamond norm is exactly known ($p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and $p_1 = p_0$), whereas for $p_0 = 0.95$, this is no longer the case.

Although we have two points for which the diamond norm is known exactly, this does not mean that the minimum diamond norm for simulating a given channel lies at either of these two points. In fact, we find numerically that the minimum of the diamond norm often lies near the minimum of the trace norm between the Choi matrices, rather than at either of these known points. We also find that for all $p_0 \leq v_1$, where $v_1$ is a function of $\xi_N$ that is always greater than $\frac{2}{5}$, the minimum of the trace norm lies at $\frac{2p_0-\xi_N}{2-\xi_N}$, and that for all $p_0 \leq v_2$, where $v_2$ is a function of $\xi_N$ that is always greater than $\frac{2}{3}$, the minimum of the trace norm lies between $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$. See Section B.1 of Appendix B for more details.

If the minimum of the trace norm lies between $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and $p_1 = p_0$, the two points at which the diamond norm is equal to the trace norm, we are guaranteed that the minimum of the diamond norm will fall between those two points, since the trace norm, which lower bounds the diamond norm, will have no local minima outside of these points. This means that the trace norm will have a negative gradient at every point below $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and a positive gradient at every point above $p_1 = p_0$. The plots in Fig. 4.2 show values of $p_0$ for which the minimum of the trace norm does not lie at $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$.

Whilst the Choi resource with $p_1$ chosen to minimise the diamond norm simulates the AD

channel better than the case of $p_1 = p_0$, the two resources tend towards each other as $N$ increases. A resource state of interest would be one that has tensor-product structure, simulates some AD channel better than the Choi resource and is distinct from the Choi resource for all $p_1$ values. We find that such a resource exists. Let $R_{\text{new}}(a)$ be a two-qubit state, defined by

$$R_{\text{new}}(a) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & -\sqrt{a(1-a)} & 0 \\ 0 & -\sqrt{a(1-a)} & 1-a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4.97}$$

where $a$ is a parameter characterising the density matrix. Consider the resource state $R_{\text{new}}(a)^{\otimes N}$ ($N$ copies of $R_{\text{new}}(a)$, such that each port is a copy of $R_{\text{new}}(a)$). This is a tensor-product resource and the state of each port is clearly different from the state in Eq. (4.88) for all parameter values except for the case of $p = 0$ and $a = \frac{1}{2}$. This resource state illustrates the importance of the explicit expressions for the components of the Choi matrix resulting from PBT: whilst it would be possible to calculate $PBT[R(p)^{\otimes N}]$ by applying an AD channel to the (known) output of the PBT channel using a maximally entangled resource, the same technique cannot be used to calculate $PBT\left[R_{\text{new}}(a)^{\otimes N}\right]$.

Carrying out PBT using this resource state, which we will call the alternate resource, results in the Choi matrix:

$$PBT\left[R_{\text{new}}(a)^{\otimes N}\right] = \begin{pmatrix} x & 0 & 0 & z \\ 0 & \frac{1}{2}-x & 0 & 0 \\ 0 & 0 & y & 0 \\ z & 0 & 0 & \frac{1}{2}-y \end{pmatrix}, \tag{4.98}$$

$$
\begin{aligned}
x = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} & a^{\frac{N+1}{2}+m}(1-a)^{\frac{N-1}{2}-m} \\
& \times \frac{N!\left[\left(\frac{N+1}{2}-s\right)^{-\frac{1}{2}}(s-m)+\left(\frac{N+3}{2}+s\right)^{-\frac{1}{2}}(s+m+1)\right]^2}{2\left(\frac{N-1}{2}-s\right)!\left(\frac{N+1}{2}+s\right)!(2s+1)} \\
& + \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N+1}{2}+m}(1-a)^{\frac{N-1}{2}-m}\frac{\left(\frac{N+1}{2}+m\right)\left(\frac{N+1}{2}-m\right)}{2N(N+1)},
\end{aligned} \tag{4.99}
$$

$$
\begin{aligned}
y = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} & a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m} \\
& \times \frac{N!(s+m)(s-m+1)\left[\left(\frac{N+1}{2}-s\right)^{-\frac{1}{2}}-\left(\frac{N+3}{2}+s\right)^{-\frac{1}{2}}\right]^2}{2\left(\frac{N-1}{2}-s\right)!\left(\frac{N+1}{2}+s\right)!(2s+1)} \\
& + \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\frac{\left(\frac{N-1}{2}+m\right)\left(\frac{N+1}{2}+m\right)}{2N(N+1)},
\end{aligned} \tag{4.100}
$$

$$
\begin{aligned}
z = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} & \frac{a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}N!}{2\left(\frac{N-1}{2}-s\right)!\left(\frac{N+1}{2}+s\right)!(2s+1)}\Bigg[\left(\frac{N+1}{2}-s\right)^{-1}(s^2-m^2) \\
& + 2\left(\frac{N+1}{2}-s\right)^{-\frac{1}{2}}\left(\frac{N+3}{2}+s\right)^{-\frac{1}{2}}(s^2+m^2+s) \\
& + \left(\frac{N+3}{2}+s\right)^{-1}((s+1)^2-m^2)\Bigg] \\
& - \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\frac{\left(\frac{N+1}{2}+m\right)\left(\frac{N+1}{2}-m\right)}{2N(N+1)},
\end{aligned} \tag{4.101}
$$

where $s_{\min}$ is 0 for odd $N$ and $\frac{1}{2}$ for even $N$. The elements of the Choi matrix have been calculated using the expressions in Eqs. (4.38) to (4.40). We can therefore write

$$
PBT\left[R_{\text{new}}(a)^{\otimes N}\right] - R'(p_0) = \begin{pmatrix} x-\frac{1}{2} & 0 & 0 & z-\frac{\sqrt{1-p_0}}{2} \\ 0 & \frac{1}{2}-x & 0 & 0 \\ 0 & 0 & y-\frac{p_0}{2} & 0 \\ z-\frac{\sqrt{1-p_0}}{2} & 0 & 0 & \frac{p_0}{2}-y \end{pmatrix}, \tag{4.102}
$$

Again, we can find the values of $a$ at which this matrix is scalar by finding the points at which $x-\frac{1}{2} = \pm\left(y-\frac{p_0}{2}\right)$. In this case, however, we have a more complicated expression in terms of $a$ and $p_0$, which depends on $N$, making it difficult to find a general (for arbitrary $N$) expression for the diamond norm at these points where the diamond norm is known exactly (however it is simple to find the expression for fixed $N$).

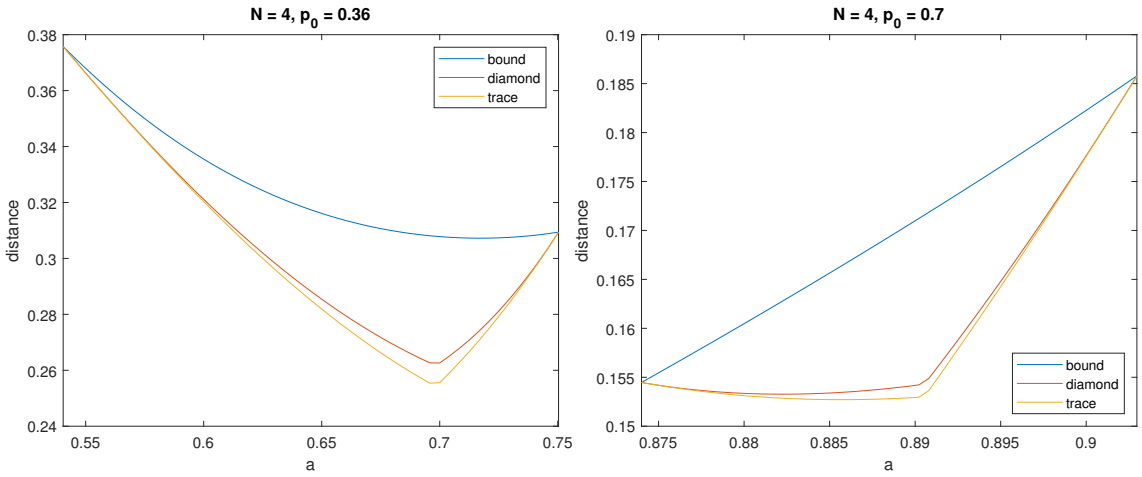


Figure 4.3: The trace norm, the numerically found diamond norm and the analytical upper bound on the diamond norm from Ref. [1] are plotted against $a$, the parameter that parametrises the state in Eq. (4.97). Comparing with Fig. 4.1, we can see that at the "known points" where the diamond norm is known analytically (where the trace norm coincides with the diamond norm), the diamond norm is significantly lower for the resource $R_{\text{new}}(a)^{\otimes N}$ than at the known points for the Choi resource. Further, the minimum diamond norm for this new resource is significantly lower than the minimum diamond norm for the Choi resource.

Using this resource, we can prove that for all $N$ and for some range of $p_0$ values, there exists some tensor-product resource, which is distinct from $R(p)^{\otimes N}$, for which the diamond norm from the AD channel can be found analytically and is smaller than the diamond norm using the resource state $R(p)^{\otimes N}$ for both $p = p_0$ and $p = \frac{p_0 - \xi_N}{1-\xi_N}$. This means that, for any finite value of $N$, there are some (low) values of $p_0$ for which we can find a tensor-product resource state that gives a diamond norm from the AD channel lower than either $D^0_\diamond$ or $D^1_\diamond$. This is demonstrated in Fig. 4.3, for $N = 4$, using the resource state $R_{\text{new}}(a)^{\otimes N}$ and is proven in Section B.2 of Appendix B.

For low $N$, the alternate resource beats the Choi resource over a large range of $p_0$ values and by a significant amount. This can be seen for the case of $N = 6$ in Fig. 4.4. Note that at $a = \frac{1}{2}$ and $p = 0$, the two resources are the same, and these parameter values are the starting points of the graphs in the figure.

Similarly to the case of the Choi resource, we find numerically that for a large range of $p_0$ values, the value of $a$ that gives the minimum of the trace norm coincides with the value that minimises the diamond norm and is the $a$ value for which $y - \frac{p_0}{2} = 0$ (just as, for the Choi resource, the minimum of the trace norm occurs at the value of $p$ that sets $e_2 = 0$, for all $p_0 < \frac{2}{5}$). Numerically we find a trend that there exists a range of $p_0$ values such that the resource state

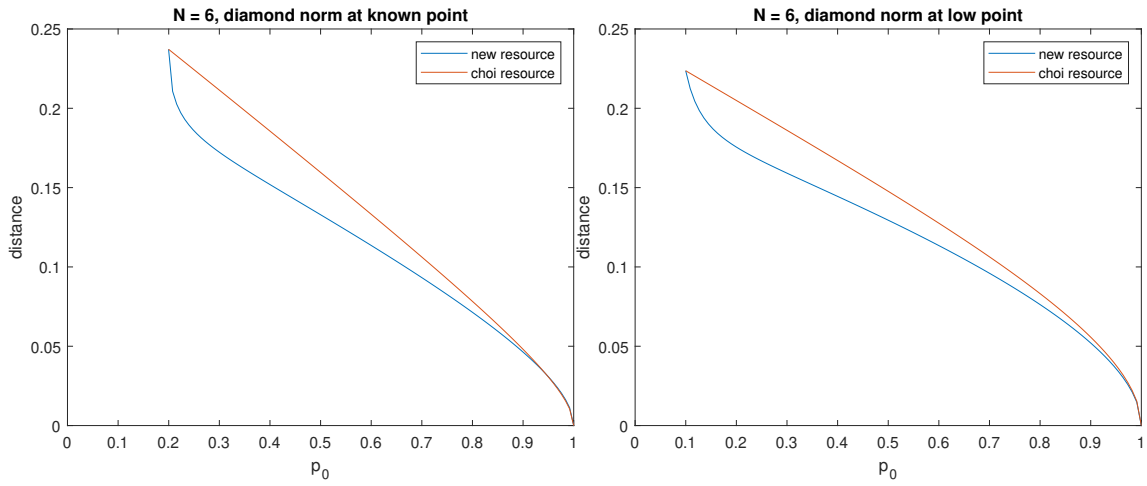


Figure 4.4: The diamond norm is plotted against the damping probability of the AD channel being simulated for PBT with the resource state $R_{\text{new}}(a)^{\otimes N}$ (new resource) and the resource state $R(p_1)^{\otimes N}$ (Choi resource). In the left-hand plot, we choose $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and choose $a$ such that $x(a) - y(a) = \frac{1-p_0}{2}$, so that the trace norm coincides with the diamond norm. In the right hand plot, we choose $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$ and choose $a$ such that $y(a) = \frac{p_0}{2}$; these are close to the optimal parameters to minimise the diamond norm. In both cases, we start at the minimum value of $p_0$ for which $p_1$ is non-negative. The new resource is better than the Choi resource for a large range of $p_0$ values and especially for low $p_0$.

$R_{\text{new}}(a)^{\otimes N}$, with $a$ chosen so that $y = \frac{p_0}{2}$, gives a better simulation of the AD channel (lower diamond norm) than $R(p_1)^{\otimes N}$, for any value of $p_1$. However, this range of $p_0$ values becomes increasingly small as $N$ increases. This has been numerically confirmed for $N < 11$. Specifically, this occurs for low $p_0$.

The explicit expressions for the Choi matrix of the PBT channel therefore allow us to calculate the diamond norm for a resource that simulates certain AD channels better than a tensor-product of Choi matrices.

## 4.5 Summary

Qubit PBT simulates a quantum channel on the teleported qubit, with the channel depending on the resource state used. Using Eqs. (4.38) to (4.41), we can find the Choi matrix for the channel simulated by a given resource state. We assume this resource state to be symmetric under exchange of labels, since this assumption does not restrict the simulable channels. We also provide a simple algorithm for converting to the alternative channel representation of Kraus operators. We show

how the Choi matrix can be easily calculated in the two port case, giving simplified expressions (namely, Eqs. (4.55), (4.57) and (4.58)).

In Eqs. (4.86) and (4.84), we give the Kraus operators that describe the PBT protocol itself (for a fixed number of ports, the square-root measurement and a resource state that is symmetric under exchange of labels). These Kraus operators characterise the map from the $2N$-qubit resource state to the two-qubit Choi matrix and thus offer a complete description of the PBT protocol. This is a complete analytical characterisation, which could be efficiently exploited in works leading on from Ref. [28], where techniques of machine learning and semi-definite programming are employed to find the optimal resource state for PBT (and other teleportation protocols).

We consider simulating the AD channel with PBT and find that, for finite numbers of ports, using $N$ copies of the Choi matrix of the simulated channel as the resource state gives a higher diamond norm than using $N$ copies of the Choi matrix of a different AD channel. We also find that there exist resource states with tensor-product structure that simulate the AD channel better than any Choi resource, in the low damping range. These improved simulations will prove useful in Chapter 5.

In this chapter, we only present results for the qubit case. Future work could explore PBT in the qudit or continuous variable cases. In the qudit case, this is complicated by the Clebsch-Gordan coefficients, which do not take the simple form they take in the qubit case. Clarifying the mathematical aspects of PBT is important for the fundamental role that this protocol plays in various areas of quantum information theory, not only in problems of ultimate channel discrimination [2] but also in communication problems such as position-based quantum cryptography [117, 118].

# Chapter 5

# Bounds on amplitude damping channel discrimination

The work in this chapter forms the basis for a paper published in Physical Review A, whose authors are (in order) Jason Pereira and Stefano Pirandola [10].

In this chapter, we start by discussing the importance of the task of discrimination between amplitude damping (AD) channels. In the second section, we describe the task of binary AD channel discrimination via an adaptive protocol and present the various bounds. We also calculate the diamond norm between AD channels. We then compare the bounds with existing bounds and apply them to two different scenarios. The final section summarises the presented work.

## 5.1 Introduction

Pure loss channels constitute an important class of quantum channels. They can be used as models in many situations in which the environmental noise is low. Examples include quantum communications [5] and quantum metrology [62] (where the parameter being measured could be the loss of the channel). Since the loss of a communications line determines the rate at which a secret key can be exchanged over it (via quantum key distribution), it would be useful to be able to accurately determine the transmissivity of a lossy channel. As seen in Chapter 3, it is difficult to bound the ultimate precision with which an adaptive protocol can discriminate between two bosonic lossy channels, because two lossy channels with different transmissivities are not jointly teleportation-covariant.

AD channels are qubit channels that act similarly to lossy channels: they can be regarded as

lossy channels that only act on qubit states. They map input states, $\rho^{\text{in}}$, according to

$$\text{AD} : \rho^{\text{in}} \to K_0 \rho^{\text{in}} K_0^\dagger + K_1 \rho^{\text{in}} K_1^\dagger, \tag{5.1}$$

$$K_0 = |0\rangle\langle 0| + \sqrt{\eta}\,|1\rangle\langle 1|\,,\; K_1 = \sqrt{1-\eta}\,|0\rangle\langle 1|\,, \tag{5.2}$$

where $p$ is a parameter of the channel and $K_0$ and $K_1$ are the Kraus operators.

In physics, they are good models for energy dissipation in qubit systems [15] and, in quantum information, they can model low noise scenarios where the number of photons passing through a quantum channel is also low. They have also been used as a model for the transfer of a qubit through a spin chain [119]. The task of discriminating between two AD channels is therefore of interest in quantum information science, so it is desirable to bound the error probability of discrimination protocols.

Since adaptivity has been shown to improve the performance of discrimination protocols [24], bounds on the distinguishability of AD channels must take this into account. Generally this can be done using teleportation stretching (see Chapter 2 for more details), but AD channels are not teleportation-covariant. This means that they are an important class of channel that cannot be simulated using standard teleportation.

Pirandola et al. used port-based teleportation (PBT), with a resource state composed of multiple copies of the Choi matrix of the channel, to simulate the AD channel in order to lower bound the error probability for the most general adaptive discrimination protocol [2]. However, this bound is not tight: there is a large gap between the upper bound and the lower bound, leaving a lot of room for improvement (although it is not immediately clear if it is the lower bound, the upper bound, or both that needs tightening).

Since we found in Chapter 4 that there exist resource states that can simulate AD channels better than multiple copies of their Choi matrices (for a fixed number of ports), it is natural to wonder whether the bounds in Ref. [2] can be tightened by simulating the AD channels using these improved resource states instead. That is the focus of this chapter.

## 5.2 Analytical results

Suppose we are given an AD channel, $\mathcal{C}$, which we know to have a transmissivity, $\eta$, equal to either $\eta_X$ or $\eta_Y$ and wish to determine which of these two values $\eta$ takes. Note that an AD channel is the qubit version of a pure loss channel, in that the pointwise application of a hard energy constraint of one photon and a pure loss channel with transmissivity $\eta$ reduces to an AD channel with transmissivity $\eta$ (or damping probability $1-\eta$). Suppose we are allowed to carry

out any protocol involving our channel, but with a maximum of $N$ channel uses. Let $\mathcal{C}_X$ be the AD channel with a transmissivity of $\eta_X$ and let $\mathcal{C}_Y$ be the AD channel with a transmissivity of $\eta_Y$. Our task is to carry out the optimal protocol for discriminating between $\mathcal{C}_X$ and $\mathcal{C}_Y$, subject to the constraint on the total number of channel uses.

A general protocol consists of quantum operations on some initial state, followed by a channel use, followed by further operations (which can include measurements) and further channel uses, until a total of $N$ channel uses have occurred [2]. At this point, a final set of quantum operations is carried out, and then a measurement is made on the final state, which we will label as $\rho_i^{N,out}$ in the case in which the channel is $\mathcal{C}_i$. Note that this protocol is allowed to be adaptive, meaning that each step in the protocol can depend on previous steps. We define the optimal protocol as the protocol for which we maximise the trace norm between $\rho_X^{N,\text{out}}$ and $\rho_Y^{N,\text{out}}$ and then carry out the most discriminating measurement possible. This optimal value of the trace norm is denoted by $D_{\mathcal{C}_X\mathcal{C}_Y}^{\text{opt},N}$. If we have $\mathcal{C}_X$ and $\mathcal{C}_Y$ with equal probabilities, this is the protocol that minimises the probability of error in identifying which channel we have. It is also worth noting that, for the optimal protocol, we can assume without loss of generality that all of the operations between channel uses are unitaries, as any other operations (such as quantum channels, of which measurements are a special case) can be modelled as unitaries, by allowing the user of the protocol to hold the distillation of all operations performed. This cannot decrease the trace norm between output states.

### 5.2.1 Bounding the maximum trace norm using channel simulation

We now apply the technique of channel simulation [2, 5, 28]. Suppose we have a qubit quantum processor $\mathcal{Q}(\pi)$, which takes the resource state $\pi$ as a program and enacts the channel $\mathcal{C}_{\mathcal{Q}(\pi)}$ on an input qubit, via some set of trace-preserving quantum operations. Suppose also that there exist program states $\pi_X$ and $\pi_Y$, such that the enacted channels, $\mathcal{C}_{\mathcal{Q}(\pi_X)}$ and $\mathcal{C}_{\mathcal{Q}(\pi_Y)}$, are sufficiently close to the two AD channels that we want to discriminate between. More precisely, suppose we can write

$$\left\|\mathcal{C}_{\mathcal{Q}(\pi_X)} - \mathcal{C}_X\right\|_\diamond \leq \epsilon_X, \tag{5.3}$$

$$\left\|\mathcal{C}_{\mathcal{Q}(\pi_Y)} - \mathcal{C}_Y\right\|_\diamond \leq \epsilon_Y, \tag{5.4}$$

where we have used the diamond norm between the channels. This is the maximum of the trace norm between the outputs of the channels, maximised over all input states (including those with idlers). Then, we replace the $N$ channel uses in our discrimination protocol with the channel enacted by the processor (with program state $\pi_{X(Y)}$ in the case in which the channel is $\mathcal{C}_{X(Y)}$) and

call the output state of the resulting protocol $\rho^{N,\text{out}}_{\mathcal{Q}(\pi_{X(Y)})}$. We can then write

$$\left\|\rho^{N,\text{out}}_{\mathcal{Q}(\pi_X)} - \rho^{N,\text{out}}_{X}\right\|_1 \leq N\epsilon_X, \tag{5.5}$$

$$\left\|\rho^{N,\text{out}}_{\mathcal{Q}(\pi_Y)} - \rho^{N,\text{out}}_{Y}\right\|_1 \leq N\epsilon_Y. \tag{5.6}$$

Using the fact that all of the operations are trace-preserving, and the only difference between the two cases is the initial program state, we can write

$$\left\|\rho^{N,\text{out}}_{\mathcal{Q}(\pi_X)} - \rho^{N,\text{out}}_{\mathcal{Q}(\pi_Y)}\right\|_1 \leq \left\|\pi_X^{\otimes N} - \pi_Y^{\otimes N}\right\|_1 \tag{5.7}$$

$$\leq 2\sqrt{1 - F\left(\pi_X^{\otimes N}, \pi_Y^{\otimes N}\right)^2}, \tag{5.8}$$

where $F(\rho_1, \rho_2)$ is the quantum fidelity, defined by

$$F(\rho_1, \rho_2) = \text{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}. \tag{5.9}$$

Using the multiplicativity of the fidelity with respect to tensor products, we get

$$\left\|\rho^{N,\text{out}}_{\mathcal{Q}(\pi_X)} - \rho^{N,\text{out}}_{\mathcal{Q}(\pi_Y)}\right\|_1 \leq 2\sqrt{1 - F\left(\pi_X, \pi_Y\right)^{2N}}. \tag{5.10}$$

Finally, using the triangle inequality, we write

$$D^{\text{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \leq N\epsilon_{XY} + 2\sqrt{1 - F\left(\pi_X, \pi_Y\right)^{2N}}, \tag{5.11}$$

$$\epsilon_{XY} = \epsilon_X + \epsilon_Y, \tag{5.12}$$

where $D^{\text{opt},N}_{\mathcal{C}_X\mathcal{C}_Y}$ is maximised over all possible protocols. The trace norm between two states, $D$, is related to the maximum probability of successfully discriminating between them, $p^{\text{succ}}$ via

$$p^{\text{succ}} = \frac{1}{2} + \frac{D}{4}, \tag{5.13}$$

and so we have an upper bound on the probability of discriminating between two AD channels $\mathcal{C}_X$ and $\mathcal{C}_Y$, which holds over all possible adaptive protocols. Alternatively, by defining

$$p^{\text{err}} = 1 - p^{\text{succ}} \tag{5.14}$$

$$= \frac{1}{2} - \frac{D}{4}, \tag{5.15}$$

we have a lower bound on the error probability.

Note that the tightness of this bound depends both on the chosen program states, $\pi_X$ and $\pi_Y$, and on the quantum processor, $\mathcal{Q}$, used. In order to attain a tight bound, we need to both minimise the simulation errors, $\epsilon_X$ and $\epsilon_Y$, and minimise the trace norm between the program states

simultaneously. For instance, we could conceive of a trivial quantum processor that measures the program state in the computational basis and then, depending on the outcome of the measurement, enacts either $\mathcal{C}_X$ or $\mathcal{C}_Y$. Choosing the program states $|0\rangle$ and $|1\rangle$, we get $\epsilon_{X(Y)} = 0$, but the trace norm between the program states is maximised, and hence our bound is too large. More useful bounds can be found with processors that use PBT, as discussed in Subsection 5.2.2, and with a different trivial processor, as discussed in Subsection 5.2.3.

### 5.2.2 Quantum processors for AD channel simulation

As previously mentioned, the tightness of the bound depends on the quantum processor and program states used to simulate the channels. We wish to minimise the simulation error whilst keeping our program states as similar to each other as possible, in order to achieve the tightest possible bound.

One idea that may be intuitively appealing is to use (standard) quantum teleportation [107] to simulate the AD channels. For certain qubit channels (namely, Pauli channels), quantum teleportation using the Choi matrix of the channel as a resource (program state) can perfectly simulate the channel (with a simulation error of 0). The Choi matrix of a qubit channel is the state obtained by sending one half of a Bell pair through the channel.

The issue with this is that standard quantum teleportation cannot simulate non-Pauli channels [78], and so we would have a very high simulation error. This would result in a bound that would be too loose to be useful.

One alternative is to use PBT [80, 81]. PBT uses a combined measurement (the square-root measurement) on an input state and $m$ ports, held by the sender, to teleport the input state to one of $m$ ports, held by the receiver. The receiver then traces over the remaining ports. The process is discussed in more detail in Chapter 4 and in Refs. [11, 80, 81]. The program state is the shared resource state of $2m$ qubits, $m$ of which constitute the sender's ports and $m$ of which constitute the receiver's ports.

A possible program state, in this case, is $m$ copies of the Choi matrix of the AD channel. It is known that in the asymptotic limit of $m \to \infty$, such a simulation becomes perfect. The issue with this is that the trace norm between the program states of the two possible channels increases as the number of copies increases, and so we cannot take the asymptotic limit of $m$. Instead, we can accept some small but non-zero simulation error and try to find the optimal value of $m$ to minimise the total value of the bound.

This is the approach taken by Pirandola et al. in Ref. [2], for calculating a lower bound for

the error probability of discriminating between two AD channels (i.e. the same type of bound that we want to calculate here). We will call the family of bounds that come from PBT simulations using the Choi matrix of the simulated channels as a resource the standard Choi bounds (and will implicitly assume that the optimal value of $m$ has been chosen).

In fact, for finite $m$, there are program states that simulate AD channels better than $m$ copies of the Choi matrix of the simulated channel. This was discussed in Chapter 4 (which draws from the work presented in Ref. [11]), where two classes of resource states capable of providing better simulations of AD channels were described.

The first class uses $m$ copies of the Choi matrix of a different AD channel from the one being simulated as a resource. Specifically, to simulate an AD channel with transmissivity $\eta$, we use $m$ copies of the Choi matrix of the AD channel with transmissivity $\eta'$, where

$$\eta' = \frac{\eta}{1-\xi_m}. \tag{5.16}$$

$\xi_m$ is the PBT coefficient for $m$ ports, as defined in Eq. (11) of Ref. [2], and represents the depolarisation probability when carrying out PBT with a maximally entangled resource state. As such, it is a number between 0 and 1, and consequently $\eta' > \eta$. Our notation here differs from Chapter 4, since we are characterising the AD channels with $\eta$ rather than the damping probability (which is $1-\eta$). Note that we also require $\eta \leq 1-\xi_m$. We will call the bounds deriving from PBT using this resource state the improved Choi bounds.

The second class uses pure resource states, parametrised by a parameter $a$, that take the form

$$R^{alt}(a) = \left(\sqrt{a}\,|01\rangle - \sqrt{1-a}\,|10\rangle\right)^{\otimes m}. \tag{5.17}$$

An advantage that comes from the fact that this resource state is pure is that the trace norm between different program states is analytically calculable (since the upper bound coming from the fidelity is tight). The value of the parameter $a$ is determined by both the damping probability of the AD channel that is being simulated and by the number of ports, $m$, and is chosen so as to minimise the simulation error. We will call the bounds deriving from PBT using this resource state the alternative resource bounds.

In all three cases, we must tune $m$ so as to obtain the tightest bound possible.

### 5.2.3 The trivial bound

We can also formulate a bound based on a trivial processor that simply always enacts the channel $\mathcal{C}_X$. In this case, we have

$$\epsilon_X = 0, \tag{5.18}$$

$$\epsilon_Y = \|\mathcal{C}_X - \mathcal{C}_Y\|_\diamond, \tag{5.19}$$

$$\left\|\pi_X^{\otimes M} - \pi_Y^{\otimes M}\right\|_1 = 0. \tag{5.20}$$

In other words, the bound in Eq. (5.12) simply becomes

$$D^{\text{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \leq N D^\diamond_{\mathcal{C}_X\mathcal{C}_Y}, \tag{5.21}$$

$$D^\diamond_{\mathcal{C}_X\mathcal{C}_Y} = \|\mathcal{C}_X - \mathcal{C}_Y\|_\diamond. \tag{5.22}$$

This is $N$ times the diamond norm between the two channels that we are trying to distinguish between. Note that this bounds is not specific to AD channels and could be applied to any binary discrimination task.

In fact, we can write an alternative and simpler proof that this bound holds. Let

$$S(N,m) = \{\mathcal{C}_X, \mathcal{C}_X, ...\mathcal{C}_X, \mathcal{C}_Y, \mathcal{C}_Y, ...\mathcal{C}_Y\} \tag{5.23}$$

be a sequence of $N$ channels that are either $\mathcal{C}_X$ or $\mathcal{C}_Y$. Specifically, the first $m$ channels are $\mathcal{C}_X$ and the next $N-m$ channels are $\mathcal{C}_Y$. Then let $\mathcal{P}(S(N,m))$ be the output of a fully general and potentially adaptive protocol $\mathcal{P}$, which has a total of $N$ channel uses, where the $i$-th channel use involves sending the signal through the channel that is the $i$-th element of $S$. E.g. if $S(3,2) = \mathcal{C}_X, \mathcal{C}_X, \mathcal{C}_Y$, $\mathcal{P}(S)$ is the output of a discrimination protocol when the channel that we are trying to identify as either $\mathcal{C}_X$ or $\mathcal{C}_Y$ (and which the protocol assumes is always the same) is $\mathcal{C}_X$ for the first two channel uses and is $\mathcal{C}_Y$ for the final channel use. We then have

$$\mathcal{P}(S(N,N)) = \mathcal{P}(\mathcal{C}_X, \mathcal{C}_X, ...\mathcal{C}_X) = \rho_X^{N,\text{out}}, \tag{5.24}$$

$$\mathcal{P}(S(N,0)) = \mathcal{P}(\mathcal{C}_Y, \mathcal{C}_Y, ...\mathcal{C}_Y) = \rho_Y^{N,\text{out}}. \tag{5.25}$$

We therefore want to upper bound the trace norm between $\mathcal{P}(S(N,N))$ and $\mathcal{P}(S(N,0))$. We start by writing

$$\|\mathcal{P}(S(N,N)) - \mathcal{P}(S(N,N-1))\|_1 \leq D^\diamond_{\mathcal{C}_X\mathcal{C}_Y}. \tag{5.26}$$

This is due to the fact that the states are identical prior to the final channel use, the states immediately after the final channel use cannot be further apart than the diamond norm between the two

channels and any subsequent post-processing is the same in both cases and so cannot increase the trace norm between the states.

By a similar argument we have

$$\|\mathcal{P}(S(N,N-1)) - \mathcal{P}(S(N,N-2))\|_1 \leq D^{\diamond}_{\mathcal{C}_X\mathcal{C}_Y}, \tag{5.27}$$

and generalising, we can write

$$\|\mathcal{P}(S(N,i)) - \mathcal{P}(S(N,i-1))\|_1 \leq D^{\diamond}_{\mathcal{C}_X\mathcal{C}_Y}. \tag{5.28}$$

Then, using the triangle inequality, we can write

$$\|\mathcal{P}(S(N,i)) - \mathcal{P}(S(N,i-j))\|_1 \leq (i-j)D^{\diamond}_{\mathcal{C}_X\mathcal{C}_Y}, \tag{5.29}$$

and therefore

$$\|\mathcal{P}(S(N,N)) - \mathcal{P}(S(N,0))\|_1 \leq ND^{\diamond}_{\mathcal{C}_X\mathcal{C}_Y}, \tag{5.30}$$

as required.

The diamond norm between any two AD channels is presented in Subsection 5.2.4.

### 5.2.4 Calculating the diamond norm between two AD channels

We start by making an ansatz that the exact diamond norm between two AD channels can be achieved using a state of the form

$$|\phi(t)\rangle = \sqrt{t}\,|00\rangle + \sqrt{1-t}\,|11\rangle\,, \tag{5.31}$$

where $0 \leq t \leq 1$. The trace norm between two AD channels for such a state, $D^{|\phi(t)\rangle,1}_{\mathcal{C}_X\mathcal{C}_Y}$, is then given by

$$D^{|\phi(t)\rangle,1}_{\mathcal{C}_X\mathcal{C}_Y} = |\eta_X - \eta_Y|(1-t)\left(1 + \sqrt{1 + \frac{4t}{(1-t)x^2}}\right), \tag{5.32}$$

$$x = \sqrt{\eta_X} + \sqrt{\eta_Y}. \tag{5.33}$$

Defining $t_{\max}$ as the value of $t$ that maximises $D^{|\phi(t)\rangle,1}_{\mathcal{C}_X\mathcal{C}_Y}$, we find

$$t_{\max} = \max\{0, 1 - \frac{1}{2-x}\}. \tag{5.34}$$

From this, we can see that the problem is split into two regimes: one in which $t_{\max} = 0$ and one in which $t_{\max} > 0$. Note that for the extremal case of $t = 0$, the idler mode is not necessary, since

the state given by Eq. (5.31) is separable for this parameter value. Consequently, probing with $|1\rangle$ is equivalent to probing with $|11\rangle$. The first regime occurs when

$$\sqrt{\eta_x} + \sqrt{\eta_Y} > 1. \tag{5.35}$$

We can then calculate the trace norms for each regime:

$$D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} = 2|\eta_X - \eta_Y|, \tag{5.36}$$

$$D^{\diamond,t>0}_{\mathcal{C}_X\mathcal{C}_Y} = \frac{2|\sqrt{\eta_X} - \sqrt{\eta_Y}|}{2 - (\sqrt{\eta_X} + \sqrt{\eta_Y})}. \tag{5.37}$$

The next step is to prove that the expressions in Eqs. (5.36) and (5.37) are the diamond norms in each regime. We do this using semidefinite programming. In Ref. [64], Watrous showed that finding the diamond norm can be reduced to a semidefinite programming problem. In a semidefinite programming problem, some matrices must be chosen, subject to constraints, to maximise or minimise a quantity that is dependent on these matrices. More specifically, every problem consists of a primal and a dual problem. Each valid solution to the primal problem provides a lower bound to the quantity, and so one maximises over the primal problem. Each valid solution to the dual problem provides an upper bound to the quantity, and so one minimises over it. Therefore, in order to show that Eqs. (5.36) and (5.37) give the diamond norm, we must find matrices satisfying the constraints of the dual problem for the diamond norm that give the expressions in Eqs. (5.36) and (5.37) as the diamond norm. The dual problem is to find positive matrices $Y_0$ and $Y_1$ that satisfy the constraint

$$M = \begin{pmatrix} Y_0 & -J(\mathcal{C}_X, \mathcal{C}_Y) \\ -J(\mathcal{C}_X, \mathcal{C}_Y) & Y_1 \end{pmatrix} > 0, \tag{5.38}$$

where $J$ is the Choi matrix of channel $\mathcal{C}_X$ minus the Choi matrix of channel $\mathcal{C}_Y$, multiplied by the dimension of the input system (which, in our case, is 2). The upper bound on the diamond norm, which we must then minimise, is

$$D^{\diamond}_{\mathcal{C}_X\mathcal{C}_Y} \leq \frac{\|\mathrm{Tr}_S(Y_0)\|_\infty + \|\mathrm{Tr}_S(Y_1)\|_\infty}{2}, \tag{5.39}$$

where the partial trace is taken over the signal (rather than the idler) mode and the norm is the operator norm (i.e. the largest eigenvalue). In our case, we have

$$J(\mathcal{C}_X, \mathcal{C}_Y) = \begin{pmatrix} 0 & 0 & 0 & \sqrt{\eta_X} - \sqrt{\eta_Y} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \eta_Y - \eta_X & 0 \\ \sqrt{\eta_X} - \sqrt{\eta_Y} & 0 & 0 & \eta_X - \eta_Y \end{pmatrix}. \tag{5.40}$$

Let us first consider the $t = 0$ case. Consider the matrices

$$Y_0^{t=0} = \begin{pmatrix} D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} & 0 & 0 & |\sqrt{\eta_X} - \sqrt{\eta_Y}| \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2}D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} & 0 \\ |\sqrt{\eta_X} - \sqrt{\eta_Y}| & 0 & 0 & \frac{1}{2}D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} \end{pmatrix}, \tag{5.41}$$

$$Y_1^{t=0} = Y_0^{t=0}. \tag{5.42}$$

We can immediately see that

$$\mathrm{Tr}_S\left(Y_0^{t=0}\right) = \mathrm{Tr}_S\left(Y_1^{t=0}\right) = D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.43}$$

and so the upper bound on the diamond norm coming from this solution is equal to the expression in Eq. (5.36). The distinct, non-zero eigenvalues of $M^{t=0}$ are

$$e^1_{M^{t=0}} = 2|\eta_X - \eta_Y|, \tag{5.44}$$

$$e^2_{M^{t=0}} = 2\left(|\eta_X - \eta_Y| - |\sqrt{\eta_X} - \sqrt{\eta_Y}|\right), \tag{5.45}$$

$$e^3_{M^{t=0}} = 2\left(|\eta_X - \eta_Y| + |\sqrt{\eta_X} - \sqrt{\eta_Y}|\right), \tag{5.46}$$

the smallest of which is $e^2_{M^{t=0}}$. Since $e^2_{M^{t=0}} > 0$ for the regime in which $t = 0$ (i.e. for $x > 1$), $M^{t=0} > 0$ in this regime, as required. The non-zero eigenvalues of $Y_0^{t=0}$ (and $Y_1^{t=0}$) are

$$e^1_{Y^{t=0}} = |\eta_X - \eta_Y|, \tag{5.47}$$

$$e^2_{Y^{t=0}} = \frac{|\eta_X - \eta_Y|}{2}\left(3 - \sqrt{1 + \frac{4}{x^2}}\right), \tag{5.48}$$

$$e^3_{Y^{t=0}} = \frac{|\eta_X - \eta_Y|}{2}\left(3 + \sqrt{1 + \frac{4}{x^2}}\right). \tag{5.49}$$

$e^2_{Y^{t=0}}$ is the smallest of these, and $e^2_{M^{t=0}} > 0$ in the regime in which $t = 0$, so both $Y_0^{t=0}$ and $Y_1^{t=0}$ are positive. Therefore, Eq. (5.36) gives the exact diamond norm for the $t = 0$ regime.

Next, we consider the $t > 0$ case. Consider the matrices

$$Y_0^{t>0} = \begin{pmatrix} D^{\diamond,t>0}_{\mathcal{C}_X\mathcal{C}_Y} & 0 & 0 & \frac{|\eta_X - \eta_Y|}{2-x} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2}D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} & 0 \\ \frac{|\eta_X - \eta_Y|}{2-x} & 0 & 0 & D^{\diamond,t>0}_{\mathcal{C}_X\mathcal{C}_Y} - \frac{1}{2}D^{\diamond,t=0}_{\mathcal{C}_X\mathcal{C}_Y} \end{pmatrix}, \tag{5.50}$$

$$Y_1^{t>0} = Y_0^{t>0}. \tag{5.51}$$

Tracing over the signal mode, we get

$$\mathrm{Tr}_S\left(Y_0^{t>0}\right) = \mathrm{Tr}_S\left(Y_1^{t>0}\right) = D^{\diamond,t>0}_{\mathcal{C}_X\mathcal{C}_Y}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{5.52}$$

and so the upper bound on the diamond norm coming from this solution is equal to the expression in Eq. (5.37). The non-zero eigenvalues of $M^{t>0}$ are

$$e^1_{M^{t=0}} = 2|\eta_X - \eta_Y|, \tag{5.53}$$

$$e^2_{M^{t=0}} = \frac{4|\sqrt{\eta_X} - \sqrt{\eta_Y}|}{2-x}, \tag{5.54}$$

$$e^3_{M^{t=0}} = \frac{4|\sqrt{\eta_X} - \sqrt{\eta_Y}|}{2-x} - 2|\eta_X - \eta_Y|, \tag{5.55}$$

which are all positive. The non-zero eigenvalues of $Y_0^{t>0}$ (and $Y_1^{t>0}$) are

$$e^1_{Y^{t>0}} = |\eta_X - \eta_Y|, \tag{5.56}$$

$$e^2_{Y^{t>0}} = \frac{|\sqrt{\eta_X} - \sqrt{\eta_Y}|}{2(2-x)}\Big(3 + (1-x)^2 - x\sqrt{4 + (2-x)^2}\Big), \tag{5.57}$$

$$e^3_{Y^{t>0}} = \frac{|\sqrt{\eta_X} - \sqrt{\eta_Y}|}{2(2-x)}\Big(3 + (1-x)^2 + x\sqrt{4 + (2-x)^2}\Big). \tag{5.58}$$

These are again all positive, proving that Eq. (5.37) gives the exact diamond norm for the $t > 0$ regime.

A logical next step would be to calculate the diamond norm between multiple uses of two AD channels, i.e. between the two channels $\mathcal{C}_X^{\otimes N}$ and $\mathcal{C}_Y^{\otimes N}$. However, this is a more difficult task. Numerically, we find that input states of the form $|1\rangle^{\otimes N}$ achieve the diamond norm in some cases, as for the single use case, but that the regimes are more complicated to characterise. Further, for large numbers of channel uses, numerically finding the diamond norm via semidefinite programming is computationally expensive.

### 5.2.5 Lower bounds on the optimal trace norm

It is helpful to also find lower bounds on the maximum trace norm between protocol outputs, $D^{\mathrm{opt},N}_{\mathcal{C}_X\mathcal{C}_Y}$ since this allows us to assess how tight our upper bounds are. One option is to find the diamond norm between $\mathcal{C}_X^{\otimes N}$ ($N$ copies of $\mathcal{C}_X$) and $\mathcal{C}_Y^{\otimes N}$. The only reason that such a lower bound would not be tight is if adaptivity between rounds adds to the discriminative power of a protocol (it is not yet known whether this is the case). The problem with using such a bound is that it is difficult to find the diamond norm for $N > 1$ (as discussed in Subsection 5.2.4).

An alternative is to consider specific protocols that could be implemented and to find the trace norms between outputs in these cases. Since we are looking for the maximum trace norm over all possible protocols, any specific protocol provides a lower bound on this maximum.

Pirandola et al. [2] provided a lower bound on the trace norm between protocol outputs, based on consideration of a non-adaptive protocol, in which $N$ copies of a Bell state are sent through the channel. The output of this protocol is $N$ copies of the Choi matrix of the channel. They found that

$$D^{\mathrm{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \geq 2\left(1 - f^{\mathrm{Choi}}(\eta_X, \eta_Y)^N\right), \tag{5.59}$$

$$f^{\mathrm{Choi}}(p,q) = \frac{1 + \sqrt{(1-p)(1-q)} + \sqrt{pq}}{2}. \tag{5.60}$$

$f^{\mathrm{Choi}}$ is the fidelity between the Choi matrices of channels $\mathcal{C}_X$ and $\mathcal{C}_Y$. We refer to this as the Bell state lower bound.

In fact, we find that we can obtain a slightly tighter bound using an alternative, non-adaptive protocol, in which $N$ copies of the state $|1\rangle$ are sent through the channel. Note that this is also the input state that achieves the maximum quantum Fisher information (QFI) per channel use [120] and so is the optimal input state for parameter estimation, at least in the asymptotic limit of a large number of channel uses. In this case, we obtain the tighter bound

$$D^{\mathrm{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \geq 2\left(1 - f^{|1\rangle}(\eta_X, \eta_Y)^N\right), \tag{5.61}$$

$$f^{|1\rangle}(p,q) = \sqrt{(1-p)(1-q)} + \sqrt{pq}. \tag{5.62}$$

This bound is again based on the fidelity between the possible outputs of the protocol.

For sufficiently small $N$, we can do better still by calculating the exact trace norm for this protocol (rather than lower bounding it). Since the output state of the protocol takes the form

$$\rho^{N,\mathrm{out}}_{X(Y)} = \left((1-\eta_{X(Y)})\,|0\rangle\langle 0| + \eta_{X(Y)}\,|1\rangle\langle 1|\right)^{\otimes N}, \tag{5.63}$$

for channel $\mathcal{C}_{X(Y)}$, the trace norm between the two possible outputs, $D^{|1\rangle,N}_{\mathcal{C}_X\mathcal{C}_Y}$, is

$$D^{|1\rangle,N}_{\mathcal{C}_X\mathcal{C}_Y} = \sum_{i=0}^{N} \binom{N}{i} \left|\eta_X^{N-i}(1-\eta_X)^i - \eta_Y^{N-i}(1-\eta_Y)^i\right|. \tag{5.64}$$

We will refer to this bound as the improved lower bound. The problem with using this bound for large $N$ is that the binomial coefficients become large and therefore computationally difficult to calculate.

When applying the trace norm bounds to channel discrimination, the improved lower bound can be approximated using the quantum Cramér-Rao bound (QCRB), as per [3]. The QCRB lower

bounds the error-variance for estimating a channel parameter, based on the QFI with respect to that parameter. In our case, we have

$$\sigma_\eta^2 \leq \frac{1}{NH_\eta}, \tag{5.65}$$

where $\sigma_\eta^2$ is the variance of an estimation of $\eta$ around its true value and $H_\eta$ is the QFI with respect to $\eta$. As shown in Ref. [120], the optimal QFI is achieved using number states with the maximum number of photons per channel use. In our case, this is the state $|1\rangle$ and the maximum QFI per channel use is

$$H_\eta^{\max} = \frac{1}{\eta(1-\eta)}. \tag{5.66}$$

The QCRB therefore takes the form

$$\sigma_\eta^2 \leq \frac{\eta(1-\eta)}{N}. \tag{5.67}$$

We can return to a binary hypothesis testing scenario by picking a threshold value, $\tau$, of $\eta$, such that if our estimation of $\eta$ is greater than $\tau$, we decide that we have channel $\mathcal{C}_X$ (for $\eta_X > \eta_Y$), and if not, we decide that we have channel $\mathcal{C}_Y$. We assume that our estimation of $\eta$, $\eta'$, follows a Gaussian probability distribution, centred on the true value of $\eta$, with a variance equal to the lower bound from Eq. (5.67). For $\mathcal{C}_{X(Y)}$, this distribution is given by

$$p_{\eta_{X(Y)}}(\eta') = \frac{1}{\sigma_\eta\sqrt{2\pi}} e^{-\frac{\left(\eta'-\eta_{X(Y)}\right)^2}{2\sigma^2_{\eta_{X(Y)}}}}. \tag{5.68}$$

Spedalieri et al. then calculated the probabilities of deciding we have channel $\mathcal{C}_Y$ when we have channel $\mathcal{C}_X$ ($p_X^{\text{err}}$) and of deciding we have channel $\mathcal{C}_X$ when we have channel $\mathcal{C}_Y$ ($p_Y^{\text{err}}$) [3]. These error probabilities are

$$p_X^{\text{err}} = \mathcal{N}_X^{-1} \int_0^\tau p_{\eta_X}(\eta')d\eta', \tag{5.69}$$

$$p_Y^{\text{err}} = \mathcal{N}_Y^{-1} \int_\tau^1 p_{\eta_Y}(\eta')d\eta', \tag{5.70}$$

$$\mathcal{N}_{X(Y)} = \int_0^1 p_{\eta_{X(Y)}}(\eta')d\eta', \tag{5.71}$$

where the normalisation factors, $\mathcal{N}_{X(Y)}$, are due to restricting the probability distributions to the range $[0, 1]$ and again assuming $\eta_X > \eta_Y$. In the case in which both channels have prior probabilities, we can then choose the value of $\tau$ that minimises the mean of these two errors, in order to find the total error probability (in the asymmetric case, we can minimise a weighted mean of the errors). This then gives us an estimate of the error probability obtained using the improved lower bound on the trace norm.

It should be noted that this estimate is only tight for a large number of channel uses. For a small number of channel uses, the QCRB is often not tight [121]. Since Eq. (5.67) lower bounds the variance of the parameter estimates (rather than upper bounding them), we do not attain an upper bound on the error probability, but rather an estimate of the error probability attained using the upper bound in Eq. (5.64) (since we are using the same input states in both cases, and the trace norm between the output states gives the lowest possible error in discriminating between them). In fact, for low $N$, the estimate of the bound based on the QCRB, which we will call the QCRB bound, underestimates the minimum error probability over a range of values. We will therefore only apply it for large $N$ ($> 100$). The advantage of using it in this range is that it is more easily calculated than the upper bound in Eq. (5.64), whilst being significantly tighter than the bound on the error probability attained using Eq. (5.61).

### 5.2.6 Upper bounds from PBT simulations

We now calculate the upper bounds based on PBT simulations of the AD channel. We consider three types of resource state, as mentioned in Subsection 5.2.2.

The first type is the Choi matrix of the simulated channels, resulting in the upper bounds in Ref. [2], which we call the standard Choi bounds. These bounds are

$$D^{\mathrm{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \leq N\epsilon^{\mathrm{std}}_{m,XY} + 2\sqrt{1 - f^{\mathrm{Choi}}(\eta_X,\eta_Y)^{2mN}}, \tag{5.72}$$

$$\epsilon^{\mathrm{std}}_m(\eta) = \xi_m\left(\frac{\eta}{2} + \sqrt{\eta}\right), \tag{5.73}$$

$$\epsilon^{\mathrm{std}}_{m,XY} = \epsilon^{\mathrm{std}}_m(\eta_X) + \epsilon^{\mathrm{std}}_m(\eta_Y), \tag{5.74}$$

where $m$ can take any positive, integer value. Note that we have a family of bounds, since we have a bound for any value of the number of ports, $m$. We must then optimise over $m$ to achieve the tightest possible bound in this family.

The second type of resource state is similar to the first, but $\eta_X$ and $\eta_Y$ have been replaced by $\eta'_X$ and $\eta'_Y$, according to Eq. (5.16). The reason we choose this value of $\eta'_{X(Y)}$ is that this is one of the points at which the diamond norm between the channel and its simulation coincides with the trace norm. This means that we have an analytical expression for the resulting family of trace norm bounds, which we call the improved Choi bounds. Further, for all values of $m \geq 6$, the

simulation errors are lower than for the standard Choi resource. These bounds are

$$D^{\mathrm{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \leq N\epsilon^{\mathrm{imp}}_{m,XY} + 2\sqrt{1 - f^{\mathrm{Choi}}(\eta'_X, \eta'_Y)^{2mN}}, \tag{5.75}$$

$$\epsilon^{\mathrm{imp}}_m(\eta) = \frac{1}{2}\left(\frac{(\eta)\xi_m}{1-\xi_m} + \sqrt{4(\eta)\left(1-\sqrt{1-\xi_m}\right)^2 + \frac{\eta^2\xi_m^2}{(1-\xi_m)^2}}\right), \tag{5.76}$$

$$\epsilon^{\mathrm{imp}}_{m,XY} = \epsilon^{\mathrm{imp}}_m(\eta_X) + \epsilon^{\mathrm{imp}}_m(\eta_Y). \tag{5.77}$$

In fact, since the chosen values of $\eta'_X$ and $\eta'_Y$ are not necessarily the values that give the tightest possible bounds, we could numerically minimise over all pairs of "Choi-like" resources simulating $\mathcal{C}_X$ and $\mathcal{C}_Y$. In other words, we could simulate $\mathcal{C}_X$ with $R^{\mathrm{Choi}}(\eta''_X)$ and $\mathcal{C}_Y$ with $R^{\mathrm{Choi}}(\eta''_Y)$, where

$$R^{\mathrm{Choi}}(\eta) = C(\eta)^{\otimes m}, \tag{5.78}$$

$$C(\eta) = \begin{pmatrix} \frac{1-\eta}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{\eta}}{2} & 0 \\ 0 & -\frac{\sqrt{\eta}}{2} & \frac{\eta}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.79}$$

We could then numerically minimise over $\eta''_X$ and $\eta''_Y$ to find the optimal resource states. However, in this case, we would not have an analytical expression for the simulation error and would need to calculate it numerically, by finding the diamond norm for both simulations; this involves maximising the trace norm between the channels and their simulations over all possible input states. We would then also need to minimise the bounds over $m$. This would involve a lot more numerical minimisation/maximisation than simply numerically optimising the standard and improved Choi bounds over $m$. This is why we do not find optimal bounds for this more general resource.

Finally, we consider resource states, $R^{\mathrm{alt}}(a)$, of the form given in Eq. (5.17). The Choi matrix of the channel simulated by carrying out PBT with this resource state is given in Chapter 4 (see also [11]). Writing the Choi matrix in the form

$$\rho^{\mathrm{Choi}}_{\mathrm{PBT}(R^{\mathrm{alt}}(a))} = \begin{pmatrix} x & 0 & 0 & z \\ 0 & \frac{1}{2}-x & 0 & 0 \\ 0 & 0 & y & 0 \\ z & 0 & 0 & \frac{1}{2}-y \end{pmatrix}, \tag{5.80}$$

where the expressions for $x$, $y$ and $z$ are functions of $a$ (and $m$), which are given in Chapter 4 (Eqs. (4.99) to (4.101)), we choose $a_{X(Y)}$ such that

$$x\left(a_{X(Y)}\right) - y\left(a_{X(Y)}\right) = \frac{\eta_{X(Y)}}{2}. \tag{5.81}$$

We then simulate $\mathcal{C}_{X(Y)}$ with $a_{X(Y)}$ and call the resulting bounds the alternative resource bounds. We choose this value of $a_{X(Y)}$ because this is one of the points at which the diamond norm between the channel and its simulation coincides with the trace norm. Similarly to the case of the "Choi-like" resources, we could minimise our bound over all possible values of $a_{X(Y)}$, rather than choosing this value, but this would again require a lot more numerical minimisation/maximisation. The simulation error is given by

$$\epsilon^{\text{alt}}(a_{X(Y)}) = 1 - \eta_{X(Y)} - 2y + \sqrt{(1 - \eta_{X(Y)} - 2y)^2 + (\sqrt{\eta_{X(Y)}} - 2z)^2}\Big|_{a=a_{X(Y)}}, \quad (5.82)$$

and the fidelity between the resource states is given by

$$f^{\text{alt}}(a_X, a_Y) = \sqrt{a_X a_Y} + \sqrt{(1 - a_X)(1 - a_Y)}. \quad (5.83)$$

The alternative resource bounds are therefore given by

$$D^{\text{opt},N}_{\mathcal{C}_X\mathcal{C}_Y} \leq N\epsilon^{\text{alt}}_{XY} + 2\sqrt{1 - f^{\text{alt}}(a_X, a_Y)^{2mN}}, \quad (5.84)$$

$$\epsilon^{\text{alt}}_{XY} = \epsilon^{\text{alt}}(a_X) + \epsilon^{\text{alt}}(a_Y). \quad (5.85)$$

Although this may not be immediately apparent from the expressions, both $\epsilon^{\text{alt}}_{XY}$ and $f^{\text{alt}}$ depend on $m$, since the expressions for $x$, $y$, $z$ and $a_{X(Y)}$ all depend on $m$. Therefore, we again want to pick the optimal value of $m$, in order to achieve the tightest possible bound. Note that the resource states are pure, meaning that our expression for the trace norm between different resource states is exact.

In order to optimise over $m$, we use analytical functions that closely approximate the standard and improved Choi bounds, but that do not feature $\xi_m$. This is done because the expression for $\xi_m$ is too complicated to easily find an analytical minimum of the full bounds. Specifically, we replace $\xi_m$ in the simulation error expressions with $m^{-1}$; this gives us expressions that we can easily locate the minima of, for fixed $\eta_X$, $\eta_Y$ and $N$. We then use the closest integer values of $m$ to our minima when calculating the actual values of the bounds (substituting them into the original expressions). When referring to the standard or improved Choi bound in Section 5.3, it is implicit that this process has been carried out, and that the bounds are calculated for the optimal value of $m$. For the alternative resource bounds, we find numerically that the bound gets tighter as $m$ increases, rather than having a maximum, so we pick a fixed, high value of $m$.

### 5.2.7 Extending to the qudit case

We will briefly consider the case in which we must discriminate between two pure loss qudit channels, rather than two AD channels (which are pure loss qubit channels). The Stinespring

dilation of such a channel is a beamsplitter acting on an environmental vacuum mode. The action of the beamsplitter can be described as

$$|n\rangle_S |0\rangle_E \to \sum_{i=0}^{n} \sqrt{\eta^{n-i}(1-\eta)^i \binom{n}{i}} |n-i\rangle_S |i\rangle_E , \tag{5.86}$$

where $S$ labels the signal mode (the input mode to the channel), $E$ labels the environmental mode and $\eta$ is the transmissivity of the beamsplitter (and the channel). The binomial coefficient on the right-hand side of the expression comes from the choice of which photons are transferred to the environmental modes. This means that a $d$-dimensional, pure loss channel, with transmissivity $\eta$, can be described by the $d$ Kraus operators

$$K_j = \sum_{i=j}^{d-1} \sqrt{\eta^{i-j}(1-\eta)^j \binom{i}{j}} |i\rangle \langle i-j| , \tag{5.87}$$

where the label $j$ ranges from 0 to $d-1$.

Calling our pure loss $d$-dimensional channels $\mathcal{C}_X^d$ and $\mathcal{C}_Y^d$, with transmissivities of $\eta_X$ and $\eta_Y$ respectively, the $J$-matrix of the two channels (the difference between the Choi matrices, multiplied by the input dimension of the channel, as per Subsection 5.2.4) can be written as

$$J(\mathcal{C}_X^d, \mathcal{C}_Y^d) = \sum_{i=0}^{d-1} \left( |v_{\eta_X}^i\rangle \langle v_{\eta_X}^i| - |v_{\eta_Y}^i\rangle \langle v_{\eta_Y}^i| \right)_{SI} , \tag{5.88}$$

$$|v_\eta^i\rangle_{SI} = \sum_{j=0}^{d-i-1} \sqrt{\eta^j(1-\eta)^i \binom{i+1}{j}} |i+j\rangle_S |j\rangle_I , \tag{5.89}$$

where $S$ labels the signal mode and $I$ labels the idler mode.

In order to calculate simulation bounds in the qudit case, we would require expressions for the output of qudit PBT channels and would require new resource states capable of simulating pure loss qudit channels. However, the trivial bound, based on the diamond norm, can still be used in the qudit case (substituting the diamond norm between channels $\mathcal{C}_X$ and $\mathcal{C}_Y$, in Eq. (5.21), with the diamond norm between $\mathcal{C}_X^d$ and $\mathcal{C}_Y^d$).

We do not have an analytical expression for the diamond norm between channels $\mathcal{C}_X^d$ and $\mathcal{C}_Y^d$, as we do for the qubit case. Instead, we can find it numerically, using semidefinite programming, using the formula for the difference between Choi matrices, given in Eq. (5.88). The issue here is the same as with finding the diamond norm for multiple channel uses. As the input dimension becomes large (i.e. for large $d$), numerically finding the diamond norm becomes computationally expensive.

An alternative is to bound the diamond norm, using a result from Ref. [1]. Nechita et al.

showed that the diamond norm between any two channels, $\mathcal{A}$ and $\mathcal{B}$, can be bounded by

$$\|\mathcal{A}-\mathcal{B}\|_{\diamond} \leq \|\mathrm{Tr}_S\,|J(\mathcal{A},\mathcal{B})|\|_{\infty}\,, \tag{5.90}$$

where $J(\mathcal{A},\mathcal{B})$ is the difference between the Choi matrices of $\mathcal{A}$ and $\mathcal{B}$, multiplied by the input dimension of the channels. Here we have first taken the absolute value of the matrix $J(\mathcal{A},\mathcal{B})$, then taken the partial trace over the signal mode. We have then taken the largest eigenvalue (the operator norm) of the resulting matrix as our bound. Note that this coincides with the trace norm and is therefore exactly the diamond norm, if the matrix is scalar after the partial trace is taken. Applying this bound to the expression in Eq. (5.88) gives a computationally cheaper (but less tight) bound on the diamond norm (and hence on the optimal trace norm between protocol outputs) than finding the diamond norm numerically, via semidefinite programming.

Numerical investigation shows that the diamond norm between two qudit channels, $\mathcal{C}_X^d$ and $\mathcal{C}_Y^d$, appears to coincide with the diamond norm between $d-1$ uses of two qubit channels, $\mathcal{C}_X^{\otimes d-1}$ and $\mathcal{C}_Y^{\otimes d-1}$ (for the same transmissivities, $\eta_X$ and $\eta_Y$). This suggests some connection between the two cases, however it is not clear what the connection is.

It is also worth noting that the (approximation of the) upper bound on the error probability of discriminating between two equiprobable channels attained by using the QCRB still holds in the qudit case (as long as the number of channels is large enough for the approximation to be valid), with the only change being to the lower bound on the channel parameter variance, in Eq. (5.67). In this equation, $N(d-1)$ is substituted for $N$. This is because the maximum QFI per channel use is

$$H_{\eta,d}^{\max} = \frac{d-1}{\eta(1-\eta)}, \tag{5.91}$$

and the input state that attains this value is $|d-1\rangle$ [3, 120]. The QFI is additive, so the maximum value of the total QFI is the same for both lossy channels and AD channels, as long as the total number of photons sent through the channel is the same.

## 5.3 Numerical investigations

Carrying out numerical PBT simulations for our three classes of resource states over a variety of $\eta_X$, $\eta_Y$ and $N$ values, we find that the improved Choi bound beats the standard Choi bound over the entire range of investigated parameter values. We also find that the alternative resource bound, for a sufficiently large number of ports, $m$, beats the both Choi bounds across almost the entire range, and that the trivial bound also beats the Choi bounds over a wide range of values.

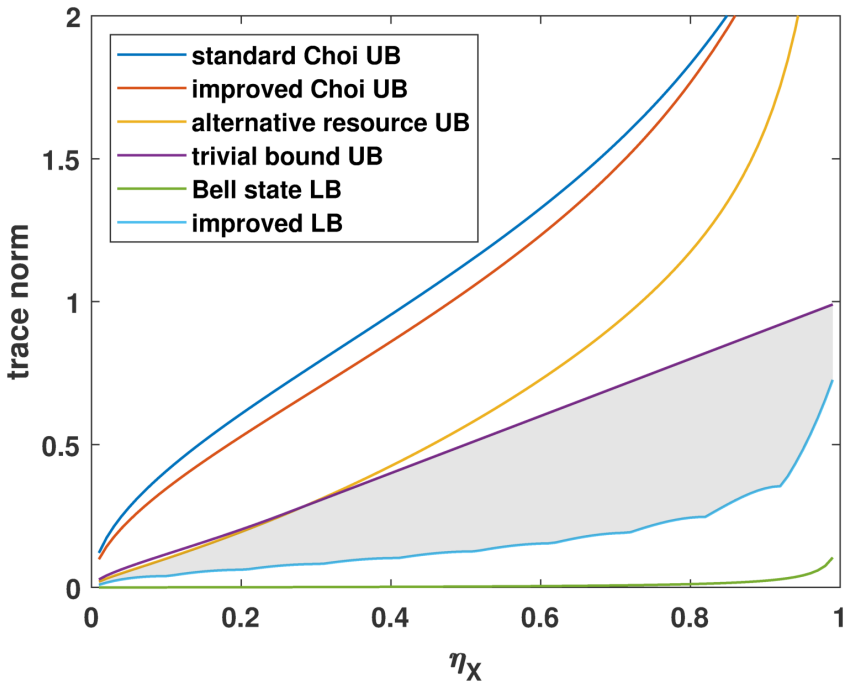


Figure 5.1: Upper and lower bounds on the maximum value of the trace norm between the two possible outputs of an adaptive discrimination protocol with no more than 10 channel uses. The channels being discriminated between are AD channels with transmissivities $\eta_X$ and $\eta_Y$, where $\eta_Y = \eta_X \eta_{XY}$. In this case, $\eta_{XY} = 0.95$. The two upper bounds based on PBT simulations using "Choi-like" resources are significantly less tight than the trivial (upper) bound and the upper bound based on PBT simulations using the alternative resource. Each of these latter two bounds is optimal over some range of $\eta_X$ values. The improved lower bound is tighter than the Bell state lower bound. The grey shaded area is the region between the tightest upper and lower bounds.

In fact, either the trivial bound or the alternative resource bound beat both of the Choi bounds over the entire range of values that was investigated. Since this was a numerical study, it is not possible to definitively say that the tightest out of the trivial bound and the alternative resource bound is always tighter than either of the Choi bounds, however this is the case for a wide range of parameter values.

In Figs. 5.1 and 5.2, we demonstrate the performance of the various bounds. Choosing $\eta_X > \eta_Y$, we decompose $\mathcal{C}_Y$ as the pointwise application of $\mathcal{C}_X$ and some other AD channel, $\mathcal{C}_{XY}$, with transmissivity

$$\eta_{XY} = \frac{\eta_Y}{\eta_X}. \tag{5.92}$$

Two specific values of $\eta_{XY}$ (one per plot) were chosen: $0.95$ and $0.9$. Two values of $N$, the total number of channel uses, were also chosen: $10$ and $30$. With these kept fixed, $\eta_X$ was then varied

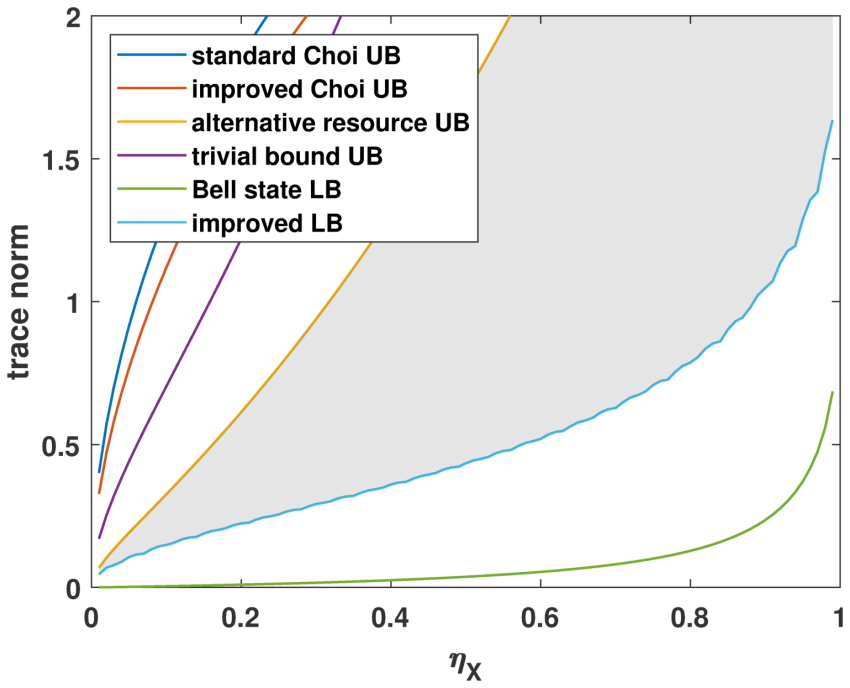


Figure 5.2: Upper and lower bounds on the maximum value of the trace norm between the two possible outputs of an adaptive discrimination protocol with no more than 30 channel uses. The channels being discriminated between are AD channels with transmissivities $\eta_X$ and $\eta_Y$, where $\eta_Y = \eta_X \eta_{XY}$. In this case, $\eta_{XY} = 0.9$. For these parameter values, the upper bound based on simulation using the alternative resource is always better than the other three upper bounds. It is to be expected that the trivial bound performs less well for high values of $N$, because it scales linearly with $N$, whilst the bounds based on PBT do not. The improved lower bound has a distinct advantage over the Bell state lower bound. The grey shaded area is the region between the tightest upper and lower bounds.

from $0.01$ to $0.99$ and the bounds were studied over this range. For the alternative resource bound, we have set $m = 150$.

As shown in the plots, the improved Choi bound performs better than the standard Choi bound, however both are beaten by either the trivial bound or the alternative resource bound (which of these is highest depends on the parameter values). The trivial bound performs better than the alternative resource bound when $\eta_X$ and $\eta_{XY}$ are large and when $N$ is small.

The new lower bound on the optimal trace norm (based on sending $N$ copies of the state $|1\rangle$ through the channel) is tighter than the lower bound from Ref. [2] (based on sending $N$ copies of a Bell state through the channel) across the entire range. It is clear, however, that there is still room to tighten either the upper or the lower bounds, since there is still a gap between the tightest upper

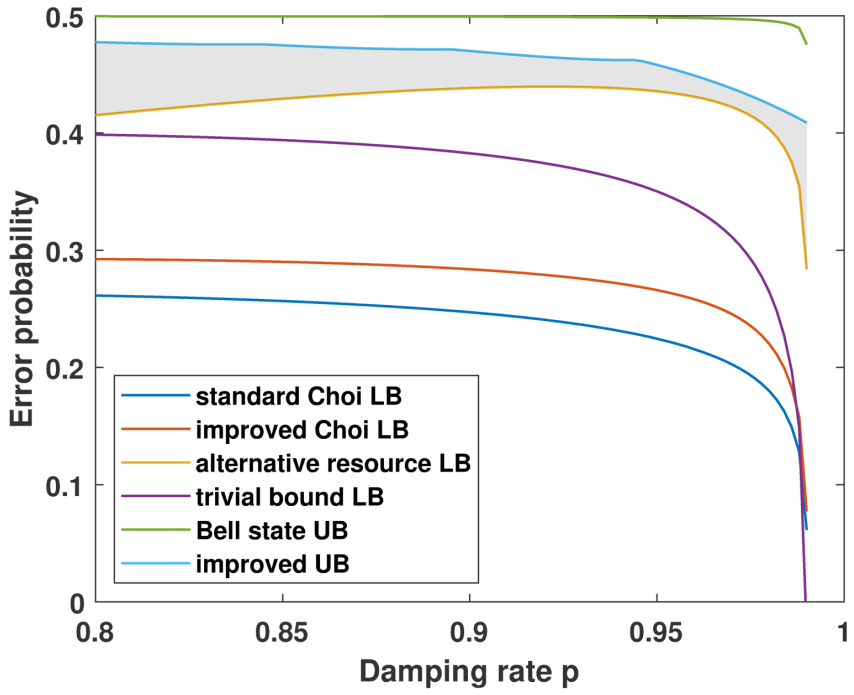


Figure 5.3: Comparison with the bounds on the error probability of discriminating between two AD channels, one with damping rate $p$ and one with damping rate $p + 0.01$, with equal prior probabilities, with no more than 20 channel uses, found in Ref. [2]. The line labelled "standard Choi" is the lower bound found in Ref. [2] and the line labelled "Bell state UB" is the upper bound from the same paper. The other lines are the new bounds presented here and the grey shaded area is the region between the tightest upper and lower bounds.

bound and the tightest lower bound, especially in Fig. 5.2.

In Fig. 5.3, we compare our new bounds to the bounds presented in Fig. 4 of Ref. [2]. We plot the error probability of discriminating between two AD channels, with equal prior probability, using the equation in Eq. (5.15). The AD channels are characterised by the damping rate, $p$, rather than by the transmittance, $\eta$, although the two quantities are trivially connected via the equation $p = 1 - \eta$. One channel has a damping rate of $p$ and the other has a damping rate of $p + 0.01$. The maximum number of channel uses is 20. The new lower bounds on the error probability (which come from the new upper bounds on the trace norm) are tighter than the lower bounds in Ref. [2] over the entire range; in this case, the alternative resource bound is the tightest lower bound. The new upper bound on the error probability (coming from the improved lower bound on the optimal trace norm) is slightly tighter than the upper bound in Ref. [2] over the entire range, but most noticeably for a high damping rate, $p$.

We now consider two examples of how these bounds might be applied to quantum information

tasks. The tasks we consider are quantum hacking and biological sensing (a quantum metrology task).

### 5.3.1 Applying the bounds to quantum hacking

Suppose a hacker, Eve, is attempting to eavesdrop on communications between a sender, Alice, and a receiver, Bob, who are implementing the BB84 protocol. Suppose also that Eve is able to send photons into Alice's device before each transmission (and to receive some return state). Eve could use this side-channel to gain more information on the states sent by Alice than is accounted for in the security proofs. For instance, Alice's basis choice could be enacted by a polariser [122]. By sending in photons with a known polarisation, Eve could glean information about Alice's basis choice based on the loss experienced by the photons (which could be basis dependent). Then, if Eve could determine Alice's basis with a high probability of success, she could carry out an intercept and resend attack on the photons sent through the main channel, without greatly disturbing them. In other words, she could measure the signal states in the basis that she believes them to have been sent in, based on her side-channel attack. Alice and Bob would only detect errors in half of the cases in which Eve incorrectly guesses Alice's basis. Since, in this scenario, Eve's error probability is low, the quantum bit error rate detected by the trusted parties would be much lower than the $25\%$ normally expected for an intercept and resend attack.

We model the attack as Eve carrying out a general, adaptive discrimination protocol with up to $N$ channel uses. We set the transmissivity of one of the channels as $\eta_Y = 0$ and then choose three different values of $\eta_X$: $10^{-5}$, $5 \times 10^{-6}$ and $10^{-6}$. We then calculate Eve's discrimination error probability, assuming equal prior probabilities of each channel occurring, for various numbers of channel uses. In this scenario, we assume that we have a perfect polariser, and so for one channel (i.e. for one polarisation), the photons sent through are completely absorbed by the polariser, whilst for the other channel, they are undisturbed by the polariser. We assume that the input states are so strongly attenuated that they can be modelled as a train of at most single photon states by the time they arrive at the polariser and hence that Eve's protocol can be modelled as a discrimination protocol between AD channels. This is a reasonable assumption, since BB84 involves the sending of single-photon states, which are often produced using strongly attenuated laser pulses. It is thus reasonable to assume that a laser pulse sent by Eve into Alice's device, through the optical fibre, would be similarly attenuated, such that the pulse arriving at the polariser could be well-modelled as a qubit state. We also assume that further attenuation occurs as the states leave the device, giving rise to the low values of $\eta_X$. This is in line with the architecture in Ref. [123], which limits

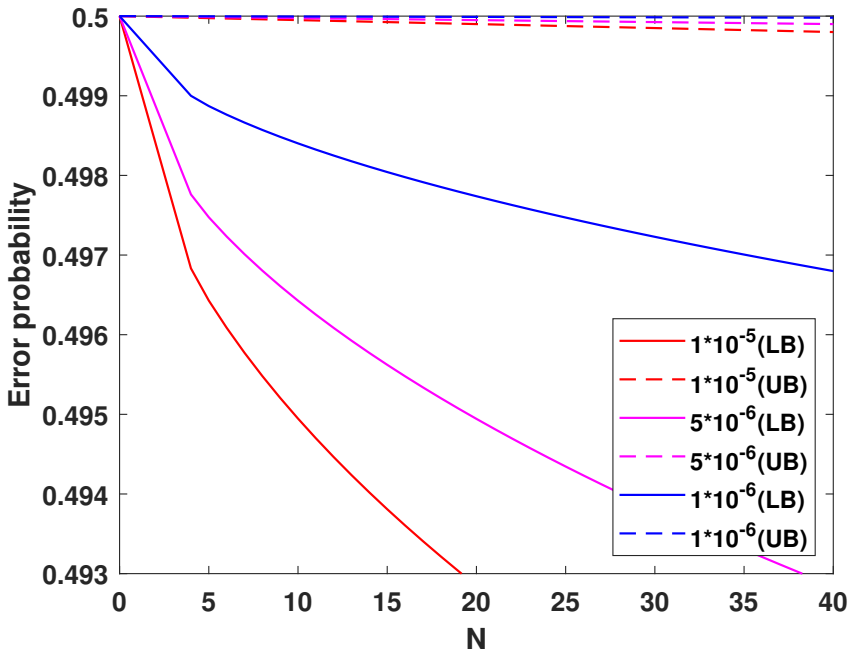


Figure 5.4: Upper and lower bounds on the discrimination error probability for an eavesdropper carrying out an adaptive protocol to discriminate between two BB84 preparation bases, with at most $N$ channel uses. We assume that Eve must send qubit states through an AD channel, in order to determine whether the channel has a transmissivity of $\eta_X$ or of $\eta_Y$. $\eta_Y = 0$, whilst $\eta_X$ takes values of $10^{-5}$, $5 \times 10^{-6}$ and $10^{-6}$; each case is represented by a different colour. The continuous lines give lower bounds on the error probability, whilst the dashed lines give upper bounds. The upper bounds are based on the improved lower bound, from Eq. (5.64). The lower bounds are based on the trivial bound, from Eq. (5.21), and the alternative resource bound, from Eq. (5.85); whichever bound has a higher value for a given $N$ is used for that value of $N$. For the alternative resource bound, $m = 150$. We find that, for all three values of $\eta_X$, the trivial bound gives a tighter bound for $N \leq 4$ and the alternative resource bound gives a tighter bound for $N > 4$.

the total mean photon number leaving Alice's device via the optical fibre, per signal sent through the main channel, to $10^{-6}$.

The assumption that Eve's states can be modelled as (up to) one-photon states probing AD channels can be justified by numerically finding the energy-constrained diamond norm [5, 65, 66] between a lossy channel and the pointwise application of a truncation channel (a channel mapping all number states of the form $|n > 1\rangle$ to $|0\rangle$) and the same lossy channel, for low transmissivities. More specifically, we use the semidefinite program for calculating the energy-constrained diamond norm given in Ref. [66]; note that the definition of the energy-constrained diamond norm used by Winter [66] (and Shirokov [65]) differs slightly from the definition given by Pirandola et al. [5].

We find that, for $\eta_X = 10^{-6}$, the truncation to one-photon states has a small effect on the error probability.[1]

Since one of the channels ($\mathcal{C}_Y$) will always output the state $|0\rangle$, we can significantly simplify the improved lower bound. Eq. (5.64) reduces to

$$D^{|1\rangle,N}_{\mathcal{C}_X\mathcal{C}_Y} = 2(1 - \eta_X^N). \tag{5.93}$$

The upper and lower bounds found are shown in Fig. 5.4. The upper bounds come from the improved lower bound and the lower bounds are based on whichever is tighter of the trivial bound and the alternative resource bound. For the chosen values of $\eta_X$, the trivial bound is tighter for $N \leq 4$. This is in line with our expectation that the trivial bound performs less well (compared to bounds based on PBT simulation) for large values of $N$, due to its linear scaling. The gap between the upper and lower bounds is small in proportion to their values, but still shows significant room for improvement, especially for large $N$. It is not clear whether it is the upper bounds, the lower bounds, or both which need tightening.

### 5.3.2 Applying the bounds to biological sensing

Quantum channel discrimination protocols have applications in biology. The concentration of bacteria in a growth medium affects the transmissivity of light through the medium. The tasks of distinguishing between the presence and absence of bacteria in a sample and of distinguishing between two possible concentrations of bacteria can therefore be considered to be quantum channel discrimination tasks, where the two possible channels are lossy channels with different transmissivities. Further, in biological applications, low photon numbers are often desirable, since intense radiation can harm the samples that are being probed. As a result, in some scenarios, modelling the task as an AD channel discrimination task may be appropriate.

In Ref. [3], Spedalieri et al. show that quantum light sources and detectors can reduce the error probability for both detecting the presence or absence of *E. coli* in a sample and determining whether a sample contains *E. coli* or *Salmonella*. They start by determining the transmissivities of growth media containing *E. coli* and *Salmonella* bacteria, as a function of time. The time-dependence comes from the changing concentrations of the bacteria in the media as they grow. The two possible types of bacteria and the case with no bacteria present therefore correspond to three different possible lossy channels. Determining whether a specific bacteria is present or absent

[1] See the supplementary data of Ref. [10] for more details.

and determining which of the two types of bacteria is present then become channel discrimination tasks.

Spedalieri et al. consider the task of parameter estimation, where the parameter to be estimated is the transmissivity of the channel. They consider both coherent state sources and the optimal input states for parameter estimation, from Ref. [120] (which are number states that send the maximum number of photons through the channel per channel use). They then bound the error probability for detecting the presence of *E. coli* and discriminating between *E. coli* and *Salmonella*, by using the expressions in Eqs. (5.69) and (5.70). In the symmetric testing case (equal prior probabilities), the mean of $p_X^{\text{err}}$ and $p_Y^{\text{err}}$ is minimised over $\tau$. Note, however, that the resulting expression (the QCRB bound) only provides an upper bound on the optimal error probability for sufficiently large $N$ (i.e. in the regime in which the QCRB is tight).

In Figs. 5.5 and 5.6, we plot upper and lower bounds on the optimal error probability for an adaptive protocol with up to $150$ channel uses, each sending at most one photon through the channel. This is reasonable, because it is desirable to send only a small amount of energy through the channels and because the error probabilities from Eqs. (5.69) and (5.70) can be achieved in this way.

Fig. 5.5 bounds the error probability over time for detecting the presence of *E. coli* in a sample. In this scenario, $\mathcal{C}_X$ is the channel corresponding to a blank sample (no bacteria present), and so $\eta_X$ has a constant value of $\eta_{\text{bk}} = 0.92$. $\mathcal{C}_Y$ is the channel corresponding to a sample with *E. coli* present and has a transmissivity of

$$\eta_Y = \eta_{E.Coli}(t) = \eta_{bk} - c_{1,E.Coli}t^2 + c_{2,E.Coli}t^3, \tag{5.94}$$

where $c_{1,E.Coli}$ and $c_{2,E.Coli}$ are constants (for a fixed type of bacteria) with values of $0.1\ \text{hrs}^{-2}$ and $0.0088\ \text{hrs}^{-3}$ respectively and where $t$ is the time, in hours, since the sample was prepared. The values of $c_{1,E.Coli}$ and $c_{2,E.Coli}$ were experimentally determined in Ref. [3], and the cubic expression for $\eta_{E.Coli}$, from Eq. (5.94) is valid for small $t$ ($\leq 3$).

The lower bound is the tightest out of the lower bounds derived from our upper bounds on the trace norm. In fact, this is always the bound based on the trivial bound (in the regime in which the lower bound is $> 0$). For the upper bound, we consider both the error probability derived from the exact form of the improved lower bound on the trace norm [from Eq. (5.64)] and the QCRB bound. Since the two bounds overlap almost perfectly, the approximation is valid in this regime ($N = 150$). It is clear that there is room for improvement of either the upper or the lower bounds on the trace norm for large $N$.

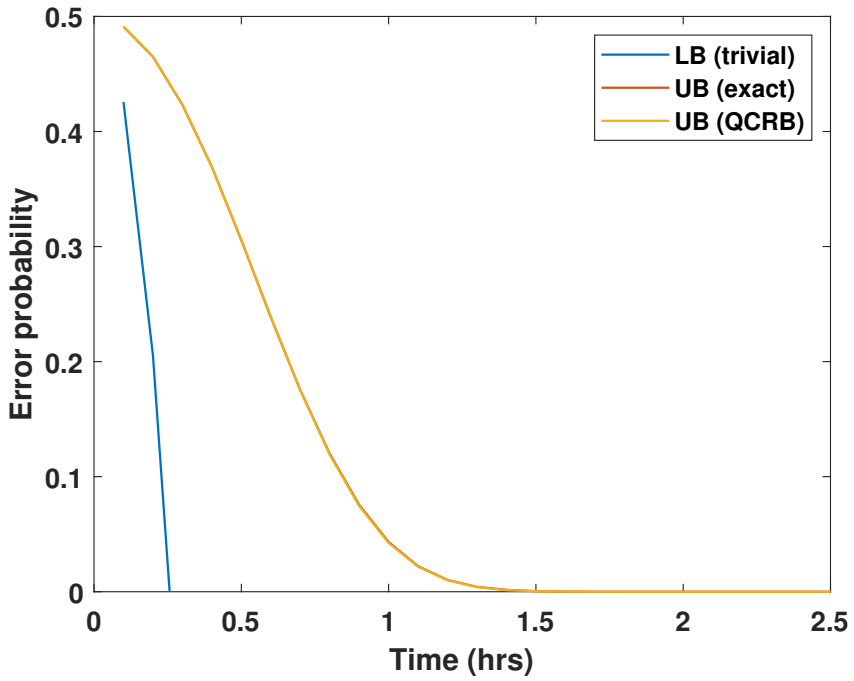


Figure 5.5: Upper and lower bounds on the error probability of detecting the presence of *E. Coli* bacteria in a sample, with a maximum of 150 channel uses (each using no more than one photon) as a function of time. The transmissivity of the blank sample is constant, whilst the transmissivity of the sample containing *E. Coli* is modelled as following a cubic equation (with respect to the time since the sample was prepared). The lower bound (denoted "LB (trivial)") is derived from the trivial bound on the trace norm. The exact form of the upper bound ("UB (exact)") is derived from the improved lower bound on the trace norm and the approximation to the upper bound ("UB (QCRB)") is based on the QCRB bound. Since the two bounds overlap almost perfectly, the approximation is valid in this regime.

Fig. 5.6 bounds the error probability over time for discriminating between samples of *E. coli* and *Salmonella*. In this scenario, $\mathcal{C}_X$ is the channel corresponding to a sample containing *E. coli* and $\mathcal{C}_Y$ is the channel corresponding to a sample containing *Salmonella*. In this case, we calculate the time-dependent transmissivities differently, by modelling the absorbances, $A$, of the samples as Gompertz functions and applying the formula

$$\eta = 10^{-A}. \tag{5.95}$$

The absorbances are modelled as following

$$A = c_1 e^{a} + A_{bk}, \tag{5.96}$$

$$a = -e^{\frac{c_2 e}{c_1}(c_3 - t)+1}, \tag{5.97}$$

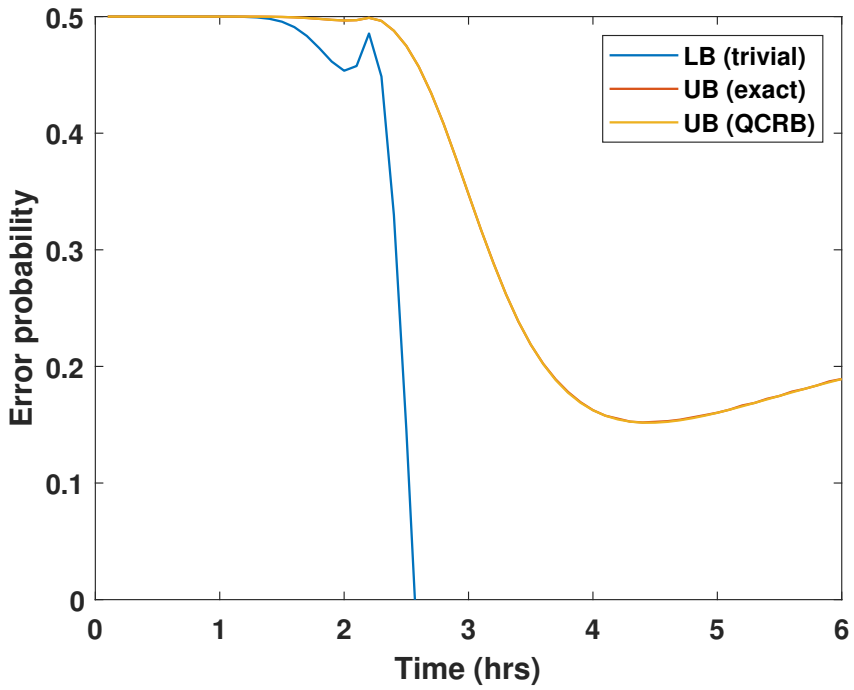


Figure 5.6: Upper and lower bounds on the error probability of discriminating between *E. Coli* and *Salmonella* bacteria in a sample, with a maximum of 150 channel uses (each using no more than one photon) as a function of time. The absorbances of the samples are modelled as following Gompertz functions. The lower bound (denoted "LB (trivial)") is derived from the trivial bound on the trace norm. The exact form of the upper bound ("UB (exact)") is derived from the improved lower bound on the trace norm and the approximation to the upper bound ("UB (QCRB)") is based on the QCRB bound. Since the two bounds overlap almost perfectly, the approximation is valid in this regime. The absorbances are initially very similar, but become more distinguishable as the time since the sample was prepared increases. We note that this plot differs from Fig. 10 in Ref. [3]; this is because Spedalieri et al. consider probing with a mean total of $10^3$ photons, whilst we only allow a maximum of 150 photons in total. They also model the transmissivities of the two samples using cubic equations, rather than Gompertz functions.

where $A_{\text{bk}}$ is the absorbance of a blank sample and $c_1$, $c_2$ and $c_3$ are experimentally determined coefficients that depend on the type of bacteria present in the sample. Spedalieri et al. found that the triple $(c_1, c_2, c_3)$ took values $(0.309, 0.139, 2.634)$ for *E. coli* and $(0.242, 0.0882, 2.672)$ for *Salmonella* [3]. $A_{bk}$ (which was the same for both samples) took the value $0.144$.

The lower bound is derived from the tightest of our upper bounds on the trace norm, which is again the trivial bound over the entire regime in which the lower bound is $> 0$. The upper bounds are calculated in the same way as for Fig. 5.5, and we again find that the exact form of the bound

and the approximation overlap almost perfectly. The bounds briefly peak after a little more than 2 hours, before decreasing again, due to the fact that the difference in the absorbances of the samples briefly decreases before increasing again. Once again, we have a large gap between the bounds, which could be improved by tightening either the lower or the upper bounds. It is not yet known which bound most needs to be tightened.

## 5.4 Summary

In this chapter, we calculated multiple new bounds on the optimal trace norm for discriminating between two AD channels. We have strengthened both the upper and the lower bounds on the optimal trace norm by presenting the improved Choi bounds, the alternative resource bounds, the trivial bound and the improved lower bound on the trace norm. We have also calculated the exact diamond norm between AD channels, thus obtaining the exact error probability for one-shot channel discrimination between any two AD channels, in analytical form.

The bounds were then numerically investigated and we found that the either the alternative resource bound or the trivial bound gave the tightest lower bound over a wide range of parameter ($\eta_X$, $\eta_Y$, and $N$) values. The bounds were applied to two different scenarios: quantum hacking of BB84 and biological quantum metrology (detecting and discriminating between bacteria in a sample). In the latter scenario, we also confirm that the QCRB bound is valid as an approximation of the discrimination error probability derived from the improved lower bound on the trace norm (and is therefore a valid upper bound on the error probability) for large $N$ (in our case, $N = 150$).

We briefly discussed how these results could be extended to pure loss qudit channels, however this is an area that is open to more research, which could find bounds on the error probability of adaptive discrimination protocols between any two lossy channels. Another area for continued research is the further tightening of either the upper or the lower bounds on the optimal trace norm, since there is still room for improvement.

This work contributes to the theory of channel simulation of AD channels and significantly improves the bounds on the optimal error probabilities for adaptive discrimination protocols between AD channels.

# Chapter 6

# Trojan horse attacks on coherent state protocols

The work in this chapter forms the basis of a paper published in Physical Review A, whose authors are (in order) Jason Pereira and Stefano Pirandola [8].

We start this chapter by discussing quantum hacking in general and the Trojan horse attack specifically. In the next section, we introduce our side channel model and calculate the key rate of a coherent state protocol when the sender's device is subject to a Trojan horse attack. We briefly discuss how the attack could be mitigated and then we summarise our work.

## 6.1 Introduction

Quantum information science [109, 111, 124] is advancing at a rapid pace. The progress of quantum computing [15] threatens to make current, classical cryptography insecure. Quantum key distribution (QKD) [125–127] is a possible solution to this problem, offering provable information security based on physical principles. It is possible to design QKD protocols that ensure that any eavesdropper can hold only an arbitrarily small amount of information about the message sent. This holds true regardless of how advanced the eavesdropper's technology is.

Security proofs for QKD protocols have a few assumptions that must hold in order for them to be valid [128]. The two trusted parties (Alice and Bob) must have isolated devices, which are inaccessible to the eavesdropper (Eve). The devices should be fully characterised, so that an adversary cannot exploit device imperfections to acquire information about the key or to alter the trusted parties' estimations of the quantum channel properties. The trusted parties must also have an authenticated (but not secure) classical channel; an eavesdropper can listen in to classical

communications along this channel, but cannot alter them. If we relax any of these conditions, the secure key rate for a protocol may change.

Current commercial implementations of discrete variable (DV) protocols, such as BB84 [51] with decoy states [129, 130], have been shown to be vulnerable to a variety of attacks that exploit device imperfections, such as "side-channels" that leak information from the trusted parties' devices to Eve [46]. These attacks include detector blinding attacks [131], time-shift attacks [132] and Trojan horse attacks [133].

A variety of attacks on continuous variable (CV) protocols have been proposed. In experimental realisations of QKD, the local oscillator (which is used by Bob to carry out his measurements) is often sent down the quantum channel; this introduces a vulnerability that an eavesdropper can exploit. Häseler et al. [134] demonstrated that Eve could disguise an intercept and resend attack by replacing the signal state and the local oscillator with squeezed states. The wavelength-dependence of beamsplitters in Bob's setup can be exploited to engineer his measurement outcomes [135,136]. Altering the shape of the local oscillator pulse can allow an eavesdropper to change Bob's estimation of the vacuum noise [137, 138]. Saturation attacks [139, 140], which push Bob's detectors out of the linear mode of operation, have also been proposed.

One way of avoiding attacks that exploit device imperfections is to use device-independent QKD [128, 141]. This is a family of protocols that do not require Alice's and Bob's devices to be trusted. Such protocols are immune from many side-channel attacks, but have significantly lower key rates than protocols that require trusted devices. Measurement-device independent (MDI) QKD protocols have been formulated for both the DV [142, 143] and the CV [144] cases and have much higher key rates than fully device-independent protocols. MDI-QKD removes threats from the detector's point of view, but still assumes that the state-preparation devices are completely trusted. Therefore, MDI-QKD is also subject to the quantum hacking described in this chapter.

In this chapter we consider a Trojan horse attack, where Eve sends extra photons into Alice's device, in order to gain information about the states being sent through the main quantum channel without disturbing the signal state. This type of attack was first considered in depth by Vakhitov et al. [145]. Such an attack may be used in DV protocols [146], in order to distinguish decoy states from signal states or to gain information about Alice's basis choice.

Gisin et al. [147] described how reflectometry could be used by Eve to gain information about Alice's phase modulator settings and analytically calculated the information leakage in terms of the photon number of the state received by Eve after the side-channel. They assumed attenuation of the side-channel mode by Alice and showed that the information leakage is reduced if Alice can

randomise the phase of the side-channel mode.

Lucamarini et al. [123] calculated the secret key rate for BB84, with and without decoy states, in the presence of a Trojan horse side-channel, in terms of the photon number of the state received by Eve. They then bounded the incoming photon number in terms of the Laser Induced Damage Threshold (LIDT) of the optical fibre and the time for which Alice's device gate is open, assuming that the Trojan horse photons are sent in via the main channel, whilst the gate is open. Based on this constraint, they designed an architecture to passively limit the photon number of the received state and hence the information leakage.

Tamaki et al. [148] found general analytical expressions for the information leakage of DV protocols due to Trojan horse attacks, in terms of the actions of the phase and intensity modulators. This allows the secret key rate of a general DV protocol in the presence of a Trojan horse side-channel to be calculated, as long as the phase and intensity modulators are well-characterised.

In Chapter 5, we describe how a discrimination protocol between two amplitude damping channels could model an attempt by an eavesdropper to gain information on the basis choice for the BB84 protocol [51], via a Trojan horse attack. We gave bounds on the discrimination error probability, which hold for the most general adaptive protocols. We thereby found ultimate bounds on the probability of an eavesdropper successfully discovering the basis choice, although our lower bounds on the error probability may not be achievable (since the bounds are not tight). In this chapter, we look in detail at a specific type of Trojan horse attack on a CV-QKD protocol.

Here we assume a CV protocol based on the modulation of coherent states [7], so that the attack is against the modulator. The experimental viability of carrying out a Trojan horse attack on the commercial CV system SeQureNet has previously been considered [149]. We assume that Eve is both hacking Alice's device with $\bar{n}$ mean photons per run and tapping the main quantum channel between Alice and Bob, which can be assumed to be a thermal-loss channel.

## 6.2 Calculating the key rate with a side channel in the sender's device

### 6.2.1 General scenario

We consider two parties, Alice and Bob, who are trying to establish a secret key, with a third party, Eve, trying to gain information about the secret key. Alice initiates a coherent state protocol [7,55]. This involves her displacing a vacuum state by a Gaussian-distributed random (two-dimensional) variable, $\alpha$. In real implementations, this displacement is generally carried out by independently modulating the phase and the intensity, so that the overall displacement has a Gaussian distribution.

She then sends the displaced vacuum state (called the signal state) to Bob, via a quantum channel. Bob then carries out a heterodyne measurement on the signal state, to obtain a value $\beta$. This process is repeated several times. Alice and Bob compare some of their values via a classical communication channel in order to establish the transmittance, $\eta$, and excess noise, $\epsilon$, of the channel. Bob and Alice then establish a secret key based on their shared knowledge of Bob's values (this is called reverse reconciliation).

Whilst the signal states are in the main quantum channel, we allow Eve to enact any unitary operation upon them. We assume that Eve can listen in on all classical communication between Alice and Bob (but cannot alter it). She can then store all states involved in the operation (except for the signal state) in a quantum memory and carry out an optimal measurement on them after all quantum and classical communication has been completed, in order to gain information about Bob's values. Alice and Bob therefore assume that all of the noise and loss of the channel has been caused by Eve's unitary operations and try to bound the maximum knowledge that Eve could have obtained about Bob's values. As long as Alice has more information about Bob's values than Eve, it is possible for Alice and Bob to obtain a secret key.

If Eve is only able to access the main channel and is not able to access Alice or Bob's devices in any way, the optimal attack on the signal state for a given attenuation and noise is an entangling cloner [150]. The secret key rate for this case has been calculated [151]. Here we instead consider the case where Eve also has access to part of Alice's device via a side-channel. Eve can send a Trojan horse mode into Alice's device, which will be displaced by $\alpha$ in the same way as the signal state. This side-channel mode contains an average number of photons $\bar{n}$, and we assume that Alice is able to monitor these photons and estimate their number. This will not be the case for most current CV-QKD implementations, especially since certain potential Trojan horse side-channels may not have been identified yet, so additional quantum metrological tools must be placed inside Alice's box in order for this assumption to be met. To represent Eve's Trojan horse mode, we assume it is part of a two-mode squeezed vacuum (TMSV) state [55] with squeezing $r$, so that $\bar{n} = \sinh^2 r$. This is an active attack when $\bar{n} > 0$ and it is a passive one when $\bar{n} = 0$, meaning that we just have a leakage mode from Alice's device.

Recently, a side-channel on CV-QKD based on leakage from a multimode modulator was considered by Derkach et al. [152], building on their previous work [153]. These works considered leakage modes prior to and after modulation of the signal state, for both the coherent state and the squeezed state protocols. However, these authors did not consider side-channels that allow Eve to send photons into Alice's device (non-zero values of $\bar{n}$). They also considered homodyne, rather

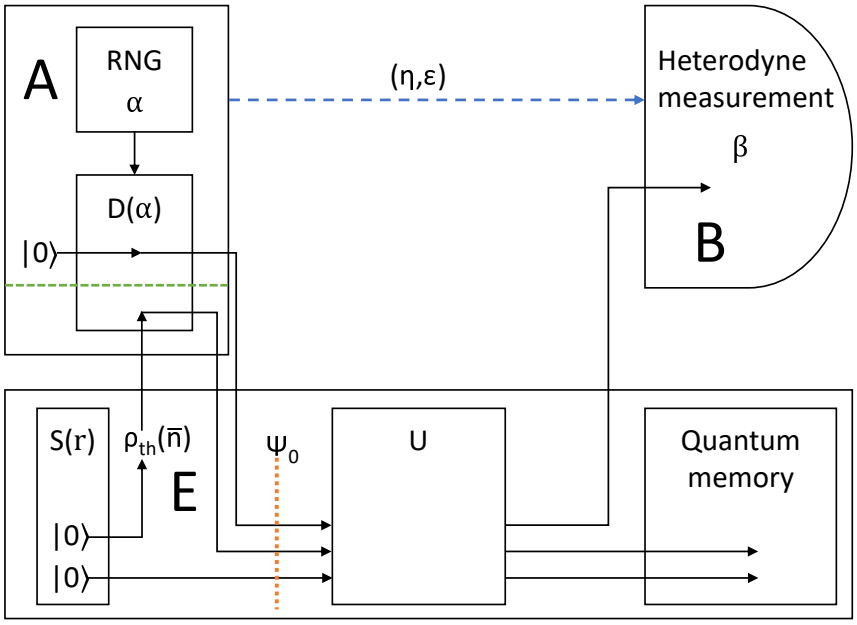


Figure 6.1: The channel setup under consideration. A is Alice's device, B is Bob's device and E is Eve's device. The dashed green line marks the part of Alice's device that is accessible to Eve. Eve sends one mode of a TMSV state into Alice's device to be displaced by $\alpha$ in the same way as the signal state. Alice knows the average photon number, $\bar{n}$, of Eve's state. The (displaced) squeezed vacuum modes and the signal state form the state $\psi_0$. Eve enacts a unitary on this total state and any ancillary modes, then sends the signal state to Bob and stores the remaining modes in a quantum memory. Bob carries out a heterodyne measurement on the signal state, obtaining $\beta$. We find the key rate assuming that the main channel is a thermal channel, with transmittance $\eta$ and excess noise $\epsilon$, as represented by the blue dashed arrow.

than heterodyne, measurements by the receiver. In this chapter, we will consider a more general scenario, where the hacking of Alice's device is active, therefore involving the use of two-mode squeezing, so that $\bar{n} > 0$ photons enter the device. We analyse the security when the side-channel mode is modulated by $\alpha$, exactly as the signal mode is (we later generalise to the case where its modulation is $m\alpha$). See Fig. 6.1 for an overview of the situation.

To find the secret key rate in reverse reconciliation, we need to calculate the mutual information between Alice and Bob $I(\alpha : \beta)$ and that between Eve and Bob. The latter is upper-bounded by the Holevo bound $I(E : \beta)$, which can be calculated as the reduction in entropy of Eve's output state when conditioned by Bob's value, $\beta$. We upper-bound Eve's knowledge of Bob's state by assuming that all noise and loss experienced by the signal state is due to Eve enacting unitary operations on the signal state and some ancillary modes, which are then stored in a quantum

memory.

Note that we set the vacuum noise equal to 1 in this chapter.

### 6.2.2 Reduction of the attack

If there are no side-channels, Eve's Holevo bound can be calculated by assuming that the signal state is entangled with some state held by Alice and that $\alpha$ is the result of a heterodyne measurement on a TMSV state [150]. In the presence of our side-channel, the initial state held by Eve prior to her enacting the main channel is tripartite and composed of the signal mode and Eve's side-channel modes. Our first step must be to determine the first and second moments of this state $\psi_0$ (see Fig. 6.1). We label the initial first moment vector $X_0$ and the initial second moment (covariance) matrix $V_0$. For a fixed value of $\alpha$, we have the conditional state $\psi_0|\alpha$ which is the tensor product of a coherent state $|\alpha\rangle\langle\alpha|$ and a TMSV state where one of the modes has also been displaced by $\alpha$. The conditional moments are given by

$$X_0|\alpha = \begin{pmatrix} \alpha \\ \alpha \\ 0 \end{pmatrix}, \quad V_0|\alpha = \begin{pmatrix} \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cosh 2r\mathbb{I} & \sinh 2r\mathbb{Z} \\ \mathbf{0} & \sinh 2r\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}, \tag{6.1}$$

where $\mathbb{I}$ is the one-mode identity matrix, $\mathbf{0}$ is the one-mode zero-matrix, and $\mathbb{Z}$ is the Pauli Z-matrix.

In order to find the elements of $V_0$, we add the expectation value of $X_0|\alpha \cdot X_0|\alpha^T$ to $V_0|\alpha$. Using $\langle\alpha\rangle = 0$ and $\langle\alpha^2\rangle = \mu$, we find

$$X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; V_0 = \begin{pmatrix} (\mu+1)\mathbb{I} & \mu\mathbb{I} & \mathbf{0} \\ \mu\mathbb{I} & (\mu+\cosh 2r)\mathbb{I} & \sinh 2r\mathbb{Z} \\ \mathbf{0} & \sinh 2r\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}. \tag{6.2}$$

From the covariance matrix $V_0$ we can compute the three symplectic eigenvalues [55]

$$v_1 = 1, \tag{6.3}$$

$$v_2 = \mu + \sqrt{1+\mu+\mu^2+\mu\cosh 2r}, \tag{6.4}$$

$$v_3 = -\mu + \sqrt{1+\mu+\mu^2+\mu\cosh 2r}, \tag{6.5}$$

and compute the entropy of the total state as [55] $S(\psi_0) = \sum_{k=1}^{3} g(v_k)$ where [154]

$$g(x) = \frac{x+1}{2}\log_2\frac{x+1}{2} - \frac{x-1}{2}\log_2\frac{x-1}{2} \tag{6.6}$$

$$\overset{x\gg 1}{\to} \log_2\frac{ex}{2} + O(x^{-1}). \tag{6.7}$$

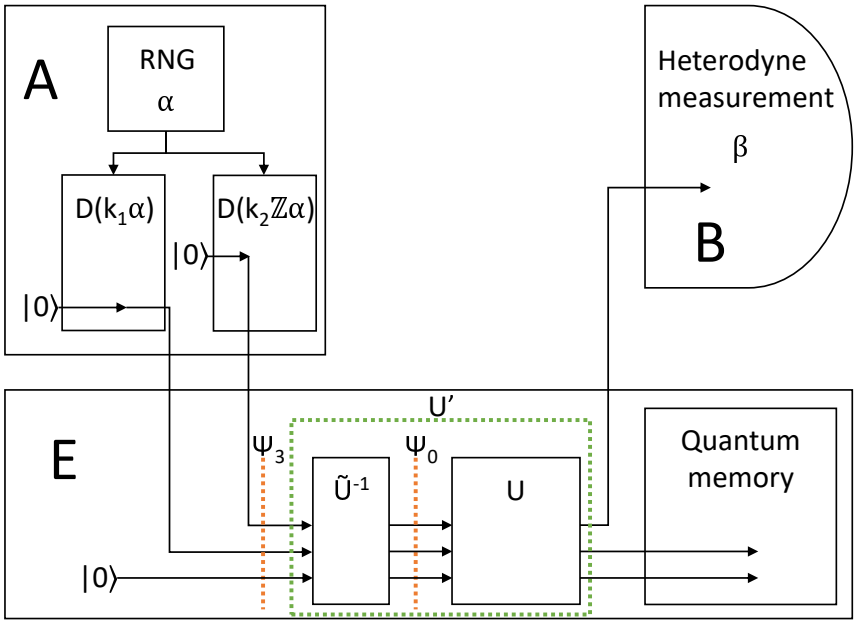


Figure 6.2: An equivalent channel to the setup in Fig. 6.1. Alice draws a two-dimensional variable, $\alpha$, from a Gaussian distribution then displaces one vacuum state by $k_1\alpha$ and another by $k_2\mathbb{Z}\alpha$. The first mode is sent through the main channel to Bob as the signal state and the second mode is leaked to Eve. The equivalence can be seen from the fact that Eve can get the initial state from Fig. 6.1, $\psi_0$, by enacting the unitary $\tilde{U}^{-1}$ and can then enact the same arbitrary unitary, $U$. We can regard this as Eve enacting a single combined unitary, $U'$.

The fact that $v_1 = 1$ tells us that there is a symplectic transformation that reduces $\psi_0$ to a tensor product of a two-mode state and a vacuum state. We can build on this observation and reduce the number of modes. In fact, we may show the reduction to the setup in Fig. 6.2, which only involves the signal mode, modulated by $k_1\alpha$ (with $k_1 > 1$), and a single Trojan horse mode, modulated by $k_2\mathbb{Z}\alpha$ (with $k_2$ real). We can design a Gaussian unitary $\tilde{U}$ that converts the initial state $\psi_0$ from Fig. 6.1 into the initial state $\psi_3$ from Fig. 6.2. This unitary operation $\tilde{U}$ is the optical circuit shown in Fig. 6.3, where we have labelled the signal state as $\psi_B$, Eve's squeezed state that enters the side-channel as $\psi_{E_1}$ and Eve's idler state (the squeezed state that does not enter the side-channel) as $\psi_{E_2}$.

To see how the circuit transforms the state, we examine it after each of the three optical components; we label the states after each component with the subscripts 1, 2 and 3. $\psi_i$ has first moments vector $X_i$ and covariance matrix $V_i$. The conditional state $\psi_i|\alpha$ is associated to $X_i|\alpha$ and $V_i|\alpha$. The symplectic matrix of the i$^{\text{th}}$ component is $S_i$ and it characterises the transformation of the state from $\psi_{i-1}$ to $\psi_i$ as follows: $V_i = S_i V_{i-1} S_i^T$ and $X_i = S_i X_{i-1}$.

The first component is a balanced beamsplitter, acting on the signal state and Eve's side-channel mode. This sets the quadratures for Eve's side-channel mode to 0. It has symplectic matrix

$$S_1 = \begin{pmatrix} \frac{1}{\sqrt{2}}\mathbb{I} & \frac{1}{\sqrt{2}}\mathbb{I} & \mathbf{0} \\ -\frac{1}{\sqrt{2}}\mathbb{I} & \frac{1}{\sqrt{2}}\mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I} \end{pmatrix}, \tag{6.8}$$

and it results in the following moments for $\psi_1|\alpha$ and $\psi_1$

$$X_1|\alpha = \begin{pmatrix} \sqrt{2}\alpha \\ 0 \\ 0 \end{pmatrix}, \tag{6.9}$$

$$V_1|\alpha = \begin{pmatrix} \cosh^2 r\mathbb{I} & \sinh^2 r\mathbb{I} & \frac{\sinh 2r}{\sqrt{2}}\mathbb{Z} \\ \sinh^2 r\mathbb{I} & \cosh^2 r\mathbb{I} & \frac{\sinh 2r}{\sqrt{2}}\mathbb{Z} \\ \frac{\sinh 2r}{\sqrt{2}}\mathbb{Z} & \frac{\sinh 2r}{\sqrt{2}}\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}, \tag{6.10}$$

$$V_1 = V_1|\alpha \oplus 2\mu \begin{pmatrix} \mathbb{I} & & \\ & \mathbf{0} & \\ & & \mathbf{0} \end{pmatrix}. \tag{6.11}$$

The second component is a two-mode squeezer, operating on Eve's modes such that one of them becomes a vacuum state. Its squeezing parameter is given by $r_2 = \log\left(\frac{\sqrt{2}\cosh r - \sinh r}{\sqrt{\cosh^2 r + 1}}\right)$, and it has symplectic matrix

$$S_2 = \begin{pmatrix} \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\sqrt{2}\cosh r}{\sqrt{\cosh^2 r+1}}\mathbb{I} & -\frac{\sinh r}{\sqrt{\cosh^2 r+1}}\mathbb{Z} \\ \mathbf{0} & -\frac{\sinh r}{\sqrt{\cosh^2 r+1}}\mathbb{Z} & \frac{\sqrt{2}\cosh r}{\sqrt{\cosh^2 r+1}}\mathbb{I} \end{pmatrix}. \tag{6.12}$$

The moments of $\psi_2|\alpha$ and $\psi_2$ are given by

$$X_2|\alpha = \begin{pmatrix} \sqrt{2}\alpha \\ 0 \\ 0 \end{pmatrix}, \tag{6.13}$$

$$V_2|\alpha = \begin{pmatrix} \cosh^2 r\mathbb{I} & \mathbf{0} & \sqrt{\cosh^4 r - 1}\mathbb{Z} \\ \mathbf{0} & \mathbb{I} & \mathbf{0} \\ \sqrt{\cosh^4 r - 1}\mathbb{Z} & \mathbf{0} & \cosh^2 r\mathbb{I} \end{pmatrix}, \tag{6.14}$$

$$V_2 = V_2|\alpha \oplus 2\mu \begin{pmatrix} \mathbb{I} & & \\ & \mathbf{0} & \\ & & \mathbf{0} \end{pmatrix}. \tag{6.15}$$

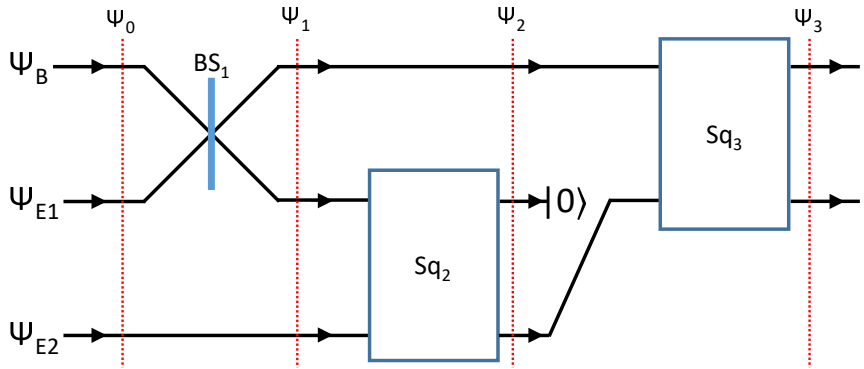


Figure 6.3: A circuit that converts the initial (pre-main channel) state from the setup in Fig. 6.1 into the initial state from the setup in Fig. 6.2. This shows that the two channel setups have the same key rate, since Eve can enact any unitary operation and hence is able to convert one into the other. We label this entire circuit $\tilde{U}$. Eve can also enact the inverse, $\tilde{U}^{-1}$. $\psi_B$ denotes the signal state, $\psi_{E_1}$ denotes Eve's squeezed state that enters the side-channel and $\psi_{E_2}$ denotes Eve's idler state. $\mathrm{BS}_1$ is a balanced beamsplitter and $\mathrm{Sq}_2$ and $\mathrm{Sq}_3$ are two-mode squeezers. $\mathrm{BS}_1$ moves all of the displacement onto the first mode, such that Eve's states are no longer displaced, $\mathrm{Sq}_2$ unsqueezes Eve's states such that one of the modes becomes a pure vacuum state and $\mathrm{Sq}_3$ unsqueezes the signal state and Eve's remaining mode such that they become pure displaced vacuum states.

Note that one of the modes has become a vacuum state. Henceforth, we neglect this mode and implicitly enact the identity operation on it. We now see that, for fixed $\alpha$, the system is a displaced TMSV state. The third component undoes the squeezing, leaving us with two displaced vacuum states. Its squeezing parameter is given by $r_3 = -\mathrm{arcsinh}\left(\frac{\sinh r}{\sqrt{2}}\right)$ and it has symplectic matrix

$$S_3 = \begin{pmatrix} \frac{\sqrt{\cosh^2 r+1}}{\sqrt{2}}\mathbb{I} & -\frac{\sinh r}{\sqrt{2}}\mathbb{Z} \\ -\frac{\sinh r}{\sqrt{2}}\mathbb{Z} & \frac{\sqrt{\cosh^2 r+1}}{\sqrt{2}}\mathbb{I} \end{pmatrix}. \tag{6.16}$$

The moments of $\psi_3|\alpha$ and $\psi_3$ are

$$X_3|\alpha = \begin{pmatrix} k_1\alpha \\ k_2\mathbb{Z}\alpha \end{pmatrix}, \; V_3|\alpha = \begin{pmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{pmatrix}, \tag{6.17}$$

$$V_3 = \begin{pmatrix} (1+k_1^2\mu)\mathbb{I} & k_1k_2\mu\mathbb{Z} \\ k_1k_2\mu\mathbb{Z} & (1+k_2^2\mu)\mathbb{I} \end{pmatrix}, \tag{6.18}$$

where we have set

$$k_1 := \sqrt{\cosh^2 r + 1}, \; k_2 := -\sinh r. \tag{6.19}$$

This concludes the proof of equivalence between the setups in Fig. 6.1 and Fig. 6.2.

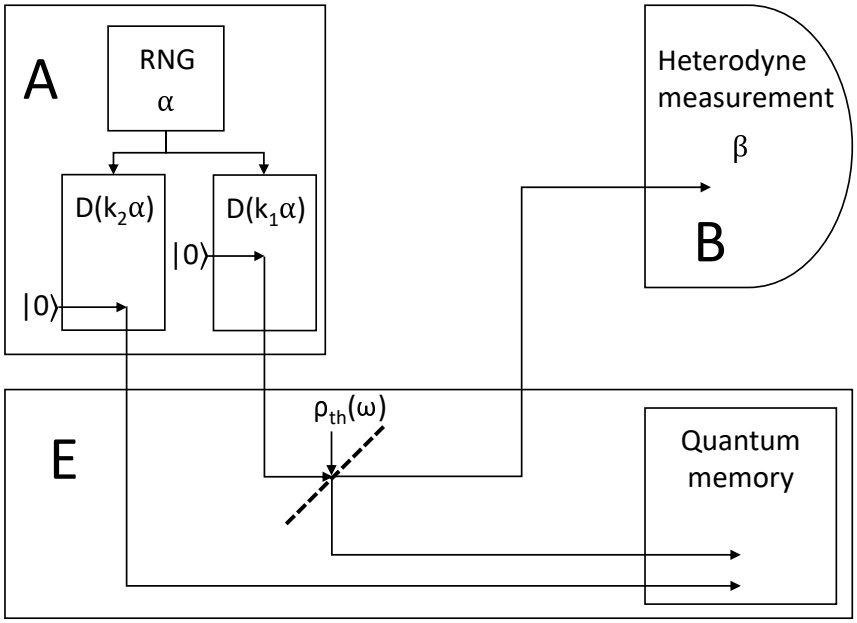


Figure 6.4: An alternative channel setup that must give the same secret key rate as the setup in Fig. 6.2 assuming the presence of a thermal-loss channel. The difference between the two setups is that in Fig. 6.2, the x-quadrature of Eve's side-channel state is modulated by $k_2\alpha_x$ and the p-quadrature is modulated by $-k_2\alpha_p$; in this figure,the x-quadrature is still modulated by $k_2\alpha_x$ but the p-quadrature is modulated by $k_2\alpha_p$. Since the two quadratures encode independent variables and since the x-quadrature is not affected by the change, the mutual informations arising from the measurement of the x-quadrature, $I^x_{AB}$ and $I^x_{EB}$, must be the same in each setup and hence the key rates must be the same. We assume that Eve beamsplits the signal state with some thermal state with variance $\omega$. This specific representation of Eve's unitary is unique up to isometries on her output ancillas. In other words, if we fix the channel to be thermal-loss, then its dilation into a beams-splitter with an environmental thermal state is fixed up to unitaries acting over Eve's entire output Hilbert space [4].

We note that the two components (quadratures) of $\alpha$ are uncorrelated with each other and have the same variance. Let us also assume that the two quadratures of Bob's outcome ($\beta$) are also uncorrelated with each other and have the same variance. This is certainly the case in the presence of a thermal-loss channel, characterised by a transmittance $\eta$ and an excess noise $\epsilon$, which is the most typical scenario in QKD. Next, we show that the setup in Fig. 6.2 has the same key rate as the setup in Fig. 6.4, in which the signal mode is modulated by $k_1\alpha$ and the side-channel mode is modulated by $k_2\alpha$ (rather than by $k_2\mathbb{Z}\alpha$). Note that in Fig. 6.4, we have also imposed that the general unitary results in a thermal-loss channel.

Since we assume that the main channel does not mix the quadratures, we can treat the two quadratures of $\alpha$, which we denote as $\alpha_x$ and $\alpha_p$, as independent variables that have been sent through the channel and measured to give the independent variables $\beta_x$ and $\beta_p$ respectively. Let $I^x_{AB}$ ($I^p_{AB}$) denote the mutual information between Alice and Bob arising from the measurement of the x-quadrature (p-quadrature) and let $I^x_{EB}$ ($I^p_{EB}$) denote the maximum mutual information between Eve and Bob arising from the measurement of the x-quadrature (p-quadrature). Since the x and p quadratures of $\alpha$ and $\beta$ are independent and identically distributed, $I_{AB}$ and $I_{EB}$ are double $I^x_{AB}$ and $I^x_{EB}$ respectively.

Let $I'_{AB}$, $I'^x_{AB}$, $I'_{EB}$ and $I'^x_{EB}$ be the counterparts of $I_{AB}$, $I^x_{AB}$, $I_{EB}$ and $I^x_{EB}$ respectively for the setup in Fig. 6.4. It is again true that $I'_{AB}$ and $I'_{EB}$ are double $I'^x_{AB}$ and $I'^x_{EB}$ respectively. Further, since the quadratures are independent and the x-quadratures of Eve's states are not affected by the change in setup (the only difference is that the p-quadrature of Eve's side-channel mode is modulated by $k_2\alpha_p$ rather than by $-k_2\alpha_p$), $I^x_{AB}$ must be the same as $I'^x_{AB}$. This means that $I_{AB}$ is the same as $I'_{AB}$ and $I_{EB}$ is the same as $I'_{EB}$. Note that this holds for all channels (not just thermal channels) that do not mix the quadratures and so the $\mathbb{Z}$ matrix in Fig. 6.2 can be neglected for any such channel.

Hence, the setup in Fig. 6.4 must give the same key rate as the setup in Fig. 6.2 and therefore the setup in Fig. 6.1. The setup in Fig. 6.4 is equivalent to a main channel setup with a higher initial modulation and a lower effective transmittance. The equivalent main channel attack is shown in Fig. 6.5. The signal state is modulated by $k\alpha$, where

$$k = \sqrt{k_1^2 + k_2^2} = \sqrt{2}\cosh r = \sqrt{2(\bar{n}+1)}, \tag{6.20}$$

and hence the modulation amplitude is $k^2\mu$. $k$ is a function of $\bar{n}$, which characterises the side-channel. We note that $k_1$ and $k_2$ are functions only of $\bar{n}$. By choosing an appropriate parameter for the beamsplitter in Fig. 6.5, Eve can get the initial state of Fig. 6.4. We then effect a thermal channel by beamsplitting with the thermal state with parameter $\omega$. We can reduce both operations to a single beamsplitter operation with some other thermal state $\omega'$ (see Fig. 6.6).

This allows us to calculate the key rate in the same way as a main channel attack but with a higher "effective modulation amplitude", $\mu'$, and a lower "effective transmittance", $\eta'$. These effective parameters (the channel parameters that the trusted parties would calculate for the setup in Fig. 6.5) are related to the measured values of $\mu$ and $\eta$ by

$$\mu' = k^2\mu, \;\; \eta' = \frac{\eta}{k^2}. \tag{6.21}$$

The effective transmittance accounts for both beamsplitters and is the transmittance that we would

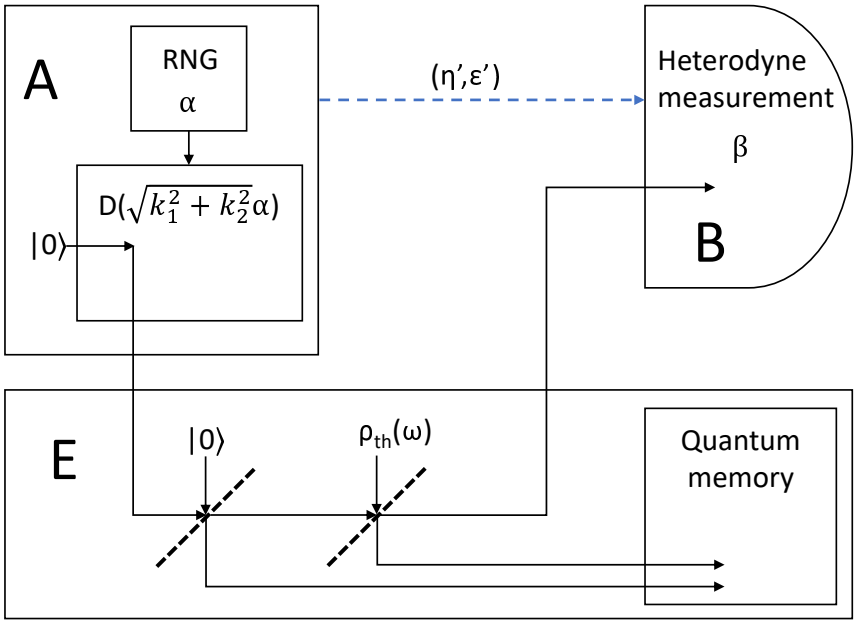


Figure 6.5: This is a setup without a side-channel that must give the same secret key rate as the setup with the side-channel. The variance of Alice's variable in this setup is higher than the actual variance of $\alpha$, and the channel transmittance for this setup is lower than the observed channel transmittance, $\eta$. The channel for this setup can be regarded as a thermal channel with parameters $\eta'$ and $\epsilon'$ (represented by the blue, dashed arrow).

observe if, instead of a setup with a signal state modulated by $\mu$ and a side-channel (as seen in Fig. 6.1), we had a setup with a signal state modulated by $\mu'$ and no side-channel, with the same measured values of $\beta$ (as seen in Fig. 6.5). This was found by multiplying the transmissions of the two beamsplitters in Fig. 6.5.

It is helpful to clarify the definition of the excess noise, $\epsilon$. To do so, we introduce the random variable $n$: this is the total relative input noise of $\beta$ around $\alpha$, including the vacuum noise. We can describe $\beta$ in terms of $n$ as $\beta = \sqrt{\eta}(\alpha + n)$. Here $n$ is characterised by its second moment $\langle n^2 \rangle = 1 + (1-\eta)/\eta + \epsilon$. We now find the effective excess noise, $\epsilon'$ (as would be observed for the setup in Fig. 6.5), using the fact that we have the same measured $\beta$ values in all representations. $\beta$ can be expressed in terms of effective parameters as $\beta = \sqrt{\eta'}(k\alpha + n')$, where the second moment of $n'$ is now given by $\langle n'^2 \rangle = 1 + (1-\eta')/\eta' + \epsilon'$. We then substitute in the definition of $\eta'$, compare the expressions for $\beta$, and solve for $\epsilon'$, i.e.

$$\epsilon' = \frac{\eta}{\eta'}\epsilon = k^2\epsilon. \tag{6.22}$$

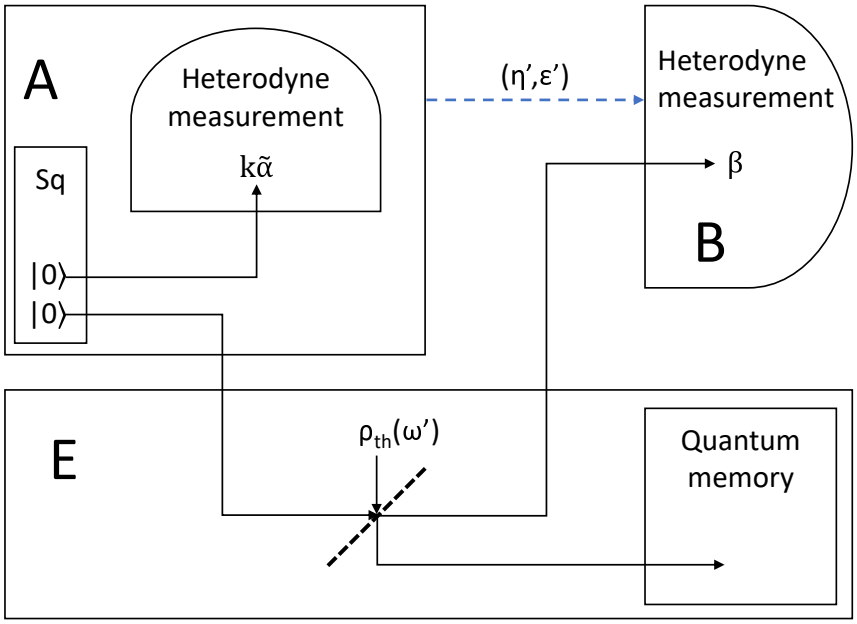


Figure 6.6: This is the entanglement-based representation of the attack in Fig. 6.5. Alice heterodynes one half of a TMSV state to get the value $k\tilde{\alpha}$, which linearly corresponds to $k\alpha$ (the displacement of the signal state). The signal state enters the channel and is subject to some thermal noise due to beamsplitting with one mode of an entangling cloner (the thermal state $\omega'$). It is then heterodyned by Bob, to obtain $\beta$. The resultant state of Alice, Bob and Eve is pure. The channel between Alice and Bob is a thermal channel, characterised by $\eta'$ and $\epsilon'$; this is represented by the blue, dashed arrow.

### 6.2.3 Computation of the key rate

To calculate the secret key rate for a main channel attack with a modulation amplitude of $\mu'$, a transmittance of $\eta'$ and an excess noise of $\epsilon'$, we can use an entanglement-based representation (rather than a prepare and measure representation) [150]. This representation is shown in Fig. 6.6 and is valid as long as $\mu > 0$.

Alice heterodynes one mode of a TMSV state, obtaining the value $k\tilde{\alpha}$ (and hence also the value of $\alpha$) and preparing the state $\rho(k\alpha)$. She then sends the prepared signal state through the channel to Bob, who heterodynes it to obtain $\beta$. In the channel, the signal state is beamsplit with the thermal state $\rho_{th}(\omega)$. The total state shared by Alice, Bob and Eve, which we denote $\rho_{ABE}$, is pure since Eve holds the purification of the channel. This means that the entropy of Eve's state, $\rho_E$, is equal to the entropy of the combined state of Alice and Bob, $\rho_{AB}$. The combined state of Alice and Eve conditioned by some value of $\beta$, $\rho_{AE}|\beta$, is also pure, so the entropy of Eve's state conditioned by $\beta$, $\rho_E|\beta$, is equal to the entropy of Alice's state conditioned by $\beta$, $\rho_A|\beta$.

The covariance matrix of $\rho_{AB}$ is

$$V_{AB} = \begin{pmatrix} (\mu'+1)\mathbb{I} & \sqrt{\eta'\mu'(\mu'+2)}\mathbb{Z} \\ \sqrt{\eta'\mu'(\mu'+2)}\mathbb{Z} & (\eta'(\mu'+\epsilon')+1)\mathbb{I} \end{pmatrix}, \tag{6.23}$$

the covariance matrices of the conditional states $\rho_A|\beta$ and $\rho_B|\alpha$ are given by

$$V_A|\beta = \left(\mu'+1-\frac{\eta'\mu'(\mu'+2)}{\eta'(\mu'+\epsilon')+2}\right)\mathbf{1}, \tag{6.24}$$

$$V_B|\alpha = (\eta'\epsilon'+1)\mathbb{I}. \tag{6.25}$$

We can calculate the symplectic eigenvalues of $V_{AB}$ using the formula in Ref. [55]. The expressions for these eigenvalues can be simplified by taking the asymptotic limit in $\mu$ (the limit as $\mu \to \infty$). In this limit, $\mu' \to \infty$ and all other parameters stay the same. We assume that $\eta' \leq 1$, since realistically, Eve will not enact a main channel that causes gain rather than loss. We denote the two symplectic eigenvalues of $V_{AB}$ in this limit as $v^{\infty}_{AB,1}$ and $v^{\infty}_{AB,2}$ and denote the symplectic eigenvalue of $V_A|\beta$ in this limit as $v^{\infty}_{A|\beta}$. We find these to be:

$$v^{\infty}_{AB,1} = 1 + \frac{\epsilon'\eta'}{1-\eta'}, \tag{6.26}$$

$$v^{\infty}_{AB,2} = \mu'(1-\eta'), \tag{6.27}$$

$$v^{\infty}_{A|\beta} = \frac{2}{\eta'} + \epsilon' - 1. \tag{6.28}$$

We calculate the mutual information between Alice and Bob, $I(\alpha : \beta)$, as the reduction in (classical) entropy of $\beta$ when conditioned with $\alpha$. The asymptotic limit of this mutual information is equal to

$$I(\alpha:\beta)^{\infty} = H(V_\beta + 1) - H(V_\beta|\alpha + 1) \tag{6.29}$$

$$= \log_2 \frac{\eta'\mu'}{\eta'\epsilon'+2}, \tag{6.30}$$

where $H$ is the Shannon entropy [13] and $V_\beta$ ($V_\beta|\alpha$) is the variance of Bob's outcome $\beta$ (conditional outcome $\beta|\alpha$). We then calculate the Holevo bound between Eve and Bob in the asymptotic limit. We find:

$$I(E:\beta)^{\infty} = \log_2 \frac{e\, v^{\infty}_{AB,2}}{2} + S_{\text{const}}, \tag{6.31}$$

where

$$S_{\text{const}} = g(v^{\infty}_{AB,1}) - g(v^{\infty}_{A|\beta}) \tag{6.32}$$

is the entropy contribution that does not scale with $\mu$. The first term of this expression comes from the asymptotic form of $g(v^{\infty}_{AB,2})$, as per Eq. (6.7).

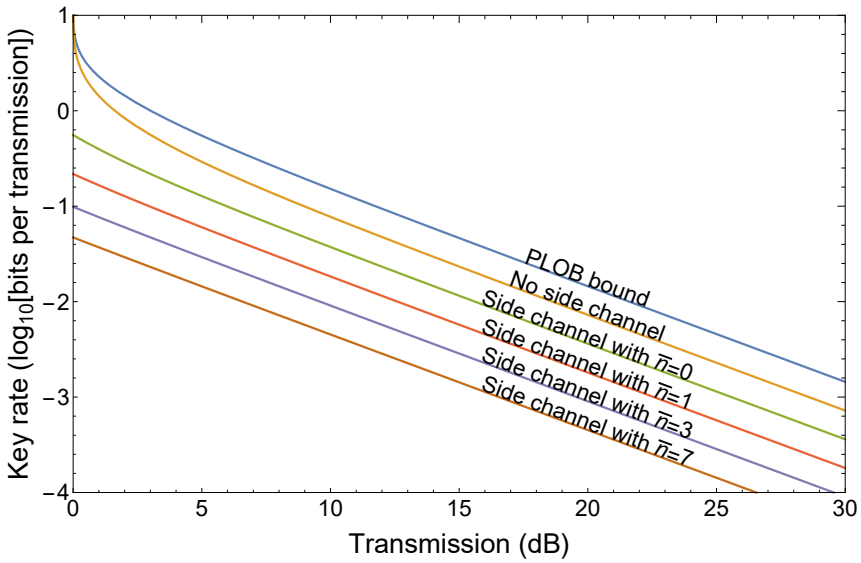


Figure 6.7: Plots of the secret key rate (in logarithmic scale) versus channel transmission $\eta$ of the main quantum channel, in the absence of excess noise (lossy channel rate). The top curve is the PLOB bound [5], which is the secret key capacity of the lossy channel, i.e. the maximum key rate achievable over this channel by any point-to-point QKD protocol in the absence of side-channels [6]. We then show the ideal rate of the coherent state protocol [7] with no side channels. Lower curves refer to the coherent state protocol in the presence of a side-channel with an increasing number of photons $\bar{n}$, ranging from the leakage mode case ($\bar{n} = 0$) to more active hacking ($\bar{n} = 1$, 3, 7). As we can see, the key rate is always positive (for any value of $\bar{n}$), but it quickly declines as $\bar{n}$ increases.

The asymptotic secret key rate is given by the difference

$$K^{\infty}(\bar{n}, \eta, \epsilon) = I(\alpha : \beta)^{\infty} - I(E : \beta)^{\infty} \tag{6.33}$$

$$= \log_2 \frac{2\eta'}{e(1-\eta')(\eta'\epsilon' + 2)} - S_{\text{const}}. \tag{6.34}$$

The extra information gained by Eve due to the side-channel is the difference between the key rate with the side-channel and the key rate without. In general, the asymptotic key rate decreases as the effective transmission decreases (either due to an increase in the average photon number of the side-channel mode or due to increased line loss) and as the channel noise increases. This is shown in the plots in Figs. 6.7 and 6.8.

The asymptotic secret key rate $K^{\infty}$ takes a particularly simple form if the channel does not

add any noise (a pure-loss channel). In fact, it becomes

$$K_{\text{lossy}}^{\infty} = -\frac{\log_2(1-\eta')}{\eta'} - \log_2 e \tag{6.35}$$

$$= \frac{2(\bar{n}+1)}{\eta}\log_2\left[1-\frac{\eta}{2(\bar{n}+1)}\right] - \log_2 e. \tag{6.36}$$

The rate $K_{\text{lossy}}^{\infty}$ is always positive and plotted in Fig. 6.7 for various mean photon numbers $\bar{n}$, where it is also compared with the ultimate point-to-point rate or PLOB bound $-\log_2(1-\eta)$ [5]. Each time $\bar{n}+1$ doubles (e.g. when $\bar{n}$ goes from 0 to 1, from 1 to 3, or from 3 to 7), the key rate $K_{\text{lossy}}^{\infty}$ decreases by approximately 3 dB.

In the low transmission regime (i.e. long distances), it is known that the PLOB bound becomes roughly linear in $\eta$ and is approximately equal to $\eta/\ln 2 \simeq 1.44\eta$ bits per transmission. It is also known that, without side-channels, the coherent state protocol has a long-distance ideal rate of about $\eta/(2\ln 2) \simeq 0.72\eta$ bits per transmission, which is half the PLOB bound. The linearity also holds when we include the side channels. In fact, for low $\eta$, we find that the key rate of Eq. (6.36) becomes

$$K_{\text{lossy}}^{\infty} \simeq \frac{\eta}{4(\bar{n}+1)\ln 2} \simeq \frac{0.36}{\bar{n}+1}\eta\,. \tag{6.37}$$

Note that with the leakage mode ($\bar{n} = 0$), this rate is half that of the coherent state protocol without side-channels. This rate keeps halving each time $(\bar{n}+1)$ doubles; this can also be seen in the constant decrease in intercept between each of the plots in Fig. 6.7.

We then calculate the threshold excess noise, $\epsilon_{\text{max}}$, for a given channel transmission, $\eta$, and side-channel parameter, $\bar{n}$. This is the value of the excess noise up to which secret key distribution is possible. The threshold condition $\epsilon_{\text{max}} = \epsilon(\eta, \bar{n})$ is given by solving $K^{\infty}(k, \eta, \epsilon) = 0$. In Fig. 6.8, we show the security threshold of the coherent state protocol [7] without side-channels and, then, in two cases with side-channel modes ($\bar{n} = 0$ and 1). The shaded regions show the regions in which secret key distribution is possible for a given side-channel.

The leakage mode case ($\bar{n} = 0$) has a significantly lower security threshold than the case with no side-channel, and increasing the average photon number further decreases the threshold, for fixed transmission. For instance, for channel transmission of 20 dB, the presence of leakage ($\bar{n} = 0$) decreases the tolerable excess noise by $\simeq 0.06$ (from about $0.12$). For active hacking with $\bar{n} = 1$ photon, we have a further decrease of $\simeq 0.03$. In other words, a side-channel with $\bar{n} = 1$ gives a $\simeq 75\%$ decrease in tolerable excess noise at this distance. If $\bar{n}$ is increased, the attack becomes even more powerful. It is then important for Alice to be able to accurately measure $\bar{n}$, by characterising her devices as accurately as possible.

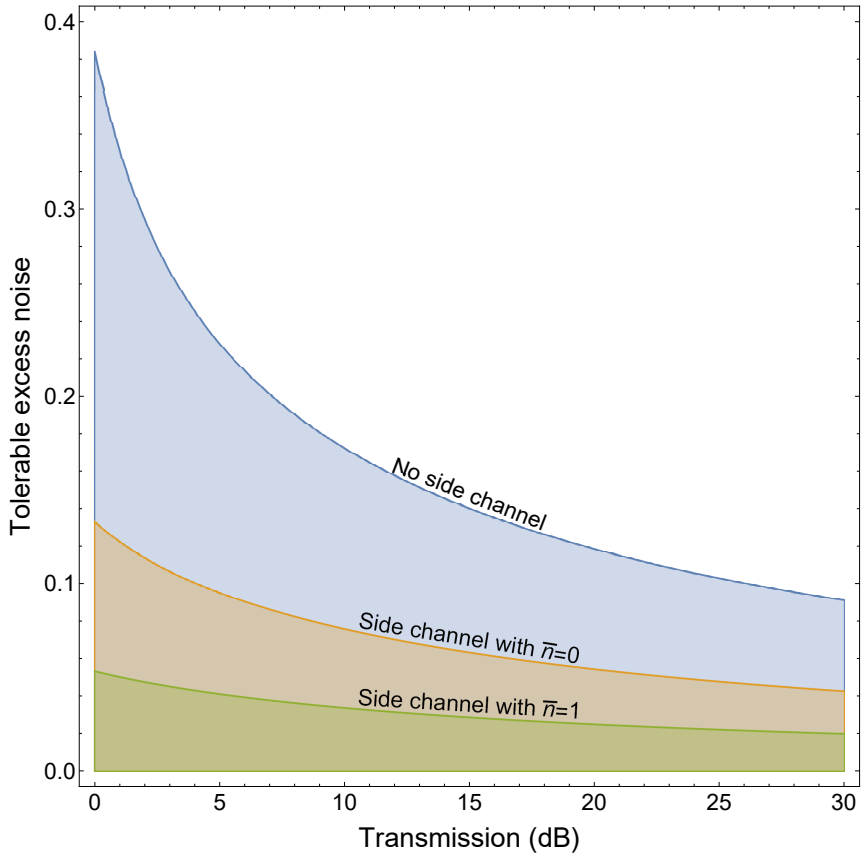


Figure 6.8: Security thresholds in terms of maximally-tolerable excess noise versus channel transmission (in decibels). The shaded regions are the regions in which secret key distribution is possible for a given side-channel. The boundaries of the regions show the values of the excess noise at which secret key distribution becomes impossible for a given transmission and side-channel. Adding the leakage mode side-channel significantly decreases the tolerable excess noise for a given transmission, and increasing the average photon number $\bar{n}$ of the side-channel further decreases it.

### 6.2.4 Generalisation of the side channel

We can also consider a simple extension, in which Eve's side-channel mode is modulated by $m\alpha$, whilst Alice's signal state is modulated by $\alpha$. $m$ is a multiplicative factor on the displacement of the Trojan state; $m = 1$ gives the case that has already been considered. This setup is shown in Fig. 6.9. Without loss of generality, we assume that $m > 0$, since Eve can always apply a phase shift of $\pi$ to her modes. Similarly to the original $m = 1$ case, we can show that this attack is equivalent to a standard attack against the main channel but with an "effective modulation amplitude", an "effective excess noise" and an "effective loss". The original and effective parameters are related by the same Eqs. (6.21) and (6.22), but where $k$ becomes the following function of both

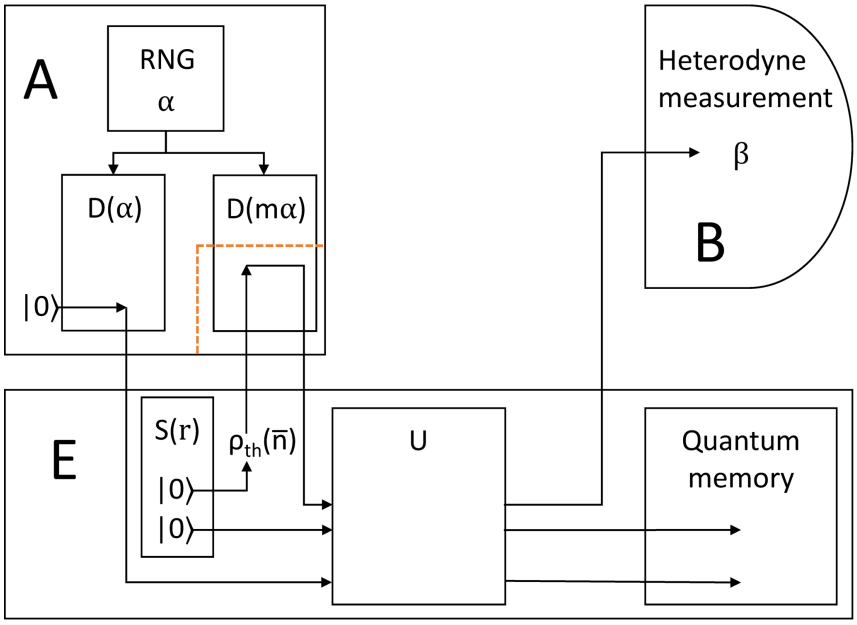


Figure 6.9: This is an extension of the original setup (Fig. 6.1), in which both the average number of photons entering Alice's device, $\bar{n}$, and the modulation amplitude of the side-channel mode, $m$, are monitored. Unlike in the original case, $m$ does not have to equal 1 and can take any real value. The dashed red line marks the part of Alice's device that is accessible to Eve. The key rate for this setup can be calculated similarly to the key rate for the original setup; the only difference is in the expression for the $k$ parameter, which affects the "effective loss", the "effective excess noise" and "effective modulation amplitude". See text for more explanation.

$\bar{n}$ and $m$[1]

$$k(\bar{n}, m) = \sqrt{m^2(2\bar{n} + 1) + 1}. \tag{6.38}$$

By monitoring both $\bar{n}$ and $m$, Alice can therefore fully quantify the effect of any single mode side-channel of this type. Alice can find $\bar{n}$ by monitoring the average photon number entering her device. There are a number of ways in which she could find $m$. For instance, she could monitor the total average outgoing photon number of her device across all modes.

[1]Note that if $m = 1$, this reduces to the previous case. Note also that if $m = 0$, we do not have a side-channel and so $k = 1$, hence the "effective loss" is equal to the observed loss, as we would expect. See Appendix C for the details of the derivation.

## 6.3 Mitigating the effects of the side channel

If the modulator can be surrounded by a passive attenuator, such that any photons not entering or leaving via the main channel are highly attenuated, the effects of any Trojan horse photons not entering via the main channel can be greatly mitigated. The attenuator can be modelled as a beamsplitting operation with a vacuum mode. The effects on the information gained by Eve are twofold: the first quadrature of her side-channel mode is scaled down by the attenuation and the correlations between Eve's Trojan horse mode and her idler mode are reduced. The conditional state $\psi_{E_S,E_I}|\alpha$ received by Eve after the side-channel (conditioned on Alice's value), will be an attenuated TMSV state. It has the covariance matrix

$$V_{E_S,E_I}|\alpha = \begin{pmatrix} (2\bar{n}+1)\mathbb{I} & 2\sqrt{\bar{n}'(\bar{n}+1)}\mathbb{Z} \\ 2\sqrt{\bar{n}'(\bar{n}+1)}\mathbb{Z} & (2\bar{n}'+1)\mathbb{I} \end{pmatrix}, \tag{6.39}$$

where $\bar{n}$ is the initial photon number of the TMSV state (prior to the side channel), $\bar{n}'$ is some positive real number less than $\bar{n}$, $E_S$ is the signal mode, and $E_I$ is the idler mode.

Let $\psi_{E_S E_I P}|\alpha$ be the purification of $\psi_{E_S,E_S}|\alpha$. Eve does not hold the purifying mode $P$, but if she did, it could only help her. Therefore, let us assume she is given it; this is equivalent to saying she is given the other output of the beamsplitter with its quadratures set to 0. Then the modes $E_I P$ together purify $E_S$. Any one-mode thermal state can be purified by a TMSV [55]. Hence, there exists some unitary acting only on the purifying systems $E_I$ and $P$ that results in a TMSV on the modes $E_S E_I$ with $\bar{n}'$ photons per mode (and a vacuum state on $P$).

If the Trojan horse mode is modulated by $m\alpha$, we can say that the first quadrature of this mode after the attenuator is $m'\alpha$, where $m'$ is some positive, real number less than $m$. Then, the key rate is lower-bounded by the key rate calculated before, but with $\bar{n}$ and $m$ replaced by $\bar{n}'$ and $m'$ respectively. More specifically, the expression for $k$ in Eq. (6.38) becomes

$$k(\bar{n}', m') = \sqrt{m'^2(2\bar{n}'+1)+1} \tag{6.40}$$

$$= \sqrt{Tm^2(2T^2\bar{n}'+1)+1}, \tag{6.41}$$

where $T$ is the transmission of the attenuator. This expression rapidly approaches unity as $T$ decreases. Here we have assumed that the Trojan horse state passes through the attenuator twice: once prior to modulation and once after modulation. Hence, we have set $m' = \sqrt{T}m$ and $\bar{n}' = T^2\bar{n}$.

The expression for the maximum secret key rate in this case is a lower bound: giving Eve access to the purification mode $P$ cannot decrease the Holevo bound on her mutual information,

but it is not immediately obvious whether it increases it. It is therefore not obvious whether this lower bound is tight or whether the power of the side-channel would be even further reduced by the attenuation; this is a question that is open to further study.

Without upper-bounding the incoming photon number, the addition of an attenuator does not provide provable security by itself, since we do not know the initial values of $m$ and $\bar{n}$, hence quantum metrological tools are still required. It may be possible to find an upper-bound on the incoming photon number for a given device, using physical considerations, such as the point at which damage to the modulator from the incoming photons would become obvious to Alice.

In order to limit the effects of a Trojan horse mode introduced via the main channel, the passive architecture to limit Trojan horse attacks in the DV case, introduced in Ref. [123], could be implemented in the CV case. The incoming photon number is bounded using the LIDT of the optical fibre constituting the main quantum channel; the photon number threshold is dependent on the frequency of the incoming photons, since lower frequency photons are less energetic, however the frequency is bounded from below by an optical fibre loop and a filtering block, which select for frequencies. There is then an attenuator, which greatly reduces the photon number of any incoming state. In this case, the attenuation will not decrease the magnitude of modulation of the Trojan horse state compared to the signal state, as it does in the case in which the Trojan horse photons do not enter via the main channel. This is because the signal state will be attenuated in the same way as the Trojan horse state. The incoming photons sent by Eve would still be attenuated, leading to damping of the off-diagonal elements of $V_{E_S,E_I}|\alpha$. By suitably choosing the attenuation, bearing in mind the maximum photon number of Eve's Trojan horse state, Alice could decrease the correlations between Eve's Trojan horse mode and her idler mode to an arbitrary degree and hence could effectively reduce Eve's side-channel to a leakage mode ($\bar{n'} = 0$).

Bounding the incoming photon number using the LIDT raises another issue, since we have assumed that modulation of the signal mode is unbounded and hence can be taken to infinity. Since the photon number of the signal state will also be limited by the LIDT, this is not entirely true. However, if the LIDT is sufficiently high, this should not greatly affect the secret key rate. Increasing the LIDT does not increase the photon number of Eve's outgoing side-channel mode as long as the attenuation is raised accordingly.

One further possible problem could occur if the attenuator itself is not properly characterised. If it re-radiates absorbed photons or scatters light in such a way that it is accessible to Eve, the attenuator itself may provide a leakage side-channel. Alternatively, it may have a much higher transmission at certain frequencies, allowing Eve to send Trojan horse photons through without

much attenuation.

## 6.4 Summary

In this chapter we have considered the effects of hacking Alice's box in one-way CV QKD, namely the coherent state protocol of Ref. [7], which is hacked while being implemented over a thermal-loss quantum communication channel. We have assumed that a Trojan horse side-channel mode is introduced in Alice's device and is modulated in the same way as the signal state. Under this condition, we have found out how quickly the key rate of the original protocol is deteriorated by increasing the mean number of photons $\bar{n}$ inserted in the device. Even the presence of a leakage mode ($\bar{n} = 0$) is able to halve the rate. Then, each time the value of $(\bar{n} + 1)$ doubles, the long-distance key rate is further halved.

Then we have also considered a direct generalisation of the basic side-channel attack in which the Trojan horse mode is modulated at a different amplitude ($m\alpha$) to the signal state. If this modulation is inefficient ($m < 1$), then the attack is weaker than the basic one. However, if $m > 1$, then the attack becomes more deleterious. In order to deal with this situation, Alice should be able to estimate not only the mean number of extra photons $\bar{n}$ entering the device, but also the mean number of extra photons leaving the device, so that she can also evaluate $m$. Therefore, it seems that quantum metrological tools [22, 30, 48, 62, 63, 155, 156] are necessary inside Alice's box, unless Eve's hacking is mitigated by other means which suitably modify the original setup and protocol.

# Chapter 7

# Conclusions

In this chapter, we will summarise the presented work and then provide possible directions that future research could take.

## 7.1 Summary of the presented work

In Chapter 1, we introduced the task of channel discrimination. We stated that the aim of this work was to improve the power of the techniques of quantum channel simulation and protocol stretching as tools for bounding the discrimination error and thereby to tighten existing bounds on the performance of the most general channel discrimination protocols.

Chapter 2 provided an overview of important quantities, formalisms, and techniques for quantum channel discrimination.

In Chapter 3, we investigated the performance of idler-free protocols for channel position finding (CPF) over a set of pure loss channels. By doing so, we showed that quantum protocol that use non-classical states can outperform classical protocols, even in the limited-technology scenario in which we are not able to store an idler. This is a useful result, because it shows that there is an advantage to developing quantum technologies for channel discrimination tasks, even if the development of a quantum memory proves difficult.

We then applied the technique of teleportation stretching to find tight bounds on the optimal output fidelity of a CPF protocol over a set of phase-insensitive Gaussian channels with fixed transmissivity. Since the lower bound on the output fidelity that we found is achievable, we calculated the exact output fidelity for the optimal (in terms of output fidelity) protocol. This also showed that the optimal protocol does not require adaptivity (this is in line with the finding by Pirandola and Lupo that the optimal protocol for channel discrimination is non-adaptive for such

channels [30]). We thereby bounded the error probability of the optimal CPF protocol. We were able to demonstrate quantum advantage over a range of parameters. Applications to a number of physical scenarios were considered, demonstrating the importance of CPF in both thermal imaging and quantum communications (in the form of eavesdropper localisation and finding the optimal - in terms of lowest induced noise - quantum communication channel).

Chapter 4 accomplishes the task of improving the power of quantum channel simulation by completely characterising qubit port-based teleportation (PBT), thereby allowing it to be used as a universal simulator of qubit channels. The analytical expression for the Choi matrix of the qubit PBT channel using the most general resource state possible is given and so is the channel from a resource state to the output Choi matrix. We then applied our analytical expressions to the task of channel simulation of the amplitude damping (AD) channel and were able to improve on existing channel simulations based on using copies of the channel's Choi matrix (i.e. achieve a lower simulation error).

Chapter 5 tightens existing lower bounds on the error probability of the most general discrimination protocols between AD channels, using the improved simulations from Chapter 4. We also tighten the existing upper bounds and show that, for a large number of channel uses, a discrimination bound based on the quantum Cramér-Rao bound, found in Ref. [3], can approximate the new upper bound. We present the diamond norm between any two AD channels and thus give the exact discrimination error probability in the one-shot case. The bounds are then applied to a variety of physical scenarios.

Chapter 6 takes a detailed look at a specific quantum hacking scenario, in which an eavesdropper has a side-channel into the sender's device and can send in Trojan states. By reducing the attack to an equivalent side-channel free setup, we compute the key rate for the scenario and show that a side channel can greatly reduce the key rate of a protocol. We therefore suggest the use of active monitoring to characterise any side-channels and passive architecture to mitigate the effect of any attack.

## 7.2 Directions for future work

CPF is a task that has not yet been well-studied. As a result, there are a number of open questions relating to it. In Chapter 3, we saw that idler-free protocols can achieve a quantum advantage when carrying out CPF on a set of pure loss channels. It may be worth investigating whether there are other protocol designs that can outperform the classical protocol - or even the bipartite entangled

protocol - whilst remaining technologically limited in some way.

PBT is a powerful tool for the simulation of qubit channels. However, it can also be used to simulate qudit channels. Its characterisation for higher dimensional input states is an open area for research and could potentially allow the construction of tighter discrimination bounds for many classes of qudit channels.

In Chapter 5, we significantly tightened the existing bounds on AD channel discrimination. There is still a gap between the upper and lower bounds (except for in the one-shot case), however, and so there is room for improvement of one or the other. Potentially, this could be accomplished using an even better simulation of the AD channel. The diamond norm for multiple copies of the AD channel (used in parallel) would give the ultimate bound on the discriminative power of a non-adaptive protocol and may be analytically calculable.

Finally, it may be possible to simulate the generalised AD channel using PBT and therefore to construct bounds on discrimination protocols between different generalised AD channels.

# Appendix A

# Appendices for Chapter 3

## A.1 Behaviour of the classical fidelity function

We now prove the statement in Subsection 3.3.3 of Chapter 3 that $F^{\text{class}}_{\text{loss/amp}} \to F^{\infty}_{\text{loss/amp}}$ as $\tau \to 0$. Substituting $\tau = 0$ into Eq. (3.80), we get

$$F^{\text{class}}_{\text{loss/amp}}(0, \epsilon_T, \epsilon_B) = \frac{\sqrt{\gamma_0 + \delta_0} + \sqrt{\gamma_0 - \delta_0}}{\delta_0}, \tag{A.1.1}$$

$$\gamma_0 = 4\epsilon_T\epsilon_B + 1,\ \delta_0 = 2\left(\epsilon_T + \epsilon_T\right). \tag{A.1.2}$$

Rearranging, we get

$$F^{\text{class}}_{\text{loss/amp}}(0, \epsilon_T, \epsilon_B) = \frac{\sqrt{2\gamma_0 + 2\sqrt{\gamma_0^2 - \delta_0^2}}}{\delta_0}, \tag{A.1.3}$$

and then, using

$$\sqrt{\gamma_0 \pm \delta_0} = \sqrt{(2\epsilon_T \pm 1)(2\epsilon_B \pm 1)}, \tag{A.1.4}$$

we get

$$\sqrt{\gamma_0^2 - \delta_0^2} = \sqrt{(4\epsilon_T^2 - 1)(4\epsilon_B^2 - 1)}. \tag{A.1.5}$$

Thus, we have

$$F^{\text{class},\tau=0}_{\text{loss/amp}} = \frac{\sqrt{4\epsilon_T\epsilon_B + 1 + \sqrt{(4\epsilon_T^2 - 1)(4\epsilon_B^2 - 1)}}}{\sqrt{2}(\epsilon_T + \epsilon_B)} \tag{A.1.6}$$

$$= F^{\infty}_{\text{loss/amp}}. \tag{A.1.7}$$

The proofs that $\frac{dF}{d\tau}$ is positive semidefinite in the range $0 \leq \tau < 1$, that $\frac{dF}{d\tau}$ is negative semidefinite in the range $\tau > 1$, and that $\frac{(F^{\text{class}})^2}{F^{\infty}} \leq 1$ for $\tau = \frac{1}{2}$ are straightforward but lengthy to write out and so are given in the supplementary Mathematica files of Ref. [9].

# Appendix B

# Appendices for Chapter 4

## B.1 Location of the minima of the trace norm, for the Choi resource

Let us calculate the trace norm by finding the eigenvalues of the matrix resulting from taking the difference of the Choi matrices of the simulated and simulating channel (i.e. the right hand side of Eq. (4.92)). This matrix has eigenvalues $e_i$, where $e_1$ and $e_2$ have already been given in Eq. (4.93). The remaining eigenvalues are:

$$e_3 = -\frac{1}{2}\left((e_1+e_2) + \sqrt{(e_1-e_2)^2+4c^2}\right), \tag{B.1.1}$$

$$e_4 = -\frac{1}{2}\left((e_1+e_2) - \sqrt{(e_1-e_2)^2+4c^2}\right). \tag{B.1.2}$$

The trace norm is the sum of the absolute values of the eigenvalues. We can show that $e_3$ is always negative and $e_4$ is always positive. We start by showing that $|e_1+e_2| \leq \sqrt{(e_1-e_2)^2+4c^2}$. Note that $e_1$ is a linear function of $p_1$ that is always positive and that $e_2$ is a linear function of $p_1$ that goes to 0 at $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$, and is negative for $p_1$ less than this value. For $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$, $e_1+e_2 = 0$, and above this value of $p_1$, it is positive. We can therefore show that $2c \geq |e_1+e_2|$ in the regime in which $e_1+e_2$ is positive, using

$$\frac{d(2c)}{dp_1} = \frac{1-\xi_N}{2\sqrt{1-p_1}}, \qquad \frac{d(e_1+e_2)}{dp_1} = \frac{1-\xi_N}{2}, \qquad \frac{d(2c)}{dp_1} \geq \frac{d(e_1+e_2)}{dp_1}, \tag{B.1.3}$$

$$2c\big|_{p_1=\frac{p_0-\xi_N}{1-\xi_N}} = \sqrt{1-p_0} - (1-\xi_N)\sqrt{1-\frac{p-\xi_N}{1-\xi_N}} = \sqrt{1-p_0}(1-\sqrt{1-\xi_N}) \geq 0. \tag{B.1.4}$$

Since the gradient of $2c$ is always larger than the gradient of $e_1+e_2$ in this regime and $c$ is positive at $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$, whilst $e_1+e_2$ is equal to 0, $2c \geq |e_1+e_2|$ for $p_1 \geq \frac{p_0-\xi_N}{1-\xi_N}$. For $p_1 < \frac{p_0-\xi_N}{1-\xi_N}$, $e_1-e_2 = \frac{p_0-p_1}{2} \geq |e_1+e_2|$, because $e_2$ is negative in this region. Hence, at all points,

$$|e_1+e_2| \leq \max[e_1-e_2, 2c] \leq \sqrt{(e_1-e_2)^2+4c^2}. \tag{B.1.5}$$

As a result, $e_3$ is always negative and $e_4$ is always positive. We therefore find

$$|e_3| + |e_4| = \sqrt{(e_1 - e_2)^2 + 4c^2}. \tag{B.1.6}$$

$|e_1| + |e_2|$ has two regimes, corresponding to $p_1 \leq \frac{2p_0 - \xi_N}{2 - \xi_N}$ and $p_1 > \frac{2p_0 - \xi_N}{2 - \xi_N}$. In the first regime, $|e_1| + |e_2| = \frac{p_0 - p_1}{2}$ and in the second, $|e_1| + |e_2| = \frac{\xi_N}{2}(1 - p_1) - \frac{p_0 - p_1}{2}$. The gradient of $|e_1| + |e_2|$ is $-\frac{1}{2}$ in the first regime and $\frac{1 - \xi_N}{2}$ in the second regime, with a discontinuity at $p_1 = \frac{2p_0 - \xi_N}{2 - \xi_N}$. Taking the second derivative of $(e_1 - e_2)^2 + 4c^2$, we find that it is always positive, so the gradient of $|e_3| + |e_4|$ is always increasing, and hence $|e_3| + |e_4|$ has at most one minimum.

The gradient of $|e_3| + |e_4|$ is given by

$$\frac{d|e_3| + |e_4|}{dp_1} = \frac{p_1 - p_0 + 2(1 - \xi_N)\left(\sqrt{\frac{1-p_0}{1-p_1}} - (1 - \xi_N)\right)}{4\sqrt{\frac{p_0 - p_1}{2}^2 + (\sqrt{1 - p_0} - (1 - \xi_N)\sqrt{1 - p_1})^2}}, \tag{B.1.7}$$

and the gradient of the total trace norm, $D_{\text{trace}}$, is given by

$$\left.\frac{dD_{\text{trace}}}{dp_1}\right|_{p_1 < \frac{2p_0 - \xi_N}{2 - \xi_N}} = \frac{d|e_3| + |e_4|}{dp_1} - \frac{1}{2}, \tag{B.1.8}$$

$$\left.\frac{dD_{\text{trace}}}{dp_1}\right|_{p_1 > \frac{2p_0 - \xi_N}{2 - \xi_N}} = \frac{d|e_3| + |e_4|}{dp_1} + \frac{1 - \xi_N}{2}. \tag{B.1.9}$$

Note that the expressions for the gradient of the trace norm are different in each regime (on either side of the discontinuity).

Consider the case in which the minimum of $|e_3| + |e_4|$ occurs "after" the discontinuity (i.e. at $p_1 > \frac{2p_0 - \xi_N}{2 - \xi_N}$). There are two possibilities: if the (second) expression for the gradient of the trace norm assessed at $p_1 = \frac{2p_0 - \xi_N}{2 - \xi_N}$ is negative, the minimum of the trace norm will lie in the region $p_1 > \frac{2p_0 - \xi_N}{2 - \xi_N}$, whereas if it is positive, there is no stationary point and the minimum of the trace norm is located exactly at the discontinuity. By numerically minimising the expression for the gradient assessed at the discontinuity over $p$ (between 0 and 1) and over $\xi_N$ (between 0 and $\frac{6 - \sqrt{3}}{6}$), we find that it is always positive. Hence, if the minimum of $|e_3 + e_4|$ occurs at $p_1 > \frac{2p_0 - \xi_N}{2 - \xi_N}$, the minimum of the trace norm lies at $\frac{2p_0 - \xi_N}{2 - \xi_N}$. Note that this is the point at which $e_2 = 0$.

Similarly, if the minimum of $|e_3| + |e_4|$ occurs "before" the discontinuity, but the (first) expression for the gradient of the trace norm remains negative up to the discontinuity, the minimum of the trace norm will be at the discontinuity. Solving for this gradient to equal 0, we get a polynomial in $\xi_N$ and $p_0$, giving the value of $p_1$ at which the minimum of the trace norm occurs (or would occur, if it is after the discontinuity). When this value becomes less than $\frac{2p_0 - \xi_N}{2 - \xi_N}$, the minimum of the trace norm lies at the value of the polynomial, rather than at the discontinuity. We can find

the value of $p_0$ at which this occurs for a given value of $\xi_N$. This is a polynomial function of $\xi_N$. Higher values of $\xi_N$ require higher values of $p_0$, and the minimum value of $p_0$ for which the minimum of the trace norm can occur in the the region $p_1 < \frac{2p_0-\xi_N}{2-\xi_N}$ is $\frac{2}{5}$. For all $p_0 < \frac{2}{5}$, the minimum trace norm always lies at $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$.

We can find the value of $p_0$ at which the minimum of the trace norm crosses the line $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$, which we denote $p_0^{\text{cross}}$. We find that we have another polynomial function of $\xi_N$:

$$p_0^{\text{cross}} = \frac{1 + 4\xi_N - 8\xi_N^2 + 5\xi_N^3 + (1-\xi_N)^{\frac{7}{2}} - \xi_N^4}{3 - 3\xi_N + \xi_N^2}. \tag{B.1.10}$$

This function has a minimum value of $\frac{2}{3}$, at $\xi_N = 0$. Note that if $p_0 \leq p_0^{\text{cross}}$, the gradient of $|e_3| + |e_4|$ is always negative in the range $p_1 < \frac{p_0-\xi_N}{1-\xi_N}$ and is always positive in the range $p_1 > p_0$, and hence the same is true of the gradient of the trace norm. Hence, for all $p_0 \leq \frac{2}{3}$, we are guaranteed that the minimum of the diamond norm lies between $p_1 = \frac{p_0-\xi_N}{1-\xi_N}$ and $p_1 = \frac{2p_0-\xi_N}{2-\xi_N}$.

## B.2 Comparison of the alternate resource and the Choi resource at known points and low damping

Carrying out PBT using a resource consisting of $N$ copies of the state in Eq. (4.97) (which we will call the alternate resource) results in the Choi matrix given in Eq. (4.98). The difference between Choi matrices with the AD channel is (as given in the main text)

$$PBT\left[R_{\text{new}}(a)^{\otimes N}\right] - R'(p_0) = \begin{pmatrix} x - \frac{1}{2} & 0 & 0 & z - \frac{\sqrt{1-p_0}}{2} \\ 0 & \frac{1}{2} - x & 0 & 0 \\ 0 & 0 & y - \frac{p_0}{2} & 0 \\ z - \frac{\sqrt{1-p_0}}{2} & 0 & 0 & \frac{p_0}{2} - y \end{pmatrix}, \tag{B.2.11}$$

with $x$, $y$ and $z$ defined in the main text. We define $a^{\text{known}}$ as the value of $a$ such that the first diagonal element of this matrix is the same as the third diagonal element. This is a value of $a$ for which the diamond norm is known analytically and is equal to the trace norm between Choi matrices; we refer to this as a known point. At the point $a^{\text{known}} = \frac{1}{2}$ the resource state is simply a maximally entangled state.

Carrying out PBT using a resource consisting of $N$ copies of the state in Eq. (4.88) (which we will call the Choi resource) results in the Choi matrix given in Eq. (4.89), and the difference

between Choi matrices is (as given in the main text)

$$PBT[R(p_1)^{\otimes N}] - R'(p_0) = \begin{pmatrix} -e_1 & 0 & 0 & -c \\ 0 & e_1 & 0 & 0 \\ 0 & 0 & e_2 & 0 \\ -c & 0 & 0 & -e_2 \end{pmatrix}, \tag{B.2.12}$$

with $e_1$, $e_2$ and $c$ defined in the main text. We define $p_1^{\text{known}}$ as the value of $p_1$ such that the first diagonal element of this matrix is the same as the third diagonal element, similarly to $a^{\text{known}}$. The minimum value of $p_1^{\text{known}}$ is 0; at this point the resource state is again a maximally entangled state.

The corresponding $p_0$ value for $a^{\text{known}} = \frac{1}{2}$ is $\frac{\xi_N}{2}$. The corresponding $p_0$ value for $p_1^{\text{known}} = 0$ is also $\frac{\xi_N}{2}$. Consequently, at this point, both resources simulate the AD channel equally well. Differentiating the expression in Eq. (4.96), we find that the gradient of the diamond norm for the Choi resource at $p_1 = p_1^{\text{known}}$, $D_\diamond^1$, is

$$\frac{dD_\diamond^1}{dp_0} = -\frac{1}{2}\left( \frac{\xi_N}{1-\xi_N} + \frac{2\left(1-\sqrt{1-\xi_N}\right)^2 + \frac{(1-p_0)\xi_N^2}{(1-\xi_N)^2}}{\sqrt{4(1-p_0)\left(1-\sqrt{1-\xi_N}\right)^2 + \frac{(1-p_0)^2\xi_N^2}{(1-\xi_N)^2}}} \right), \tag{B.2.13}$$

which is finite and negative for all $\xi_N < 1$ (a condition which holds for all $N \geq 2$). We will now show that the gradient of the diamond norm for the alternate resource at $a = a^{\text{known}}$, which we will denote as $D_\diamond^2$, diverges as $a^{\text{known}}$ tends to $\frac{1}{2}$ from above.

We first find that $D_\diamond^2$ takes the form

$$D_\diamond^2 = p_0 - 2y + \sqrt{(p_0-2y)^2 + (\sqrt{1-p_0}-2z)^2}\Big|_{a=a^{\text{known}}(p_0)}, \tag{B.2.14}$$

by using the fact that the eigenvalues of a matrix of the form

$$\begin{pmatrix} x_1 & 0 & 0 & x_2 \\ 0 & -x_1 & 0 & 0 \\ 0 & 0 & x_1 & 0 \\ x_2 & 0 & 0 & -x_1 \end{pmatrix} \tag{B.2.15}$$

are $\{\pm x_1, \sqrt{x_1^2 + x_2^2}\}$. We then differentiate $D_\diamond^2$, getting

$$\begin{aligned}\frac{dD_\diamond^2}{dp_0} &= 1 - 2\frac{dy}{da}\frac{da^{\text{known}}}{dp_0} \\ &\quad + \frac{(p_0 - 2y)\left(1 - 2\frac{dy}{da}\frac{da^{\text{known}}}{dp_0}\right) + (\sqrt{1-p_0} - 2z)\left(\frac{-1}{2\sqrt{1-p_0}} - 2\frac{dz}{da}\frac{da^{\text{known}}}{dp_0}\right)}{\sqrt{(p_0-2y)^2 + (\sqrt{1-p_0} - 2z)^2}} \\ &= \left(1 + \frac{(p_0 - 2y) - \frac{1}{2} + \frac{2z}{2\sqrt{1-p_0}}}{\sqrt{(p_0-2y)^2 + (\sqrt{1-p_0} - 2z)^2}}\right) \\ &\quad - 2\frac{da^{\text{known}}}{dp_0}\left(\frac{dy}{da} + \frac{(p_0 - 2y)\frac{dy}{da} + (\sqrt{1-p_0} - 2z)\frac{dz}{da}}{\sqrt{(p_0-2y)^2 + (\sqrt{1-p_0} - 2z)^2}}\right)\end{aligned} \tag{B.2.16}$$

where $y$ and $z$ are evaluated at $a = a^{\text{known}}(p_0)$. We will show that the term in the right-hand bracket of Eq. (B.2.16) is positive sufficiently close to $a = \frac{1}{2}$. Note that since $x \leq \frac{1}{2}$ and $x - \frac{1}{2} = y - \frac{p_0}{2}$, $p_0 - 2y \geq 0$

Let us find an expression for $\frac{dy}{da}$. Recall that $y$ is given by

$$\begin{aligned}y = &\sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m} \\ &\qquad \times \frac{N!(s+m)(s-m+1)\left[\left(\frac{N+1}{2} - s\right)^{-\frac{1}{2}} - \left(\frac{N+3}{2} + s\right)^{-\frac{1}{2}}\right]^2}{2\left(\frac{N-1}{2} - s\right)!\left(\frac{N+1}{2} + s\right)!(2s+1)} \\ &+ \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\frac{\left(\frac{N-1}{2} + m\right)\left(\frac{N+1}{2} + m\right)}{2N(N+1)},\end{aligned} \tag{B.2.17}$$

and define $\text{cont}_1^y(s,m)$ and $\text{cont}_2^y(m)$ such that

$$\begin{aligned}y = &\sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\text{cont}_1^y(s,m) \\ &+ \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\text{cont}_2^y(m),\end{aligned} \tag{B.2.18}$$

noting that $\text{cont}_1^y(s,m)$ and $\text{cont}_2^y(m)$ have no $a$-dependence. Hence, applying the product rule of differentiation,

$$\begin{aligned}\frac{dy}{da} = &\frac{N(1-2a)}{2a(1-a)}y + \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\frac{2m-1}{2a(1-a)}\text{cont}_1^y(s,m) \\ &+ \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\frac{2m-1}{2a(1-a)}\text{cont}_2^y(m).\end{aligned} \tag{B.2.19}$$

Note that if $m$ goes to $1-m$, $\text{cont}_1^y(s,m)$ is unchanged (i.e. $\text{cont}_1^y(s,m) = \text{cont}_1^y(s,1-m)$) and $2m-1$ goes to $-(2m-1)$. Note too that $\text{cont}_1^y(s,-s) = 0$ and that $m = \frac{1}{2}$ sets $2m-1$ to 0, meaning that we can write

$$\begin{aligned}\frac{dy}{da} = \frac{N(1-2a)}{2a(1-a)}y + \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=\{1,\frac{3}{2}\}}^{s} \Big( & a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m} \\ & - a^{\frac{N+1}{2}-m}(1-a)^{\frac{N-1}{2}+m}\Big)\frac{2m-1}{2a(1-a)}\text{cont}_1^y(s,m) \\ & + \sum_{m=\{1,\frac{3}{2}\}}^{\frac{N+1}{2}} \Big( a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\text{cont}_2^y(m) \\ & - a^{\frac{N+1}{2}-m}(1-a)^{\frac{N-1}{2}+m}\text{cont}_2^y(1-m)\Big)\frac{2m-1}{2a(1-a)},\end{aligned} \tag{B.2.20}$$

where the minimum value of $m$ is 1 for odd $N$ and $\frac{3}{2}$ for even $N$. We now note that, for $a \geq \frac{1}{2}$,

$$a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m} \geq a^{\frac{N+1}{2}-m}(1-a)^{\frac{N-1}{2}+m}, \tag{B.2.21}$$

with equality only at $a = \frac{1}{2}$, meaning that sufficiently close to $a = \frac{1}{2}$, the second sum in Eq. (B.2.19) dominates. Note too that $\text{cont}_2^y(m) > \text{cont}_2^y(1-m)$ (with a finite difference between $\text{cont}_2^y(m)$ and $\text{cont}_2^y(1-m)$ that does not depend on $a$), and hence $\frac{dy}{da} > 0$ for $a$ sufficiently close to $\frac{1}{2}$.

Let us now find an expression for $\frac{dz}{da}$. Recall that $z$ is given by

$$\begin{aligned}z = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} & \frac{a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}N!}{2\left(\frac{N-1}{2}-s\right)!\left(\frac{N+1}{2}+s\right)!(2s+1)}\Bigg[\left(\frac{N+1}{2}-s\right)^{-1}(s^2-m^2) \\ & + 2\left(\frac{N+1}{2}-s\right)^{-\frac{1}{2}}\left(\frac{N+3}{2}+s\right)^{-\frac{1}{2}}(s^2+m^2+s) \\ & + \left(\frac{N+3}{2}+s\right)^{-1}((s+1)^2-m^2)\Bigg] \\ & - \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\frac{\left(\frac{N+1}{2}+m\right)\left(\frac{N+1}{2}-m\right)}{2N(N+1)},\end{aligned} \tag{B.2.22}$$

and define $\text{cont}_1^z(s,m)$ and $\text{cont}_2^z(m)$ such that

$$z = \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\text{cont}_1^z(s,m) + \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\text{cont}_2^z(m). \tag{B.2.23}$$

Differentiating, we get

$$\begin{aligned}\frac{dz}{da} = \frac{N(1-2a)}{2a(1-a)}z + \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=-s}^{s} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\frac{m}{a(1-a)}\text{cont}_1^z(s,m) \\ + \sum_{m=-\frac{N+1}{2}}^{\frac{N+1}{2}} a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m}\frac{m}{a(1-a)}\text{cont}_2^z(m).\end{aligned} \tag{B.2.24}$$

Note that $\text{cont}_1^z(s,m) = \text{cont}_1^z(s,-m)$ and $\text{cont}_2^z(s,m) = \text{cont}_2^z(s,-m)$. Hence, we can write

$$\begin{aligned}\frac{dz}{da} = \frac{N(1-2a)}{2a(1-a)}z + \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=\{1,\frac{3}{2}\}}^{s} \Big(a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m} \\ - a^{\frac{N}{2}-m}(1-a)^{\frac{N}{2}+m}\Big)\frac{m}{a(1-a)}\text{cont}_1^z(s,m) \\ + \sum_{m=\{1,\frac{3}{2}\}}^{\frac{N+1}{2}} \Big(a^{\frac{N}{2}+m}(1-a)^{\frac{N}{2}-m} - a^{\frac{N}{2}-m}(1-a)^{\frac{N}{2}+m}\Big)\frac{m}{a(1-a)}\text{cont}_2^z(s,m).\end{aligned} \tag{B.2.25}$$

Note that this approaches 0 as $a$ approaches $\frac{1}{2}$, hence there exists some finite, positive $\epsilon$ such that for all $\frac{1}{2} \leq a \leq \frac{1}{2} + \epsilon$, we have

$$\frac{dy}{da} + \frac{(p_0 - 2y)\frac{dy}{da} + (\sqrt{1-p_0} - 2z)\frac{dz}{da}}{\sqrt{(p_0-2y)^2 + (\sqrt{1-p_0}-2z)^2}} > 0. \tag{B.2.26}$$

It now suffices to show that $\frac{da^{\text{known}}}{dp_0}$ diverges as $a$ tends to $\frac{1}{2}$ from above. We write

$$\frac{da^{\text{known}}}{dp_0} = \left(\frac{dp_0}{da^{\text{known}}}\right)^{-1} = \frac{d}{da}\left(1 - 2(x-y)\right) = -2\frac{d}{da}(x-y). \tag{B.2.27}$$

Using the symmetry of the PBT protocol, we can see that $x[a] = \frac{1}{2} - y[1-a]$. We can therefore write

$$\frac{dp_0}{da^{\text{known}}} = 2\frac{d}{da}\left(y[a] + y[1-a]\right). \tag{B.2.28}$$

The differential $\frac{dy[a]}{da}$ is given in Eq. (B.2.19), and we can similarly write

$$\begin{aligned}\frac{dy[1-a]}{da} =& \frac{N(1-2a)}{2a(1-a)}y[1-a] + \sum_{s=s_{\min}}^{\frac{N-1}{2}} \sum_{m=\{1,\frac{3}{2}\}}^{s} \Big(a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m} \\ & - a^{\frac{N+1}{2}-m}(1-a)^{\frac{N-1}{2}+m}\Big)\frac{2m-1}{2a(1-a)}\text{cont}_1^y(s,m) \\ & + \sum_{m=\{1,\frac{3}{2}\}}^{\frac{N+1}{2}} \Big(a^{\frac{N-1}{2}+m}(1-a)^{\frac{N+1}{2}-m}\text{cont}_2^y(1-m) \\ & - a^{\frac{N+1}{2}-m}(1-a)^{\frac{N-1}{2}+m}\text{cont}_2^y(m)\Big)\frac{2m-1}{2a(1-a)}.\end{aligned} \tag{B.2.29}$$

The expression $y[a] + y[1-a]$ is symmetric around $a = \frac{1}{2}$ and both $\frac{dy[a]}{da}$ and $\frac{dy[1-a]}{da}$ are finite at this point, so $a = \frac{1}{2}$ is either a maximum or a minimum of this expression.

Suppose that it is a minimum. Numerically, we find a clear trend indicating that this is the case for all $N$, with the second differential tending towards 1 from below (from a value of 0 at $N = 2$) as $N$ increases. Then, $\frac{da^{\text{known}}}{dp_0}$ diverges to positive infinity as $a$ approaches $\frac{1}{2}$ from above. Consequently, $\frac{dD_\diamond^2}{dp_0}$ diverges to negative infinity. Hence, there exists some finite positive $\epsilon$ such that the gradient of the diamond norm for the Choi resource, assessed at $p_0 = \frac{\xi_N}{2} + \delta$ is less negative than the gradient of the diamond norm for the alternate resource, assessed at the same point, for all positive $\delta < \epsilon$. Consequently, the diamond norm for the Choi resource at the known point is less than the diamond norm for the alternate resource for all $\frac{\xi_N}{2} < p_0 \leq \epsilon$.

Suppose instead that it is a maximum. Then, $\frac{da^{\text{known}}}{dp_0}$ diverges to negative infinity as $a$ approaches $\frac{1}{2}$ from above, and $\frac{dD_\diamond^2}{dp_0}$ diverges to positive infinity. However, in this case, increasing $a$ by a small amount from $\frac{1}{2}$ decreases $p_0$, since $\frac{dp_0}{da^{\text{known}}}$ is negative. Consequently, there exists some finite positive $\epsilon$ such that $D_\diamond^2$ assessed at $p_0 = \frac{\xi_N}{2} - \delta$ is lower than $D_\diamond^1$ assessed at $p_0 = \frac{\xi_N}{2} + \delta$ for all positive $\delta < \epsilon$. In this case, an AD channel applied to the output of the PBT channel, with the damping probability $p'$ chosen such that total channel simulates an AD channel with $p_0 = \frac{\xi_N}{2} + \delta$ would result in $D_\diamond^2 < D_\diamond^1$. This is equivalent to using the tensor-product resource composed of $N$ copies of

$$R'_{\text{new}}(a) = \begin{pmatrix} p'(1-a) & 0 & 0 & 0 \\ 0 & a & -\sqrt{a(1-a)(1-p')} & 0 \\ 0 & -\sqrt{a(1-a)(1-p')} & (1-a)(1-p') & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{B.2.30}$$

which is still distinct from any state of the form in Eq. (4.88).

Hence, in either case and for any $N$, there exists some tensor-product resource that simulates the AD channel better than the Choi resource, at either of its known points, for some range of $p_0$ values.

# Appendix C

# Appendices for Chapter 6

## C.1 Calculation of the k-value for any m-value

The steps to study the setup in Fig. 6.9 are very similar to those for the $m = 1$ case. By using a beamsplitter on modes 1 and 2 followed by two-mode squeezers on modes 2 and 3 and then on modes 1 and 3, we can show that the setup is equivalent to one in which the signal state is modulated by $k_1\alpha$ and a single pure side-channel mode is modulated by $k_2\mathbb{Z}\alpha$ (as in Fig. 6.2, but with different values for $k_1$ and $k_2$). We then again use the fact that this gives the same key rate as a setup in which the side-channel mode is modulated by $k_2\alpha$ instead of by $k_2\mathbb{Z}\alpha$ and hence that it gives the same key rate as one in which the signal state is modulated by $k = \sqrt{k_1^2 + k_2^2}$, with a beamsplitter in the main channel.

We label the initial covariance matrix of the total state as $V_0^{m\neq 1}$, the initial covariance matrix for fixed $\alpha$ as $V_0^{m\neq 1}|\alpha$, and the initial quadratures for fixed $\alpha$ as $X_0^{m\neq 1}|\alpha$ and then use the subscripts 1, 2 and 3 to denote these objects after the beamsplitter, the first two-mode squeezer and the second two-mode squeezer respectively. The optical circuit is the same as in Fig. 6.3; only the parameters of the optical components are changed for the $m \neq 1$ case.

The first and second moments of the initial state are

$$X_0^{m\neq 1}|\alpha = \begin{pmatrix} \alpha \\ m\alpha \\ 0 \end{pmatrix}, \tag{C.1.1}$$

$$V_0^{m\neq 1}|\alpha = \begin{pmatrix} \mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cosh 2r\mathbb{I} & \sinh 2r\mathbb{Z} \\ \mathbf{0} & \sinh 2r\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}, \tag{C.1.2}$$

$$V_0^{m\neq 1} = \begin{pmatrix} (\mu+1)\mathbb{I} & m\mu\mathbb{I} & \mathbf{0} \\ m\mu\mathbb{I} & (m^2\mu+\cosh 2r)\mathbb{I} & \sinh 2r\mathbb{Z} \\ \mathbf{0} & \sinh 2r\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}. \tag{C.1.3}$$

The first optical component is a beamsplitter that sets the quadratures of modes 2 and 3 to 0 (moves the entire displacement onto mode 1). This beamsplitter has angle

$$\theta_1^{m\neq 1} = \arccos \frac{1}{\sqrt{m^2+1}}, \tag{C.1.4}$$

and changes the first and second moments of the state to

$$X_1^{m\neq 1}|\alpha = \begin{pmatrix} \sqrt{m^2+1}\alpha \\ 0 \\ 0 \end{pmatrix}, \tag{C.1.5}$$

$$V_1^{m\neq 1}|\alpha = \begin{pmatrix} \frac{m^2\cosh 2r+1}{m^2+1}\mathbb{I} & \frac{2m\sinh^2 r}{m^2+1}\mathbb{I} & my^{(1)}\mathbb{Z} \\ \frac{2m\sinh^2 r}{m^2+1}\mathbb{I} & \frac{m^2+\cosh 2r}{m^2+1}\mathbb{I} & y^{(1)}\mathbb{Z} \\ my^{(1)}\mathbb{Z} & y^{(1)}\mathbb{Z} & \cosh 2r\mathbb{I} \end{pmatrix}, \tag{C.1.6}$$

$$V_1^{m\neq 1} = V_1^{m\neq 1}|\alpha \oplus (m^2+1)\mu \begin{pmatrix} \mathbb{I} & & \\ & \mathbf{0} & \\ & & \mathbf{0} \end{pmatrix}, \tag{C.1.7}$$

where $y^{(1)} = (\sinh 2r)/\sqrt{m^2+1}$.

The next component purifies the second mode, reducing the state to a bipartite state. It acts on the second and third modes and has squeezing parameter $r_2^{m\neq 1} = -\text{arcsinh}\frac{\sqrt{2}\sinh r}{\sqrt{m^2\cosh 2r+m^2+2}}$

. The first and second moments become

$$X_2^{m\neq 1}|\alpha = \begin{pmatrix} \sqrt{m^2+1}\alpha \\ 0 \\ 0 \end{pmatrix}, \tag{C.1.8}$$

$$V_2^{m\neq 1}|\alpha = \begin{pmatrix} \frac{m^2\cosh 2r+1}{m^2+1}\mathbb{I} & \mathbf{0} & y^{(2)}\mathbb{Z} \\ \mathbf{0} & \mathbb{I} & \mathbf{0} \\ y^{(2)}\mathbb{Z} & \mathbf{0} & \frac{m^2\cosh 2r+1}{m^2+1}\mathbb{I} \end{pmatrix}, \tag{C.1.9}$$

$$V_2^{m\neq 1} = V_2^{m\neq 1}|\alpha \oplus (m^2+1)\mu \begin{pmatrix} \mathbb{I} & & \\ & \mathbf{0} & \\ & & \mathbf{0} \end{pmatrix}, \tag{C.1.10}$$

where

$$y^{(2)} = \frac{\sqrt{2}m\sinh r\sqrt{m^2\cosh 2r+m^2+2}}{m^2+1}. \tag{C.1.11}$$

The final component unsqueezes the remaining two modes, such that the state for fixed $\alpha$ is a vacuum state. The squeezing parameter is $r_3^{m\neq 1} = -\text{arcsinh}\frac{m\sinh r}{\sqrt{m^2+1}}$. The first and second moments become

$$X_3^{m\neq 1}|\alpha = \begin{pmatrix} \frac{\sqrt{m^2\cosh 2r+m^2+2}}{\sqrt{2}}\alpha \\ -m\sinh r\mathbb{Z}\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} k_1^{m\neq 1}\alpha \\ k_2^{m\neq 1}\mathbb{Z}\alpha \\ 0 \end{pmatrix}, \tag{C.1.12}$$

$$V_3^{m\neq 1}|\alpha = \begin{pmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{pmatrix}, \; V_3^{m\neq 1} = \begin{pmatrix} x_+\mathbb{I} & y^{(3)}\mathbb{Z} \\ y^{(3)}\mathbb{Z} & x_-\mathbb{I} \end{pmatrix}, \tag{C.1.13}$$

where

$$x_\pm = \frac{1}{2}(m^2\mu\cosh 2r \pm m^2\mu + 2), \tag{C.1.14}$$

$$y^{(3)} = -\frac{m\mu\sinh r\sqrt{m^2\cosh 2r+m^2+2}}{\sqrt{2}}. \tag{C.1.15}$$

Since we have shown that there is an optical circuit that reversibly converts the initial state of the setup in Fig. 6.9 to the initial state of the setup in Fig. 6.2, the two setups must have the same secret key rate for the same thermal noise. As shown in the main text, this also means that the setup in Fig. 6.9 has the same secret key rate as the side-channel-free setup with an "effective modulation" of $\mu' = k^2\mu$, an "effective channel loss" of $\eta' = \frac{\eta}{k^2}$ and an "effective excess noise"

of $\epsilon' = k^2\epsilon$, where

$$k = \sqrt{k_1^2 + k_2^2} \tag{C.1.16}$$

$$= \sqrt{\frac{1}{2}(m^2 \cosh 2r + m^2 + 2) + m^2 \sinh^2 r} \tag{C.1.17}$$

$$= \sqrt{m^2(2\bar{n} + 1) + 1}. \tag{C.1.18}$$

This is the result given in the main text.

# Abbreviations

**AD** Amplitude Damping

**CPF** Channel Position Finding

**CV** Continuous Variable

**DV** Discrete Variable

**LB** Lower Bound

**LIDT** Laser Induced Damage Threshold

**LOCC** Local Operations and Classical Communications

**MLE** Maximum-Likelihood Estimation

**PBT** Port-Based Teleportation

**PLOB** Pirandola-Laurenza-Ottaviani-Banchi

**QCRB** Quantum Cramér-Rao Bound

**QFI** Quantum Fisher Information

**QKD** Quantum Key Distribution

**REE** Relative Entropy of Entanglement

**RNG** Random Number Generator

**TMSV** Two-Mode Squeezed Vacuum

**UB** Upper Bound

# References


[1] Nechita, I., Puchała, Z., Pawela, Ł., and Życzkowski, K. (May, 2018) Almost All Quantum Channels Are Equidistant. *J. Math. Phys,* **59**(5), 052201.

[2] Pirandola, S., Laurenza, R., Lupo, C., and Pereira, J. L. (June, 2019) Fundamental Limits to Quantum Channel Discrimination. *NPJ Quantum Inf.,* **5**(1), 50.

[3] Spedalieri, G., Piersimoni, L., Laurino, O., Braunstein, S. L., and Pirandola, S. (November, 2020) Detecting and Tracking Bacteria with Quantum Light. *Phys. Rev. Research,* **2**(4), 043260.

[4] Pirandola, S., Braunstein, S. L., and Lloyd, S. (November, 2008) Characterization of Collective Gaussian Attacks and Security of Coherent-State Quantum Cryptography. *Phys. Rev. Lett.,* **101**(20).

[5] Pirandola, S., Laurenza, R., Ottaviani, C., and Banchi, L. (April, 2017) Fundamental Limits of Repeaterless Quantum Communications. *Nat. Commun.,* **8**(1), 15043.

[6] Pirandola, S., Braunstein, S. L., Laurenza, R., Ottaviani, C., Cope, T. P. W., Spedalieri, G., and Banchi, L. (May, 2018) Theory of Channel Simulation and Bounds for Private Communication. *Quantum Sci. Technol.,* **3**(3), 035009.

[7] Weedbrook, C., Lance, A. M., Bowen, W. P., Symul, T., Ralph, T. C., and Lam, P. K. (October, 2004) Quantum Cryptography Without Switching. *Phys. Rev. Lett.,* **93**(17), 170504.

[8] Pereira, J. and Pirandola, S. (December, 2018) Hacking Alice's Box in Continuous-Variable Quantum Key Distribution. *Phys. Rev. A,* **98**(6), 062319.

[9] Pereira, J. L., Zhuang, Q., and Pirandola, S. (November, 2020) Optimal Environment Localization. *Phys. Rev. Research,* **2**(4), 043189.

[10] Pereira, J. L. and Pirandola, S. (February, 2021) Bounds on Amplitude-Damping-Channel Discrimination. *Phys. Rev. A,* **103**(2), 022610.

[11] Pereira, J. L., Banchi, L., and Pirandola, S. (2021) Characterising Port-Based Teleportation as a Universal Simulator of Qubit Channels. *J. Phys. A,* (in press).

[12] Pereira, J. L., Banchi, L., Zhuang, Q., and Pirandola, S. (October, 2020) Idler-Free Channel Position Finding. *arXiv:2010.10547,*.

[13] Cover, T. M. and Thomas, J. A. (November, 2012) Elements of Information Theory, John Wiley & Sons, USA.

[14] MacKay, D. J. C. (2003) Information Theory, Inference and Learning Algorithms, Cambridge University Press, USA.

[15] Nielsen, M. A. and Chuang, I. L. (2010) Quantum Computation and Quantum Information: 10th Anniversary Edition, Cambridge University Press, USA tenth edition.

[16] Gyongyosi, L., Imre, S., and Nguyen, H. V. (Secondquarter 2018) A Survey on Quantum Channel Capacities. *IEEE Commun. Surv. Tutor.,* **20**(2), 1149–1205.

[17] Bennett, C. H., DiVincenzo, D. P., and Smolin, J. A. (April, 1997) Capacities of Quantum Erasure Channels. *Phys. Rev. Lett.,* **78**(16), 3217–3220.

[18] Siudzińska, K. (October, 2020) Classical Capacity of Generalized Pauli Channels. *J. Phys. A,* **53**(44), 445301.

[19] Bennett, C. H. and Wiesner, S. J. (1992) Communication via One-and Two-Particle Operators on Einstein-Podolsky-Rosen States. *Phys. Rev. Lett.,* **69**(20), 2881.

[20] Pirandola, S., Pirandola, S., Andersen, U. L., Banchi, L., Berta, M., Bunandar, D., Colbeck, R., Englund, D., Gehring, T., Lupo, C., Ottaviani, C., Pereira, J. L., Razavi, M., Shaari, J. S., Shaari, J. S., Tomamichel, M., Tomamichel, M., Usenko, V. C., Vallone, G., Villoresi, P., and Wallden, P. (December, 2020) Advances in Quantum Cryptography. *Adv. Opt. Photonics,* **12**(4), 1012–1236.

[21] Wang, G. and Ying, M. (April, 2006) Unambiguous Discrimination among Quantum Operations. *Phys. Rev. A,* **73**(4), 042301.

[22] Laurenza, R., Lupo, C., Spedalieri, G., Braunstein, S. L., and Pirandola, S. (April, 2018) Channel Simulation in Quantum Metrology. *Quantum Meas. Quantum Metrol.,* **5**(1), 1–12.

[23] Chiribella, G., D'Ariano, G. M., and Perinotti, P. (August, 2008) Quantum Circuit Architecture. *Phys. Rev. Lett.,* **101**(6), 060401.

[24] Harrow, A. W., Hassidim, A., Leung, D. W., and Watrous, J. (March, 2010) Adaptive versus Nonadaptive Strategies for Quantum Channel Discrimination. *Phys. Rev. A,* **81**(3), 032339.

[25] Pirandola, S., Bardhan, B. R., Gehring, T., Weedbrook, C., and Lloyd, S. (December, 2018) Advances in Photonic Quantum Sensing. *Nat. Photonics,* **12**(12), 724–733.

[26] Cope, T. P. W., Hetzel, L., Banchi, L., and Pirandola, S. (August, 2017) Simulation of Non-Pauli Channels. *Phys. Rev. A,* **96**(2), 022323.

[27] Pirandola, S., Laurenza, R., and Banchi, L. (January, 2019) Conditional Channel Simulation. *Ann. Phys. (N. Y.),* **400**, 289–302.

[28] Banchi, L., Pereira, J., Lloyd, S., and Pirandola, S. (2020) Convex Optimization of Programmable Quantum Computers. *NPJ Quantum Inf.,* **6**, 42.

[29] Helstrom, C. W. (June, 1969) Quantum Detection and Estimation Theory. *J. Stat. Phys.,* **1**(2), 231–252.

[30] Pirandola, S. and Lupo, C. (March, 2017) Ultimate Precision of Adaptive Noise Estimation. *Phys. Rev. Lett.,* **118**(10), 100502.

[31] Lloyd, S. (September, 2008) Enhanced Sensitivity of Photodetection via Quantum Illumination. *Science,* **321**(5895), 1463–1465.

[32] Tan, S.-H., Erkmen, B. I., Giovannetti, V., Guha, S., Lloyd, S., Maccone, L., Pirandola, S., and Shapiro, J. H. (December, 2008) Quantum Illumination with Gaussian States. *Phys. Rev. Lett.,* **101**(25), 253601.

[33] Shapiro, J. H. and Lloyd, S. (June, 2009) Quantum Illumination versus Coherent-State Target Detection. *New J. Phys.,* **11**(6), 063045.

[34] Barzanjeh, S., Guha, S., Weedbrook, C., Vitali, D., Shapiro, J. H., and Pirandola, S. (February, 2015) Microwave Quantum Illumination. *Phys. Rev. Lett.,* **114**(8), 080503.

[35] Zhuang, Q., Zhang, Z., and Shapiro, J. H. (2017) Entanglement-Enhanced Lidars for Simultaneous Range and Velocity Measurements. *Phys. Rev. A,* **96**(4), 040304.

[36] Zhuang, Q., Zhang, Z., and Shapiro, J. H. (August, 2017) Quantum Illumination for Enhanced Detection of Rayleigh-Fading Targets. *Phys. Rev. A,* **96**(2), 020302.

[37] Wilde, M. M., Tomamichel, M., Lloyd, S., and Berta, M. (September, 2017) Gaussian Hypothesis Testing and Quantum Illumination. *Phys. Rev. Lett.,* **119**(12), 120501.

[38] Zhuang, Q., Zhang, Z., and Shapiro, J. H. (August, 2017) Entanglement-Enhanced Neyman-Pearson Target Detection Using Quantum Illumination. *J. Opt. Soc. Am. B,* **34**(8), 1567–1572.

[39] De Palma, G. and Borregaard, J. (2018) Minimum Error Probability of Quantum Illumination. *Phys. Rev. A,* **98**(1), 012101.

[40] Nair, R., Nair, R., Nair, R., Gu, M., Gu, M., Gu, M., and Gu, M. (July, 2020) Fundamental Limits of Quantum Illumination. *Optica,* **7**(7), 771–774.

[41] Karsa, A., Spedalieri, G., Zhuang, Q., and Pirandola, S. (June, 2020) Quantum Illumination with a Generic Gaussian Source. *Phys. Rev. Research,* **2**(2), 023414.

[42] Lopaeva, E. D., Ruo Berchera, I., Degiovanni, I. P., Olivares, S., Brida, G., and Genovese, M. (April, 2013) Experimental Realization of Quantum Illumination. *Phys. Rev. Lett.,* **110**(15), 153603.

[43] Zhang, Z., Tengner, M., Zhong, T., Wong, F. N. C., and Shapiro, J. H. (July, 2013) Entanglement's Benefit Survives an Entanglement-Breaking Channel. *Phys. Rev. Lett.,* **111**(1), 010501.

[44] Zhang, Z., Mouradian, S., Wong, F. N. C., and Shapiro, J. H. (March, 2015) Entanglement-Enhanced Sensing in a Lossy and Noisy Environment. *Phys. Rev. Lett.,* **114**(11), 110506.

[45] Pirandola, S. (2011) Quantum Reading of a Classical Digital Memory. *Phys. Rev. Lett.,* **106**(9), 090504.

[46] Scarani, V. and Kurtsiefer, C. (December, 2014) The Black Paper of Quantum Cryptography: Real Implementation Problems. *Theor. Comput. Sci.,* **560**, 27–32.

[47] Jain, N., Stiller, B., Khan, I., Elser, D., Marquardt, C., and Leuchs, G. (July, 2016) Attacks on Practical Quantum Key Distribution Systems (and How to Prevent Them). *Contemp. Phys.,* **57**(3), 366–387.

[48] Giovannetti, V., Lloyd, S., and Maccone, L. (April, 2011) Advances in Quantum Metrology. *Nat. Photonics,* **5**(4), 222–229.

[49] Shor, P. W. (October, 1997) Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput.,* **26**(5), 1484–1509.

[50] Bernstein, D. J. (2009) Introduction to Post-Quantum Cryptography. In Bernstein, D. J., Buchmann, J., and Dahmen, E., (eds.), *Post-Quantum Cryptography*, pp. 1–14 Springer Berlin Heidelberg Berlin, Heidelberg.

[51] Bennett, C. H. and Brassard, G. (December, 2014) Quantum Cryptography: Public Key Distribution and Coin Tossing. *Theor. Comput. Sci.,* **560**, 7–11.

[52] Bennett, C. H. (May, 1992) Quantum Cryptography Using Any Two Nonorthogonal States. *Phys. Rev. Lett.,* **68**(21), 3121–3124.

[53] Boaron, A., Boso, G., Rusca, D., Vulliez, C., Autebert, C., Caloz, M., Perrenoud, M., Gras, G., Bussières, F., Li, M.-J., Nolan, D., Martin, A., and Zbinden, H. (November, 2018) Secure Quantum Key Distribution over 421 Km of Optical Fiber. *Phys. Rev. Lett.,* **121**(19), 190502.

[54] Braunstein, S. L. and van Loock, P. (June, 2005) Quantum Information with Continuous Variables. *Rev. Mod. Phys.,* **77**(2), 513–577.

[55] Weedbrook, C., Pirandola, S., García-Patrón, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H., and Lloyd, S. (May, 2012) Gaussian Quantum Information. *Rev. Mod. Phys.,* **84**(2), 621–669.

[56] Zhang, Y., Li, Z., Chen, Z., Weedbrook, C., Zhao, Y., Wang, X., Huang, Y., Xu, C., Zhang, X., Wang, Z., Li, M., Zhang, X., Zheng, Z., Chu, B., Gao, X., Meng, N., Cai, W., Wang, Z., Wang, G., Yu, S., and Guo, H. (May, 2019) Continuous-Variable QKD over 50 Km Commercial Fiber. *Quantum Sci. Technol.,* **4**(3), 035006.

[57] Brassard, G., Lütkenhaus, N., Mor, T., and Sanders, B. C. (August, 2000) Limitations on Practical Quantum Cryptography. *Phys. Rev. Lett.,* **85**(6), 1330–1333.

[58] Lütkenhaus, N. (April, 2000) Security against Individual Attacks for Realistic Quantum Key Distribution. *Phys. Rev. A,* **61**(5), 052304.

[59] Adesso, G., Ragy, S., and Lee, A. R. (June, 2014) Continuous Variable Quantum Information: Gaussian States and Beyond. *Open Syst. Inf. Dyn.,* **21**(01n02), 1440001.

[60] Olivares, S. (April, 2012) Quantum Optics in the Phase Space - A Tutorial on Gaussian States. *Eur. Phys. J. Spec. Top.,* **203**(1), 3–24.

[61] Horodecki, M., Horodecki, P., and Horodecki, R. (November, 1996) Separability of Mixed States: Necessary and Sufficient Conditions. *Phys. Lett. A,* **223**(1), 1–8.

[62] Braunstein, S. L. and Caves, C. M. (May, 1994) Statistical Distance and the Geometry of Quantum States. *Phys. Rev. Lett.,* **72**(22), 3439–3443.

[63] Braunstein, S. L., Caves, C. M., and Milburn, G. J. (April, 1996) Generalized Uncertainty Relations: Theory, Examples, and Lorentz Invariance. *Ann. Phys. (N. Y.),* **247**(1), 135–173.

[64] Watrous, J. (July, 2013) Simpler Semidefinite Programs for Completely Bounded Norms. *Chic. J. Theor. Comput. Sci.,* **2013**.

[65] Shirokov, M. E. (January, 2018) On the Energy-Constrained Diamond Norm and Its Application in Quantum Information Theory. *Probl. Inf. Transm.,* **54**(1), 20–33.

[66] Winter, A. (December, 2017) Energy-Constrained Diamond Norm with Applications to the Uniform Continuity of Continuous Variable Channel Capacities. *arXiv:1712.10267,*.

[67] Schäfer, J., Karpov, E., García-Patrón, R., Pilyavets, O. V., and Cerf, N. J. (July, 2013) Equivalence Relations for the Classical Capacity of Single-Mode Gaussian Quantum Channels. *Phys. Rev. Lett.,* **111**(3), 030503.

[68] Grosshans, F. and Grangier, P. (2002) Continuous Variable Quantum Cryptography Using Coherent States. *Phys. Rev. Lett.,* **88**(5), 057902.

[69] Hillery, M. (January, 2000) Quantum Cryptography with Squeezed States. *Phys. Rev. A,* **61**(2), 022309.

[70] Gottesman, D. and Preskill, J. (January, 2001) Secure Quantum Key Distribution Using Squeezed States. *Phys. Rev. A,* **63**(2), 022309.

[71] Navascués, M. and Acín, A. (January, 2005) Security Bounds for Continuous Variables Quantum Key Distribution. *Phys. Rev. Lett.,* **94**(2), 020505.

[72] Renner, R. and Cirac, J. I. (March, 2009) De Finetti Representation Theorem for Infinite-Dimensional Quantum Systems and Applications to Quantum Cryptography. *Phys. Rev. Lett.,* **102**(11), 110504.

[73] Pirandola, S., Serafini, A., and Lloyd, S. (May, 2009) Correlation Matrices of Two-Mode Bosonic Systems. *Phys. Rev. A,* **79**(5), 052327.

[74] Simon, R. (March, 2000) Peres-Horodecki Separability Criterion for Continuous Variable Systems. *Phys. Rev. Lett.,* **84**(12), 2726–2729.

[75] Holevo, A. S., Sohma, M., and Hirota, O. (March, 1999) Capacity of Quantum Gaussian Channels. *Phys. Rev. A,* **59**(3), 1820–1828.

[76] Banchi, L., Braunstein, S. L., and Pirandola, S. (December, 2015) Quantum Fidelity for Arbitrary Gaussian States. *Phys. Rev. Lett.,* **115**(26), 260501.

[77] Nielsen, M. A. and Chuang, I. L. (July, 1997) Programmable Quantum Gate Arrays. *Phys. Rev. Lett.,* **79**(2), 321–324.

[78] Bowen, G. and Bose, S. (December, 2001) Teleportation as a Depolarizing Quantum Channel, Relative Entropy, and Classical Capacity. *Phys. Rev. Lett.,* **87**(26), 267901.

[79] Leditzky, F., Leung, D., and Smith, G. (October, 2018) Dephrasure Channel and Superadditivity of Coherent Information. *Phys. Rev. Lett.,* **121**(16), 160501.

[80] Ishizaka, S. and Hiroshima, T. (December, 2008) Asymptotic Teleportation Scheme as a Universal Programmable Quantum Processor. *Phys. Rev. Lett.,* **101**(24), 240501.

[81] Ishizaka, S. and Hiroshima, T. (April, 2009) Quantum Teleportation Scheme by Selecting One of Multiple Output Ports. *Phys. Rev. A,* **79**(4), 042306.

[82] Vaidman, L. (February, 1994) Teleportation of Quantum States. *Phys. Rev. A,* **49**(2), 1473–1476.

[83] Braunstein, S. L. and Kimble, H. J. (1998) Teleportation of Continuous Quantum Variables. *Phys. Rev. Lett.,* **80**(4), 869–872.

[84] Holevo, A. S. (1978) On Asymptotically Optimal Hypothesis Testing in Quantum Statistics. *Theory Probab. Its Appl.,* **23**, 411–415.

[85] Montanaro, A. (August, 2007) On the Distinguishability of Random Quantum States. *Commun. Math. Phys.,* **273**(3), 619–636.

[86] Zhuang, Q. and Pirandola, S. (August, 2020) Ultimate Limits for Multiple Quantum Channel Discrimination. *Phys. Rev. Lett.,* **125**(8), 080505.

[87] Mukamel, S., Freyberger, M., Schleich, W., Bellini, M., Zavatta, A., Leuchs, G., Silberhorn, C., Boyd, R. W., Sánchez-Soto, L. L., Stefanov, A., Barbieri, M., Paterova, A., Krivitsky, L., Shwartz, S., Tamasaku, K., Dorfman, K., Schlawin, F., Sandoghdar, V., Raymer, M., Marcus, A., Varnavski, O., Goodson, T., Zhou, Z.-Y., Shi, B.-S., Asban, S., Scully, M., Agarwal, G., Peng, T., Sokolov, A. V., Zhang, Z.-D., Zubairy, M. S., Vartanyants, I. A., del Valle, E., and Laussy, F. (March, 2020) Roadmap on Quantum Light Spectroscopy. *J. Phys. B,* **53**(7), 072002.

[88] Zhuang, Q. and Pirandola, S. (June, 2020) Entanglement-Enhanced Testing of Multiple Quantum Hypotheses. *Commun. Phys.,* **3**(1), 103.

[89] Sudarshan, E. C. G. (April, 1963) Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams. *Phys. Rev. Lett.,* **10**(7), 277–279.

[90] Glauber, R. J. (September, 1963) Coherent and Incoherent States of the Radiation Field. *Phys. Rev.,* **131**(6), 2766–2788.

[91] Boisselle, J. Port-Based Teleportation of Continuous Quantum Variables. Master's thesis University of Waterloo (2014).

[92] Stinespring, W. F. (1955) Positive Functions on C*-Algebras. *Proc. Am. Math. Soc.,* **6**(2), 211–216.

[93] Cho, Y.-W., Campbell, G. T., Everett, J. L., Bernu, J., Higginbottom, D. B., Cao, M. T., Geng, J., Robins, N. P., Lam, P. K., and Buchler, B. C. (January, 2016) Highly Efficient Optical Quantum Memory with Long Coherence Time in Cold Atoms. *Optica,* **3**(1), 100–107.

[94] Vernaz-Gris, P., Huang, K., Cao, M., Sheremet, A. S., and Laurat, J. (January, 2018) Highly-Efficient Quantum Memory for Polarization Qubits in a Spatially-Multiplexed Cold Atomic Ensemble. *Nat. Commun.,* **9**(1), 363.

[95] Wang, Y., Li, J., Zhang, S., Su, K., Zhou, Y., Liao, K., Du, S., Yan, H., and Zhu, S.-L. (May, 2019) Efficient Quantum Memory for Single-Photon Polarization Qubits. *Nat. Photonics,* **13**(5), 346–351.

[96] Montanaro, A. (2008) A Lower Bound on the Probability of Error in Quantum State Discrimination. In *2008 IEEE Information Theory Workshop* IEEE pp. 378–380.

[97] Barnum, H. and Knill, E. (2002) Reversing Quantum Dynamics with Near-Optimal Quantum and Classical Fidelity. *J. Math. Phys.,* **43**(5), 2097–2106.

[98] Serafini, A., Adesso, G., and Illuminati, F. (March, 2005) Unitarily Localizable Entanglement of Gaussian States. *Phys. Rev. A,* **71**(3), 032349.

[99] Choi, M.-D. (1975) Completely Positive Linear Maps on Complex Matrices. *Linear Algebra Its Appl.,* **10**, 285–290.

[100] Jamiołkowski, A. (1972) Linear Transformations Which Preserve Trace and Positive Semidefiniteness of Operators. *Rep. Math. Phys.,* **3**, 275–278.

[101] Belavkin, V. P. and Staszewski, P. (1986) Radon-Nikodym Theorem for Completely Positive Maps. *Rep. Math. Phys.,* **24**, 49–55.

[102] Marian, P. and Marian, T. A. (August, 2012) Uhlmann Fidelity between Two-Mode Gaussian States. *Phys. Rev. A,* **86**(2), 022340.

[103] Helstrom, C. W. (1976) Quantum Detection and Estimation Theory, Academic Press, USA.

[104] Siegman, A. E. (1986) Lasers, University Science Books, USA.

[105] Schützhold, R. (June, 2003) Pattern Recognition on a Quantum Computer. *Phys. Rev. A,* **67**(6), 062311.

[106] Schaller, G. and Schützhold, R. (July, 2006) Quantum Algorithm for Optical-Template Recognition with Noise Filtering. *Phys. Rev. A,* **74**(1), 012303.

[107] Bennett, C. H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A., and Wootters, W. K. (March, 1993) Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels. *Phys. Rev. Lett.,* **70**(13), 1895–1899.

[108] Pirandola, S., Eisert, J., Weedbrook, C., Furusawa, A., and Braunstein, S. L. (October, 2015) Advances in Quantum Teleportation. *Nat. Photonics,* **9**(10), 641–652.

[109] Watrous, J. (2018) The Theory of Quantum Information, Cambridge University Press, USA.

[110] Bengtsson, I. and Zyczkowski, K. (2006) Geometry of Quantum States: An Introduction to Quantum Entanglement, Cambridge University Press, USA.

[111] Holevo, A. S. (July, 2019) Quantum Systems, Channels, Information, De Gruyter, Berlin, Boston.

[112] Andersen, U. L., Neergaard-Nielsen, J. S., van Loock, P., and Furusawa, A. (September, 2015) Hybrid Discrete- and Continuous-Variable Quantum Information. *Nat. Phys.,* **11**(9), 713–719.

[113] Gottesman, D. and Chuang, I. L. (November, 1999) Demonstrating the Viability of Universal Quantum Computation Using Teleportation and Single-Qubit Operations. *Nature,* **402**(6760), 390–393.

[114] Pirandola, S. (May, 2019) End-to-End Capacities of a Quantum Communication Network. *Commun. Phys.,* **2**(1), 1–10.

[115] Pirandola, S. (September, 2019) Bounds for Multi-End Communication over Quantum Networks. *Quantum Sci. Technol.,* **4**(4), 045006.

[116] Christandl, M., Leditzky, F., Majenz, C., Smith, G., Speelman, F., and Walter, M. (January, 2021) Asymptotic Performance of Port-Based Teleportation. *Commun. Math. Phys.,* **381**(1), 379–451.

[117] Beigi, S. and König, R. (September, 2011) Simplified Instantaneous Non-Local Quantum Computation with Applications to Position-Based Cryptography. *New J. Phys.,* **13**(9), 093036.

[118] Buhrman, H., Chandran, N., Fehr, S., Gelles, R., Goyal, V., Ostrovsky, R., and Schaffner, C. (2011) Position-Based Quantum Cryptography: Impossibility and Constructions. In Rogaway, P., (ed.), *Advances in Cryptology – CRYPTO 2011*, Berlin, Heidelberg: Springer Berlin Heidelberg pp. 429–446.

[119] Bose, S. (November, 2003) Quantum Communication through an Unmodulated Spin Chain. *Phys. Rev. Lett.,* **91**(20), 207901.

[120] Adesso, G., Dell'Anno, F., De Siena, S., Illuminati, F., and Souza, L. A. M. (April, 2009) Optimal Estimation of Losses at the Ultimate Quantum Limit with Non-Gaussian States. *Phys. Rev. A,* **79**(4), 040305.

[121] Braunstein, S. L., Lane, A. S., and Caves, C. M. (October, 1992) Maximum-Likelihood Analysis of Multiple Quantum Phase Measurements. *Phys. Rev. Lett.,* **69**(15), 2153–2156.

[122] Kim, Y. S., Jeong, Y. C., and Kim, Y. H. (June, 2008) Implementation of Polarization-Coded Free-Space BB84 Quantum Key Distribution. *Laser Phys.,* **18**(6), 810.

[123] Lucamarini, M., Choi, I., Ward, M. B., Dynes, J. F., Yuan, Z. L., and Shields, A. J. (September, 2015) Practical Security Bounds Against the Trojan-Horse Attack in Quantum Key Distribution. *Phys. Rev. X,* **5**(3), 031030.

[124] Hayashi, M. (2017) Quantum Information Theory: Mathematical Foundation, Springer-Verlag, Berlin Heidelberg second edition.

[125] Gisin, N., Ribordy, G., Tittel, W., and Zbinden, H. (March, 2002) Quantum Cryptography. *Rev. Mod. Phys.,* **74**(1), 145–195.

[126] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M. (September, 2009) The Security of Practical Quantum Key Distribution. *Rev. Mod. Phys.,* **81**(3), 1301–1350.

[127] Diamanti, E. and Leverrier, A. (September, 2015) Distributing Secret Keys with Quantum Continuous Variables: Principle, Security and Implementations. *Entropy,* **17**(9), 6072–6092.

[128] Pironio, S., Acin, A., Brunner, N., Gisin, N., Massar, S., and Scarani, V. (April, 2009) Device-Independent Quantum Key Distribution Secure against Collective Attacks. *New J. Phys,* **11**(4), 045021.

[129] Hwang, W.-Y. (August, 2003) Quantum Key Distribution with High Loss: Toward Global Secure Communication. *Phys. Rev. Lett.,* **91**(5), 057901.

[130] Lo, H.-K., Ma, X., and Chen, K. (June, 2005) Decoy State Quantum Key Distribution. *Phys. Rev. Lett.,* **94**(23), 230504.

[131] Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., and Makarov, V. (October, 2010) Hacking Commercial Quantum Cryptography Systems by Tailored Bright Illumination. *Nat. Photonics,* **4**(10), 686–689.

[132] Zhao, Y., Fung, C.-H. F., Qi, B., Chen, C., and Lo, H.-K. (October, 2008) Quantum Hacking: Experimental Demonstration of Time-Shift Attack against Practical Quantum-Key-Distribution Systems. *Phys. Rev. A,* **78**(4), 042333.

[133] Jain, N., Anisimova, E., Khan, I., Makarov, V., Marquardt, C., and Leuchs, G. (2014) Trojan-Horse Attacks Threaten the Security of Practical Quantum Cryptography. *New J. Phys,* **16**(12), 123030.

[134] Häseler, H., Moroder, T., and Lütkenhaus, N. (March, 2008) Testing Quantum Devices: Practical Entanglement Verification in Bipartite Optical Systems. *Phys. Rev. A,* **77**(3), 032303.

[135] Huang, J.-Z., Weedbrook, C., Yin, Z.-Q., Wang, S., Li, H.-W., Chen, W., Guo, G.-C., and Han, Z.-F. (June, 2013) Quantum Hacking of a Continuous-Variable Quantum-Key-Distribution System Using a Wavelength Attack. *Phys. Rev. A,* **87**(6), 062329.

[136] Ma, X.-C., Sun, S.-H., Jiang, M.-S., and Liang, L.-M. (May, 2013) Wavelength Attack on Practical Continuous-Variable Quantum-Key-Distribution System with a Heterodyne Protocol. *Phys. Rev. A,* **87**(5), 052309.

[137] Jouguet, P., Kunz-Jacques, S., and Diamanti, E. (June, 2013) Preventing Calibration Attacks on the Local Oscillator in Continuous-Variable Quantum Key Distribution. *Phys. Rev. A,* **87**(6), 062313.

[138] Huang, J.-Z., Kunz-Jacques, S., Jouguet, P., Weedbrook, C., Yin, Z.-Q., Wang, S., Chen, W., Guo, G.-C., and Han, Z.-F. (March, 2014) Quantum Hacking on Quantum Key Distribution Using Homodyne Detection. *Phys. Rev. A,* **89**(3), 032304.

[139] Qin, H., Kumar, R., and Alléaume, R. (July, 2016) Quantum Hacking: Saturation Attack on Practical Continuous-Variable Quantum Key Distribution. *Phys. Rev. A,* **94**(1).

[140] Qin, H., Kumar, R., Makarov, V., and Alléaume, R. (July, 2018) Homodyne-Detector-Blinding Attack in Continuous-Variable Quantum Key Distribution. *Phys. Rev. A,* **98**(1), 012312.

[141] Ekert, A. K. (August, 1991) Quantum Cryptography Based on Bell's Theorem. *Phys. Rev. Lett.,* **67**(6), 661–663.

[142] Braunstein, S. L. and Pirandola, S. (March, 2012) Side-Channel-Free Quantum Key Distribution. *Phys. Rev. Lett.,* **108**(13), 130502.

[143] Lo, H.-K., Curty, M., and Qi, B. (March, 2012) Measurement-Device-Independent Quantum Key Distribution. *Phys. Rev. Lett.,* **108**(13), 130503.

[144] Pirandola, S., Ottaviani, C., Spedalieri, G., Weedbrook, C., Braunstein, S. L., Lloyd, S., Gehring, T., Jacobsen, C. S., and Andersen, U. L. (2015) High-Rate Measurement-Device-Independent Quantum Cryptography. *Nat. Photonics,* **9**(6), 397–402.

[145] Vakhitov, A., Makarov, V., and Hjelme, D. R. (November, 2001) Large Pulse Attack as a Method of Conventional Optical Eavesdropping in Quantum Cryptography. *J. Mod. Opt.,* **48**(13), 2023–2038.

[146] Vinay, S. and Kok, P. (April, 2018) Burning the Trojan Horse: Defending against Side-Channel Attacks in QKD. *Phys. Rev. A,* **97**(4).

[147] Gisin, N., Fasel, S., Kraus, B., Zbinden, H., and Ribordy, G. (February, 2006) Trojan-Horse Attacks on Quantum-Key-Distribution Systems. *Phys. Rev. A,* **73**(2), 022320.

[148] Tamaki, K., Curty, M., and Lucamarini, M. (2016) Decoy-State Quantum Key Distribution with a Leaky Source. *New J. Phys,* **18**(6), 065008.

[149] Jain, N., Stiller, B., Khan, I., Makarov, V., Marquardt, C., and Leuchs, G. (May, 2015) Risk Analysis of Trojan-Horse Attacks on Practical Quantum Key Distribution Systems. *IEEE J. Sel. Top. Quantum Electron.,* **21**(3), 168–177.

[150] Grosshans, F., Cerf, N. J., Wenger, J., Tualle-Brouri, R., and Grangier, P. (October, 2003) Virtual Entanglement and Reconciliation Protocols for Quantum Cryptography with Continuous Variables. *Quantum Inf. Comput.,* **3**(7), 535–552.

[151] Usenko, V. C. and Filip, R. (January, 2016) Trusted Noise in Continuous-Variable Quantum Key Distribution: A Threat and a Defense. *Entropy,* **18**(1), 20.

[152] Derkach, I., Usenko, V. C., and Filip, R. (December, 2017) Continuous-Variable Quantum Key Distribution with a Leakage from State Preparation. *Phys. Rev. A,* **96**(6), 062309.

[153] Derkach, I., Usenko, V. C., and Filip, R. (March, 2016) Preventing Side-Channel Effects in Continuous-Variable Quantum Key Distribution. *Phys. Rev. A,* **93**(3), 032309.

[154] Grosshans, F. (January, 2005) Collective Attacks and Unconditional Security in Continuous Variable Quantum Key Distribution. *Phys. Rev. Lett.,* **94**(2).

[155] Paris, M. G. A. (January, 2009) Quantum Estimation for Quantum Technology. *Int. J. Quantum Inf.,* **07**(supp01), 125–137.

[156] Braun, D., Adesso, G., Benatti, F., Floreanini, R., Marzolino, U., Mitchell, M. W., and Pirandola, S. (September, 2018) Quantum-Enhanced Measurements without Entanglement. *Rev. Mod. Phys.,* **90**(3), 035006.