UNIVERSITY OF SHEFFIELD

DOCTORAL THESIS

The Statistics and Security of Quantum Key Distribution

Author:
Scott E. Vinay

Supervisors:
Dr. Pieter Kok,
Dr. Stefano Pirandola

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Low-Dimensional Semiconductor Devices group
Department of Physics and Astronomy

December 2018

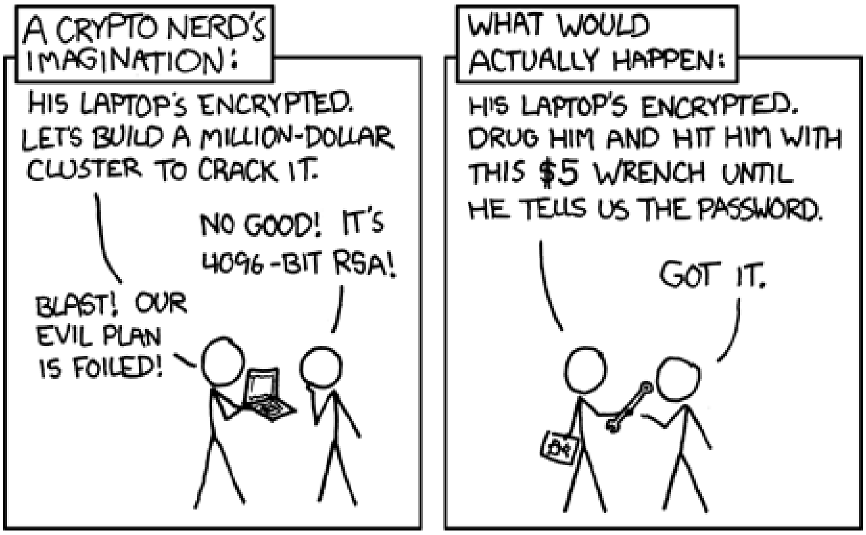
Declaration of Authorship

I, Scott Vinay, declare that this thesis titled, The Statistics and Security of Quantum Key Distribution, and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:


A CRYPTO NERD'S IMAGINATION:
HIS LAPTOP'S ENCRYPTED. LET'S BUILD A MILLION-DOLLAR CLUSTER TO CRACK IT.
NO GOOD! IT'S 4096-BIT RSA!
BLAST! OUR EVIL PLAN IS FOILED!
WHAT WOULD ACTUALLY HAPPEN:
HIS LAPTOP'S ENCRYPTED. DRUG HIM AND HIT HIM WITH THIS $5 WRENCH UNTIL HE TELLS US THE PASSWORD.
GOT IT.

https://xkcd.com/538/, this work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License.

There once was a lass named Alice,
who's friend, Bob, lived over in Dallas.
They wanted to chat,
but they didn't know that,
sneaky Eve was listening with malice...

Abstract

The Statistics and Security of Quantum Key Distribution

by Scott E. Vinay

In this work our aim has been to elucidate our theoretical developments that bolster the efficiency of quantum key distribution systems leading to more secure communication channels, as well as develop rigorous methods for their analysis. After a review of the necessary mathematical and physical preliminaries and a discussion of the present state of quantum communication technologies, we begin by investigating the Trojan Horse Attack, a form of side-channel attack that could threaten the security of existing key distribution protocols. We examine the secret key rates that may be achieved when an eavesdropper may use any Gaussian state in the presence of thermal noise, and prove that the coherent state is optimal in this case. We then allow the eavesdropper to use any separable state, and show that this gives a key rate bound close to that of the coherent state.

We develop a protocol for a quantum repeater that makes use of the double-heralding procedure for entanglement-generation. In our analysis, we include statistical effects on the key rate arising from probabilistic entanglement generation, which results in some quantum memories decohering while other sections complete their entanglement generation attempts. We show that this results in secure communication being possible over thousands of kilometres, allowing for intercontinental key distribution.

Finally, we investigate in more depth the statistical issues that arise in general quantum repeater networks. We develop a framework based on Markov chains and probability generating functions, to show how one may easily calculate an analytic expression for the completion time of a probabilistic process. We then extend this method to show how one may track the distribution of the number of errors that accrue in operating such a process. We apply these methods to a typical quantum repeater network to get new tight bounds on the achievable key rates.

Publications and conferences

Throughout the course of this doctorate, some of the work has been published and presented in journals and conferences.

Published works

Scott E. Vinay and Pieter Kok. Practical repeaters for ultralong-distance quantum communication. *Physical Review A*, 95(5), 2017.
(Mostly discussed in Chapter 5.)

Scott E. Vinay and Pieter Kok. Extended analysis of the Trojan-horse attack in quantum key distribution. *Physical Review A*, 97(4), 2018.
(Mostly discussed in Chapter 4.)

Scott E. Vinay and Pieter Kok. Statistical analysis of quantum entangled network generation." *Physical Review A*, 99(4), 2019.
(Mostly discussed in Chapter 6.)

Conferences

Quantum Roundabout 6th – 8th July 2016, Nottingham, UK. Poster presented: Highly loss-tolerant quantum repeaters.

SPIE West 28th January – 2nd February 2017, San Francisco, USA. Talk by Pieter Kok: Practical repeaters for ultralong-distance quantum communication.

Winter school on Complex Networks: From Classical to Quantum 3rd – 7th April 2017, Obergurgl, Austria. Poster presented: Highly loss-tolerant quantum repeaters.

QCrypt 18th – 22nd September 2017, Cambridge, UK. Poster presented: The Trojan Horse Attack against QKD.

Acknowledgements

The extent to which I am indebted to so many of the people in my life for supporting, encouraging, entertaining and uplifting me these past few years, in both an academic and a personal context, is surely too great to be expressed on this page, or indeed in this entire volume. I would, however, like to indicate a particular gratitude to a few individuals.

Firstly, to my family, and especially my parents. I certainly would not have reached this stage without the incredible support and unconditional love that you have shown to me. Your unwavering selflessness is a true inspiration to me, and is a model of personhood to which I aspire.

I wish to deeply thank everyone who has helped to make my time in Sheffield an immensely enjoyable one, and for standing by my side in all occasions, whether joyous or trying. A particularly notable proper subset of this may be written as { David Hurst, Earl Campbell, Giuseppe Buonaiuto, Jasminder Sidhu, Luke Heyfron, Mark Howard, Mark Pierce, Mike Roche, Sam Coveny, Zixin Huang, and all comrades of the IMT }. My deepest love goes out also to all my other friends. Whether of Derby, Sheffield, or elsewhere in the world, those I have known a few weeks, to those I have known for many years, whether entering my life or leaving it: my life would contain but a fraction of its happiness were it not for the rich fabric of joy that you have all woven throughout it.

And last but certainly not least, I wish to extend an appreciation of the most inestimable order to my supervisor, Pieter Kok. It is hard to imagine a person with more unshakable optimism and faith in his students, which did not falter even when I doubted myself entirely. For a busy person to consistently find so much time for his students is a rare thing indeed. I have unquestionably grown as a researcher during my time here, and our discussions on the philosophy of quantum mechanics have shaped my thinking of the world.

It has been an honour to be a member of this research group, and a student at this university. I dedicate this work to all those who are mentioned here, and to all those who are not.

Contents

List of Abbreviations and Symbols

QKD	**Q**uantum **K**ey **D**istribution
BB84	**B**ennet and **B**rassard 19**84**
E91	**E**kert 19**91**
SCA	**S**ide **C**hannel **A**ttack
THA	**T**rojan **H**orse **A**ttack
CHSH	**C**lauser **H**orne **S**himony **H**olt inequality
PGF	**P**robability **G**enerating **F**unction

Alice	Legitimate QKD user (sender)
Bob	Legitimate QKD user (receiver)
Eve	Eavesdropper
$\|\psi\rangle_i$	General pure state on mode(s) i (mode omitted when clear)
ρ_i	General mixed state on mode(s) i
$\|\Phi^{\pm}\rangle, \|\Psi^{\pm}\rangle$	Bell state
$h_2(\cdot)$	Binary entropy function
K	Secret key rate
ϵ	Error rate
N_S	Number of repeater sections
L_0	Inter-repeater distance
L_{att}	Channel attenuation distance
η	$\eta_T \times \eta_D$
η_T, η_D	Channel transmissivity, detector efficiency
Δ	Distinguishability of qubits by Eve
$\mathcal{P}$	General Markov process
T	Random variable describing the completion time of $\mathcal{P}$
p_t	Probability that $T = t$
p_c	Probability of success of single entanglement generation attempt
q	Number of parallel channels between repeater stations

Chapter 1

Introduction

1.1 The need for a new kind of communication

Quantum mechanics has long had a reputation as a more acute form of the image of theoretical physics as a whole; as an esoteric and "pure" science. Quantum *communication*, on the other hand, is the result of the intersection of a pure science with the distrust inherent in the real world. The germination of such an idea was the invention of quantum money in 1970 ([Wie83], published 1983), whereby a currency can be made from a quantum state such that it cannot be forged. This did not get much attention at the time, and it was not until the conception of *quantum computers* in a 1982 talk by Richard Feynman [Fey82] that it was realised that such ideas were not simply interesting, but a necessary component of the quantum-technological landscape.

In his talk Feynman suggested the possibility of a radically new type of computer, where the bits, logic gates and read-outs were replaced by quantum states, unitary operations and measurement operators. This was originally proposed for the purpose of simulating quantum systems; a task that soon becomes impractical for a classical computer, due to the fact that the number of variables needed to simulate an n-bit quantum state scales exponentially with n. If you want to accurately simulate a large quantum state, then surely the best way is to build a large quantum state yourself!

Whilst the potential that quantum computers hold for simulation purposes is important, and subject to a large and active area of research, it is arguably not the reason why they have attracted so much attention. In 1994 Peter Shor devised the process now known as Shor's algorithm [Sho94]. This is a method by which a quantum computer can find the prime factors for some integer N in a number of operations that scales poly-logarithmically with N. Whilst this may sound like little more than an academic curiosity, it actually poses a significant threat to global cybersecurity due to the fact that prime factorisation underpins the security of the Rivest–Shamir–Adleman (RSA) system — a public-key cryptography scheme that is ubiquitously used for encoding messages sent over the Internet [RSA78]. All known methods for trying to crack RSA with classical computers take exponentially long in N, which means that any realistically long message encoded with RSA would take many millions of years to crack with a classical computer. However, this can be reduced to minutes or less with a sufficiently advanced quantum computer.

Of course, once a technology has been invented, even in concept, it cannot be uninvented, no matter how dangerous it is. The task instead fell to the academic community to devise new methods of cryptography that could not be broken, even by quantum computers. Progress towards this aim has proceeded primarily along two main paths. One of these is the field of *post-quantum cryptography* (PQC) [Ber09]. Proponents of PQC aim to replace RSA with different classical algorithms of sufficient mathematical complexity that the time required to break them scales exponentially with the message length, whether attacked by a classical or quantum computer. The main advantage of PQC schemes is that they can be run on all of the existing hardware and infrastructure that already exists — the wires and fibre-optic cables and so on that constitute the connections of the Internet. The disadvantage is that PQC schemes may be *computationally secure*, but they can never be *information-theoretically secure*.[1] This means that there is no guarantee that they will not be broken in the future, and we may be forever stuck in an arms-race between the code-makers and code-breakers.

The second avenue of investigation for quantum-safe encryption, and the focus of this thesis, is that of using keys shared by *quantum key distribution* (QKD), which, unlike PQC, *is* information-theoretically secure. This means that an eavesdropper, Eve, who only intercepts the signals sent between the two legitimate parties Alice and Bob will never be able to decode the message, even given unbounded computational power, providing the system was implemented without errors. QKD systems are of a fundamentally different character to classical encryption systems. They take advantage of the quantum properties of light, such as the non-comeasurability of conjugate variables of the optical state. This does, however, mean that an entirely new network, both of hardware and software, is required by all users.

Whilst the primary purpose of QKD is to encrypt data in the face of possible attacks from the computers of the future, this does not mean that we can consider it merely as an after-thought, or as something to only be dealt with if and when we manage to build functional quantum computers. Given the knowledge that quantum computers are likely to be technologically viable in the near future, nefarious agencies could intercept and store encrypted data, and retroactively decode it when quantum computers become available. Therefore, the development of robust, long-range, secure, loophole-free QKD systems has become a unifying project of utmost urgency for the global physics, computer science, and cryptography communities.

As such, the development of functional and efficient QKD has become a focus of concerted effort for a great number of researchers. Due to the problem lying at the intersection of multiple disciplines, these come from a diverse range of academic backgrounds, and have thus far made a great amount of headway in both the theoretical

[1] See the box in Section 3.

and experimental aspects. At the time of writing, prototype-stage systems have begun to be deployed by telecommunications companies such as Toshiba [SFI+11]. The excitement around the promise of such rigorous security has lead to the conceptual ideas behind QKD branching out to other fields, with some hoping that a quantum-secured blockchain [KPA+18] will prove to be the future of a decentralised monetary system.

However, this does not mean that quantum cryptography is a closed case. Although it is provably secure against the type of attacks to which classical cryptography is vulnerable, it carries with it its own set of challenges. The fact that many protocols involve information transfer via individual photons of light means that they are highly susceptible to attenuation within the optical fibres in a way that classical communication is not. Additionally, the introduction of the new hardware necessary for the operation of QKD protocols opens up entirely new vectors for attack, which must be carefully defended against. In this thesis we introduce original theoretical developments to help address these shortcomings, to aid in the task of the physics community of bringing to fruition a new era in secure communication.

1.2 The structure of this thesis

We begin in Chapter 2 with an overview of some of the fundamentals of quantum mechanics. The concepts introduced here have applicability in many areas of quantum physics research. However, we have focused primarily on aspects which are particularly relevant to the fields of quantum information theory and cryptography. This includes the states most commonly used to describe information encoded in photons, such as Fock states, Gaussian states, and qubits, as well as metrics by which we may compare them. We also introduce the concept of entanglement, and look at how it may be constructed, manipulated, quantified, and used in key distribution. We also briefly discuss some of the implications that quantum mechanics, and entanglement in particular, has for the way in which we think about the fundamental nature of the world.

In Chapter 3 we continue to discuss important known concepts, but here we narrow our focus to the field of quantum communication in particular. We discuss two of the most common QKD protocols: BB84 and E91, and show how they lead to a security that is not conditional on an adversary's computational abilities. We show how the details of a certain protocol allow us to calculate a bound on the rate at which it may be used to securely communicate. It is here that we introduce the important concepts of side-channels and quantum repeaters. These are, respectively, manners by which Eve may attempt to learn the key without directly measuring in-transit information, and manners by which the range of quantum communication by be extended beyond the limit imposed by the attenuation of a single optical fibre. It is these aspects of QKD that we focus on with our original contributions, so we spend some time discussing how they affect the operation of a protocol. In particular, we discuss a few specific repeater protocols, focussing on a prototypical example known as the Innsbruck protocol. We also discuss a protocol known as double-heralding, which is a powerful method for probabilistically generating entanglement — a concept that we focus on in Chapters 5 and 6.

In Chapter 4 we address the issue of side-channels within BB84. We specifically investigate an threat known as the Trojan-Horse Attack, which involves Eve sending her own signal into Alice's apparatus and measuring the reflected state. Due to the fact that this may partially encode the information as to which settings Alice has used to encode her qubit, this presents a threat to the security of the protocol. We investigate the optimal states for Eve to use, and we consider some novel defences that Alice may implement against such attacks.

Whilst a quantum communication protocol may be more secure if shown to be safe against side-channel attacks, it is of limited practical utility if it cannot then realise this powerful security for communication over long ranges. Therefore, in Chapter 5 we present our design for a quantum repeater, including descriptions of how all elements of the protocol may be realised. In particular, we focus on using double-heralding as the primitive mode of communication between stations. We give a complete and detailed analysis of this, deriving bounds on the rates at which the end users may securely communicate. In particular we investigate the issue of timings, whereby different sections of a repeater network will complete at different times due to their probabilistic nature.

Upon developing the aforementioned protocol, it became clear that the issue of the unequal times at which repeater section will connect has the potential to be a major issue in the construction of a network. We also found it to be the case that a simple analysis that only considers average values for various quantities would fail to capture the most interesting behaviour. In Chapter 6 we describe and construct a full mathematical apparatus for analysing such systems. Based on the concept of Markov chains, we find results for detailed probability distributions for how long protocols will take to complete, as well as the number of errors that they will accumulate in doing so. We apply these results to the Innsbruck protocol for secure communication, giving an analysis that captures a far richer behaviour than would be otherwise possible.

Chapter 2

Quantum mechanics preliminaries

A good understanding of quantum communication and cryptography must build upon a good understanding of quantum mechanics in general. Here we discuss the physical and mathematical structures required to give an accurate description of a quantum-mechanical system.

2.1 Quantum states

A quantum state is the fundamental starting point of all quantum mechanical calculations. It is the conception of systems as quantum states that has lead us away from the safe and sensible world of classical Newtonian mechanics, and into the metaphysically uncomfortable but physically undeniable realm of quantum phenomena. The nature of quantum states is a vast area of research in and of itself, and is possibly the most fundamental question in quantum research. In this section, we will primarily be describing their mathematical formulation, without delving too much into understanding how these formulations were developed over many decades.

Quantum states may be considered in two distinct but related forms: pure and mixed, dependent on whether they embody only quantum uncertainty, or a combination of quantum and classical uncertainty respectively.

2.1.1 Pure states

The question of how best to define a quantum state is one with many different answers. Ultimately, we want to arrive at a vector space, and a set of permissible operations upon that space, which accurately describes the dynamics of the system. However, it is not immediately clear how such a space should be constructed. To answer this, we shall take the concept of a *observable* to be the primitive object from which we construct the set of state vectors, rather than the other way around which is sometimes more common.

We choose to do this from a physically-motivated perspective. Physics is ultimately an experimental investigation of Nature. Theory, in this case, should always be constructed to explain the results of measurements and to inform predictions. As such, it seems to us to be most logical to define our mathematical constructs with respect to our experimental outcomes.

A pure state of a quantum system may therefore be defined with respect to a set of measurements that may be made upon it, along with a partitioning into a number of *modes*. A mode is a variable which may, in a classical sense, allow for distinguishability. For example, two particles in two separate locations may be considered to be understood by quantum states on separate modes.

Suppose we have some set of observable,[1] $\hat{M}_1$, $\hat{M}_2$, $\hat{M}_3$, $\cdots$, $\hat{M}_n$, describing observables that we wish to measure on n modes. For each such observable, there are a set of eigenstates — these are the states of the mode that are not changed after an application of $\hat{M}$. For example, we may have a system that comprises a single photon that exists in a superposition of up to n different frequencies (modes), where $\hat{M}_i$ measures the polarisation of frequency i as either left or right.

Mathematically, a observable is some operation that acts on some space. These may, for example, be differential operators acting on a space of functions. However, in quantum information theory, the observables are typically described as matrices. An observable operator $\hat{M}_i$ that is used to *define* a space can be written as a $d \times d$ diagonal vector, with diagonal elements m_i^1, m_i^2, $\cdots$, m_i^d corresponding to the results of the measurement of the observable,[2] where d is the number of possible measurement outcomes. This "defines" the space in the sense that we may now talk about the set of all convex complex combinations of the eigenvectors of this operator. We may then use this space to talk about the results of other operations (which, when constructed as matrices in this same basis, will in general be non-diagonal).

The quantum state for a single mode, i, may then initially be thought of as an element of the complex vector space, $\mathbb{C}^d$. In this formulation, each unique basis vector corresponds to an eigenstate of the defining observable. Other operators may then be uniquely defined on that space by their set of eigenvectors. Such quantum state vectors are typically notated as a *ket*:

$$\text{state} = |\text{label}\rangle . \tag{2.1}$$

The corresponding *bra*, $\langle\text{label}|$, denotes the complex-conjugated transpose of $|\text{label}\rangle$. However, the overall complex phase of a state cannot be measured. To see this, note that the expectation of an observable $\hat{M}$ with respect to a state $|\psi\rangle$ is given by $\langle\psi| \hat{M} |\psi\rangle$, so the phases will cancel out. Therefore, the space for a single mode is the *projectivisation* of the complex space:

$$P(\mathbb{C}^d) = \mathbb{CP}^{d-1} \tag{2.2}$$

[1] These may be written as a linear sum of idempotent operators, with well-defined eigenstates. Physically, this means that there exists states which, when operated upon by the observable, do not change (i.e. measurement statistics for all other operators on such a state do not change). In practical situations, these may be very difficult to construct. For example, almost any measurement that is made on a photon involves the absorption, and therefore destruction, of the photon. In these circumstances, measurements are more accurately described by *Positive Operator-Valued Measures* (POVMs). However, these details are not important for the task of defining the Hilbert space.

[2] For example, the outcomes of the measurement of the polarisation of a photon may be assigned to the two values of -1 and $+1$.

The construction $\mathbb{CP}^{d-1}$ is known as a *complex projective space*, and equal to the equivalence class

$$\{\mathbf{z} \in \mathbb{C}^d | \mathbf{z} \sim \lambda \mathbf{z}, \lambda \in \mathbb{C}\}. \tag{2.3}$$

In the general physics literature, such spaces are typically referred to by the more general name of Hilbert spaces. As such, we shall denote $\mathbb{CP}^{d-1}$ by $\mathcal{H}_d$, where the use of d rather than $d-1$ highlights that this is a d-dimensional space.

The Hilbert space for many modes may then be built up from these smaller modes, as

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_{d_1} \otimes \mathcal{H}_{d_2} \otimes \cdots \otimes \mathcal{H}_{d_n}, \\ &= \mathcal{H}_{d_1 \times d_2 \times \cdots \times d_n}. \end{aligned} \tag{2.4}$$

The states on these composite spaces may then be written as:

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle_{1,2,\cdots,n}, \tag{2.5}$$

or, for short:

$$|\psi\rangle = |\psi_1, \psi_2, \cdots, \psi_n\rangle_{1,2,\cdots,n}. \tag{2.6}$$

where the subscripts indicate the mode labels. Throughout, we shall use $|\psi\rangle$ to indicate a general quantum state.[3]

One can see from this that the dimensionality of the Hilbert space grows exponentially with the number of modes. It is this fact that underpins the power of quantum computers, and their key difference from classical computers. Both forms of computer are Turing-equivalent [Deu85], which means that any program that is run on one can, in principle, be run on the other. However, to fully emulate a quantum computer's algorithms on a classical computer will take an infeasible length of time. A linear increase in the resources required for a quantum algorithm will necessitate an exponential increase in the resources required for the classical equivalent.

Observables

Suppose we wish to perform some measurement of an observable, $\hat{M}$, on our state. With $\hat{M}$ written as a Hermitian matrix operator[4] on $\mathcal{H}_d$, it may always be decomposed into a set of projectors onto orthogonal eigenstates. If we let $|e_i\rangle$ be the eigenstates of $\hat{M}$ with eigenvalues E_i respectively, then we may write

$$\hat{M} = \sum_{i=1}^{d} E_i |e_i\rangle\langle e_i|. \tag{2.7}$$

[3] We may omit the commas between modes for brevity and clarity where appropriate.

[4] A Hermitian matrix is one for which $\hat{M} = \hat{M}^\dagger$, where $\hat{M}^\dagger$ is the complex-conjugated transpose of $\hat{M}$.

In measuring $\hat{M}$, we always project onto one of the eigenstates, returning a measurement result of the corresponding eigenvalue. However, this is done probabilistically. If some state $|\psi\rangle$ is operated on by $\hat{M}$, then the result E_i is returned with probability

$$p(E_i) = |\langle e_i|\psi\rangle|^2 . \tag{2.8}$$

This implies that we can not deterministically predict the outcome of a measurement, unless the state vector happens to be equal to one of the observable's eigenstates. This uncertainty appears to be fundamental. No matter how well we improve our measurement apparatus, there will always be a result that is unknowable, even in principle, until the experiment is actually carried out. We call this the inherent *quantum uncertainty* in the state. The calculations of a quantum physicist, therefore, will not typically ask *What result will we get?*, but instead ask *What is the probability distribution of results that we will get?*[5]

2.1.2 Mixed states

A pure state vector and the matrix of an observable can fully characterise the quantum uncertainty in a system. However, this is not the only kind of uncertainty that we may encounter. I am uncertain as to what the weather will be like tomorrow. That does not imply that the weather is dependent on some superposed quantum state and its fundamental unknowability, it just means that I do not have the capability to fully calculate the future behaviour of such a complex, albeit classical, system. However, uncertainty may be even simpler still. For example, I do not know what the weather was *yesterday* in Paris. No future calculations are necessary to fix the value of yesterday's weather, it has already happened. It is just the case that I do not know what it was. We wish for our mathematical apparatus to take such classical uncertainties into account, on an equal yet distinct footing to the quantum uncertainties, but as a single unified mathematical entity.

This is particularly important since we are generally interested in calculating the *probabilities* of things. Suppose we ask "What is the probability that I get measurement result X when I measure state Y?". It would not do for us to say, "Well this thing has a quantum probability of p_Q and a classical probability of p_C." No, in an actual experiment, result X either happens or it does not. And when we repeat the measurement many times, the fraction for which we measure X is given by a single number, its probability, not as a quantum and classical pair. However, we still wish to preserve the *quantumness* of quantum uncertainty. The fact that quantum probabilities are related to complex amplitudes and not real numbers allows for quantum

[5] This carries over to the realm of quantum computing, and the characterisation of their power in terms of complexity classes. For a classical computer, the set P is the set of all problems that may be solved in a number of operations that scales polynomially with the size of the problem. For a quantum computer, the corresponding set is BQP, or bounded-error quantum polynomial [NC02]. It is the set of problems for which the likelihood that a quantum algorithm that completes in a polynomial number of steps will get the answer wrong is less than some ϵ that does not vary with problem size.

states to interfere with each other, and give rise to statistical outcomes that are impossible with classical mechanics. In Section 2.2 we discuss this in more detail.

The problem of finding such a mathematical construct that satisfies such subtle probabilistic requirements was solved in 1927 by John von Neuman [VN27] with the introduction of a *mixed state*. When mixed, the status of a state is elevated from a vector on $\mathcal{H}_d$ to a $d \times d$ matrix, on equal standing now with its operators.

Suppose we are given some quantum mode or modes. We may believe that the quantum state on these modes is $|\psi_1\rangle$ with probability p_1, $|\psi_2\rangle$ with probability p_2, and so on up to $|\psi_n\rangle$, where $\sum_{i=1}^n p_i = 1$.[6] Our state may be now written as a *density matrix*:

$$\rho = \sum_{i=1}^{n} p_i \, |\psi_i\rangle \langle \psi_i| \, . \tag{2.9}$$

Note that $|\psi_i\rangle$ do not need to be mutually orthogonal. However, by construction, ρ is Hermitian, meaning it can be diagonalised. Therefore, we can always re-express ρ as a new set of probabilities over some orthogonal basis. This highlights an important fact about density matrices: they do not uniquely identify the states from which they are built. Once we introduce classical uncertainty, that information is lost.[7] As an example, suppose Alice and Bob got together and examined a state ρ. Alice might say: "This state has a 50% chance of being $|\psi\rangle$, and a 50% chance of being $|\phi\rangle$." Bob might say, "No, this state has a 50% chance of being $(|\psi\rangle + |\phi\rangle)/\sqrt{2}$, and a 50% chance of being $(|\psi\rangle - |\phi\rangle)/\sqrt{2}$!" In fact, they would both be correct. Both distributions over states give the same density matrix, and no experiment could be performed to distinguish between these two hypotheses.

Given a multi-partite state across modes held between, say, Alice and Bob, we may also talk about the local view of the state. This is the mixed state that correctly describes the probabilities of outcomes of measurements that Alice will find when she performs operations on her part of the state, whilst being ignorant to any measurements or operations that Bob is performing.

If we write some state, ρ, as

$$\rho = \sum_i p_i \, \sigma_{A,i} \otimes \sigma_{B,i}, \tag{2.10}$$

where $\sigma_{A,i}$ and $\sigma_{B,i}$ are positive semi-definite operators on Alice's and Bob's subspaces respectively, then Alice's local state may be written as

$$\rho_A = \sum_i p_i \, \mathrm{Tr}[\sigma_{B,i}] \, \sigma_{A,i}. \tag{2.11}$$

[6] Note how these are true probabilities, and not probability *amplitudes*, as they were for the coefficients of a quantum state.

[7] In contrast, consider a pure state, $|\psi\rangle$. If this is superposed with another pure state $|\phi\rangle$, no information is lost. An examination of the mathematical form of the combined state reveals uniquely that $|\psi\rangle$ and $|\phi\rangle$ were its constituent parts.

This construction gives the state that represents Alice's best description of the system given the incomplete information available to her.

2.1.3 Discrete-variable quantum states

All quantum systems that we consider in this thesis are ones that may be described by a countable number of modes, either finite or infinite.[8] However, within such states, we can still talk about *discrete* and *continuous* states, but where this label applies to the kinds of variables that we use to characterise the states.

Fock states

One of the most fundamental types of quantum state is the Fock state. The Hilbert space comprising the set of all Fock states is accordingly known as the Fock space. Here, the Hilbert space of a single mode has an infinite number of basis states, denoted by $|0\rangle, |1\rangle, |2\rangle, \cdots$. The form $|k\rangle$ represents the state with precisely k photons in the mode. The ability to lump together these k photons into a single mode necessarily means they are indistinguishable in every regard, and any distinguishability (such as photons existing in different places or times) should be represented as occupations of different modes.

An operator that defines this space is the number operator,

$$\hat{n} = \sum_{n=0}^{\infty} n\, |n\rangle \langle n| \,. \tag{2.12}$$

Of course, we cannot actually write this (or any other operator that acts on all number states) out fully in the number basis, since the space is infinite dimensional. Instead, we often use *construction operators*, which, for fixed mode, k, may be divided into creation operators, $\hat{a}_k^\dagger$, and annihilation operators, $\hat{a}_k$. These act on the Fock states as follows:

$$\begin{aligned} \hat{a}_k^\dagger |n\rangle_k &= \sqrt{n+1}\, |n+1\rangle_k \,, \\ \hat{a}_k |n\rangle_k &= \sqrt{n}\, |n-1\rangle_k \,. \end{aligned} \tag{2.13}$$

Note that operator $\hat{a}_k^\dagger$ only acts non-trivially on mode k, while acting as an identity operator on all other modes.

These are not Hermitian, so are not directly measurable, although they can be combined to form *any* operator on the Fock space. As a proof of this, consider that the number operator may be written as $\hat{n} = \hat{a}^\dagger \hat{a}$. The $|n\rangle\langle n|$ part of each term is idempotent, so

[8] Countable = Bijective with $\mathbb{N}$. For systems with an uncountably infinite number of modes, then Eq. 2.4 must be replaced with a more intricate construction, which we shall not elaborate on here.

$$(\hat{a}^\dagger \hat{a})^p = \hat{n}^p = \sum_{n=0}^{\infty} n^p \left|n\right\rangle \left\langle n\right| . \tag{2.14}$$

We may write a weighted sum over such operators as

$$\sum_{p=0}^{\infty} \alpha_p \hat{n}^p = \sum_{n=0}^{\infty} \left[\sum_{p=0}^{\infty} \alpha_p n^p \right] \left|n\right\rangle \left\langle n\right| , \tag{2.15}$$

for arbitrary constants $\{\alpha_p\}_p$. Let the quantity in the square brackets be written as $\Lambda(n)$. This is, by definition, simply an arbitrary analytic function of n. Suppose that the operator we wish to construct takes values $\{\beta_n | n = 0, 1, 2, \cdots\}$ on the diagonal. We then only need to choose $\{\alpha_p\}_p$ such that $\Lambda(n) = \beta_n$ on the non-negative integers, thus constructing an arbitrary operator that is diagonal in the Fock basis from creation and annihilation operators. To construct the offset diagonal that is m elements away from the diagonal, we perform the same procedure, but pre-multiply the end result by $\hat{a}^m / \sqrt{m!}$ or $\hat{a}^{\dagger m} / \sqrt{m!}$.

The Fock space is typically one of the first and most foundational concepts that a student of quantum optics will encounter. This fact can actually highlight key conceptual differences between the quantum physics of optics and of solid-state physics. In the latter, the number of particles mostly remains fixed. We do not, at typical experimental temperatures, see the presence of entire atoms fluctuating in and out of existence. We may therefore, for example, point to a proton, and ask "what is the state of this proton?" This stands in contrast to the bosonic case, where one cannot point unambiguously to a photon, but only to the photonic field, which can contain a superposition of different numbers of photons. In the solid-state case, then the different modes may represent the different distinct particles, and the states of each mode be the different states of such a particle. If these particles are indistinguishable, then this indistinguishability would be recaptured by symmetrising or anti-symmetrising the state, depending on whether they were bosons or fermions.

So we may have a state where, in a solid-state situation, we describe the modes as being different particles, and the states as different frequencies, whereas in the bosonic case we may have the different modes being different frequencies, and the different states being the number of excitations, or particles, in that mode. We may see that the two paradigms may almost be thought of as being complementary of each other. This duality in the expression of a multi-particle multi-mode quantum state has the potential to be confusing, so a careful study of the Fock space should be a priority for all new students.

This duality in fact highlights a major advantage of Fock states — they are extremely general. We may always recast a fixed Hilbert space of the solid-state form into a Fock space form, but the reverse is not always true.[9] The Fock-space allows for what is known as a dual-rail or multi-rail encoding of states, where the basis states are expanded out into seperate modes. For example, consider a photon that exists in an equal superposition of two locations, x_1 and x_2. Its state may be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|x_1\rangle + |x_2\rangle\right). \tag{2.16}$$

By representing these two locations as separate modes, we can write the state as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|1,0\rangle_{x_1,x_2} + |0,1\rangle_{x_1,x_2}\right). \tag{2.17}$$

Note here that what appears as superposition in Eq. 2.16 appears as *entanglement* in Eq. 2.17 (see Section 2.2). This is a first clue that different aspects of quantum weirdness that we can exploit to build quantum technologies, such as superpostion, entanglement, and non-commutivity of observables, may in fact be simply different facets of a single underlying quantum advantage. This is investigated in the setting of QKD in Section 3.1.2, and in the setting of quantum computation in the work of Howard et al. [HWVE14].

Qubits

The fundamental unit of a classical computer is a bit — a simple mathematical variable that takes on the value 0 or 1, realised in hardware by electrical signals. In a quantum computer, and indeed quantum technology more generally, these are typically replaced by *qubits*, which are elements of the 2-dimensional Hilbert space $\mathcal{H}_2$. The two canonical basis states (also known as the computational basis) are denoted $|0\rangle$ and $|1\rangle$, and these are the eigenstates of the Pauli Z matrix, which is given along with the Pauli X and Y matrices as:[10]

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.18}$$

A pure qubit may be represented as a two-element complex vector, $[\alpha, \beta]^T$, such that $|\alpha|^2 + |\beta|^2 = 1$. This set of co-ordinates may be represented as a point on an ordinary sphere.[11]

[9] Proof: Let $\mathcal{H}_S$ be a Hilbert space of the first kind, and $\mathcal{H}_F$ be a Fock space. Let $\mathcal{H}_S$ have n modes, and d eigenstates for each mode, so $\dim(\mathcal{H}_S) = d^n$. In the Fock space picture, each combination of eigenstates on different modes would become a mode, and the eigenstates within each subspace be different occupation numbers, of which there are infinite, giving $\dim(\mathcal{H}_F) = \infty^{dn}$. Therefore there always exists an injective but not bijective mapping from $\mathcal{H}_S \to \mathcal{H}_F$.

[10] Note that in some works, the Pauli X, Y, Z matrix is denoted $\sigma_{X,Y,Z}$. Here we use the former since there is no ambiguity with other variables.

[11] Mathematically, we can say that $\mathbb{CP}^1$ is diffeomorphic to the 2-sphere, $\mathbb{S}^2$ [Ken77].

The Pauli matrices find a wide array of uses within quantum information. In particular, we often want to perform the measurements corresponding to these operators. The Z operator measures in the computational basis given above, whilst X measures in the basis given by the states

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \\ |-\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right), \end{aligned} \tag{2.19}$$

which are states that we use throughout this thesis. Section 2.2.3 discusses some of the most interesting uses of these various measurement operators on complex entangled states.

Physically, qubits may be implemented by any quantum system that exhibits two degrees of freedom, or can be effectively limited to a two-dimensional subspace. However, it is a current matter of great research interest as to which physical form is most suitable for implementation in quantum technology. Current solid-state candidates include superconducting qubits [MCG+07, DS13], ion traps [CZ95, KMW02] and quantum dots [IAB+99, FRS+03, HK18]. Photons are used to carry the quantum information between distant parties, though there are multiple ways in which photons may be encoded here. They may be encoded in the polarisation [PBG+05] or the phase between different modes [LCW+15] (discussed further in Section 3.1.1). The qubit can even be encoded in occupation number of a mode, if this is restricted to be at most one. However, a problem with this final encoding is that one cannot distinguish between a qubit state $|0\rangle$ and a lost photon.

2.1.4 Continuous-variable quantum states

Frequently, quantum photonic experiments involve pulses with very large numbers of photons. Whilst a complete description of the corresponding pure state could be given by writing down the coefficients of each Fock state, but this would be very lengthy. Instead, they may be characterised by certain macroscopic continuous variables. Therefore, although they exist on the same discrete Hilbert space as the Fock states, we shall call these continuous-variable (CV) states.

When dealing with such states, it is common to talk about the *quadrature* operators, which are usually denoted position and momentum, given respectively by:[12]

$$\begin{aligned} \hat{x} &= \frac{1}{2}\left(\hat{a} + \hat{a}^{\dagger}\right), \\ \hat{p} &= \frac{-i}{2}\left(\hat{a} - \hat{a}^{\dagger}\right). \end{aligned} \tag{2.20}$$

[12] Defining $\hbar = 1$. Note that some authors replace the factor of $1/2$ with $1/\sqrt{2}$.

Note that these do not actually measure the position or momentum of the pulse,[13] but are called as such because they obey the same commutation relations:

$$[\hat{x}, \hat{p}] = \frac{i}{2}. \tag{2.21}$$

Coherent states

The coherent states are distributions over *all* Fock states, and are particularly notable as being the output of an ideal laser. They are described by a single complex number, α, and are given by:

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{2.22}$$

The vacuum state (Fock state with $n = 0$) may also be thought of as a coherent state with $\alpha = 0$. From here, we can produce a coherent state $|\alpha\rangle$ by applying the unitary *displacement* operator:

$$\hat{D}(\alpha) = \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a}). \tag{2.23}$$

More generally, the displacement can map from a coherent state to another one:

$$\hat{D}(\beta) |\alpha\rangle = |\alpha + \beta\rangle . \tag{2.24}$$

The phase may also be rotated by application of the rotation operator:

$$\hat{R}(\theta) |\alpha\rangle = \exp\left(i\theta \hat{a}^\dagger \hat{a}\right) |\alpha\rangle = |e^{i\theta}\alpha\rangle . \tag{2.25}$$

The coherent states are often called the most classical states, and there are a number of reasons for this. The first of these is their association with arguably the most classical system of all: the simple harmonic oscillator. The most basic Hamiltonian that might act on a single mode is one that simply counts the number of excitations in the mode and assigns each one the photonic energy $E = \omega$, described by $\hat{H} = \omega \hat{a}^\dagger \hat{a}$. The unitary which describes evolution under this Hamiltonian may be seen to be the rotation operator, $\hat{R}(\omega)$, which maps the coherent state to a coherent state with another time-dependent phase, $\alpha(t)$. The coherent state therefore maintains its structure under time-evolution. The position and momentum quadratures, $\langle \alpha(t)| \hat{x} |\alpha(t)\rangle$ and $\langle \alpha(t)| \hat{p} |\alpha(t)\rangle$, will similarly be described by $\alpha \cos(\omega\, t)$ and $\alpha \sin(\omega\, t)$ respectively, exactly as a mass on a spring.

In addition to the mean values of these operators, we can also look at their uncertainties, Δx and Δp, where for some state $|\psi\rangle$,

$$\Delta x = \sqrt{\langle \psi| \hat{x}^2 |\psi\rangle - \langle \psi| \hat{x} |\psi\rangle^2}, \tag{2.26}$$

[13] The momentum of a photon is given by $\hbar k$. Its position is not so well-defined, and problems arise due to its fully-relativistic speed [Pau12, Haw99].

and similar for Δp. For a coherent state, there is the additional specification that the uncertainty is evenly distributed between the two quadratures. Additionally, the coherent states have *minimal uncertainty*.[14] That is to say, they tightly satisfy Heisenberg's uncertainty principle:

$$\Delta x \Delta p \geq \frac{1}{2}. \tag{2.27}$$

Since such fundamental uncertainty is a characteristic of the quantum world, this adds to the justification of denoting them as highly-classical.

The coherent states are often visualised in the x, p phase space. They may be drawn as a circle of uncertainty. That is to say, it is a circle centred on the co-ordinates

$$\langle \psi | \hat{x} | \psi \rangle, \langle \psi | \hat{p} | \psi \rangle = \text{Re}(\alpha), \text{Im}(\alpha) \tag{2.28}$$

with a width and height of Δx and Δp respectively. This represents the region of the phase space for which the Wigner function of the coherent state is above some fixed value.[15] Of course, there are many boundary shapes with such a specified width and height, so how can we be sure that the coherent state is rotationally symmetric? To see this, note that the largest-area shape that has equal and fixed width and height is a square. We can then say that the state's *region of uncertainty*[16] is the intersection of the interior of the squares defined by the uncertainties of the rotated quadratures $\hat{x}_\theta$ and $\hat{p}_\theta$, given by:

$$\begin{aligned} \hat{x}_\theta &= \cos(\theta)\hat{x} + \sin(\theta)\hat{p}, \\ \hat{p}_\theta &= \cos(\theta)\hat{p} - \sin(\theta)\hat{x}. \end{aligned} \tag{2.29}$$

This is shown in Fig. 2.1. Here we illustrate the vacuum state at the origin of the phase space, which is displaced to form a coherent state, rotated, and then squeezed (squeezing is discussed in Section 2.1.4).

Another reason for which coherent states are known as the most classical states is their photon-counting statistics. If $|k\rangle$ is a pure Fock state, then

$$|\langle k|\alpha\rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}. \tag{2.30}$$

This is a Poisson distribution, with average photon number $|\alpha|^2$. A Poisson distribution describes the number of occurrences of a given event in some period when the events are independent of each other, and is widely used in classical statistics

[14] This is how the coherent states were first discovered, when Schrödinger was searching for such minimal-uncertainty states [Sch26]. They were later found to satisfy a different problem: when Glauber asked what state might have $g^{(n)}(t) = 1$ for all t and n, where $g^{(n)}(t)$ is the n^{th}-order coherence function, in his important work characterising the optical field; [Gla63].

[15] The Wigner function for a coherent state is given by $W(x,p) = \frac{2}{\pi} e^{(x^2+p^2)/2}$. If the circle of uncertainty is defined by Eq. 2.27, giving it a radius of $2^{-3/2}$, then the circle of uncertainty represents the area for which the Wigner function is greater than $\frac{2}{\pi} e^{-1/16}$).

[16] i.e. the area bounded by the circle of uncertainty

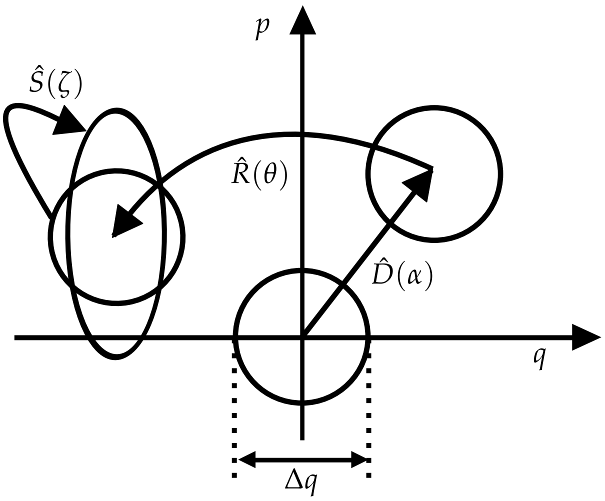


FIGURE 2.1: Illustration of the effect of displacement, $(\hat{D})$, rotation, $(\hat{R})$, and squeezing, $(\hat{S})$, operators on Gaussian states in phase space. The initial state is the vacuum state, $|0\rangle$.

[Ord67]. In contrast, a single-photon source exhibits *anti-bunching*, where the number of co-occurrences of events (i.e. photon arrivals) is less than what would be expected if they were classical and independent [MW95]. This anti-bunching is seen as characteristic of the quantum nature of the single-photon state. The Poisson distribution, therefore, may be considered to be the most classical description for the statistics of photon-counting.

Squeezed states

The squeezed states are another class of state that saturate the uncertainty relation given in Eq. 2.27. These, however, are *squeezed* in one of the quadratures, giving the state a lower uncertainty for that measurement, while having a higher uncertainty in the conjugate measurement. The simplest form of these is the *single-mode squeezed states*, which exist on the same phase space as the coherent states, and, as the name suggests, are quantum states on a single mode. They are produced by taking the single-mode squeezing operator, $\hat{S}_1(\zeta)$, given for an arbitrary complex squeezing parameter ζ by

$$\hat{S}_1(\zeta) = \exp\left(\zeta\hat{a}^{\dagger 2} - \zeta^*\hat{a}^2\right), \tag{2.31}$$

and applying it to a coherent state. Note that the vacuum state itself may be squeezed. The result is a state that still has a mean quadrature values of $\langle\hat{x}\rangle = 0$ and $\langle\hat{p}\rangle = 0$, yet now has a non-zero average photon number

Such squeezed states find use when we wish to measure some variable with great accuracy, without needing to measure the conjugate variable. For example, squeezed states have application in advanced laser interferometers, that need to be able to measure a slight change in a path length, but without needing to know the rate at which that path length is changing to the same accuracy [Sch17]. This is to be implemented in the Advanced LIGO interferometer [AAA+13], which will enable an increased power to detect gravitational waves.

Squeezing may also be applied to pairs of quantum modes. Two-mode squeezed states are generated on modes 1 and 2 by the operator

$$\hat{S}_2(\zeta) = \exp\left(\zeta \hat{a}_1{}^\dagger \hat{a}_2{}^\dagger - \zeta^* \hat{a}_1 \hat{a}_2\right), \tag{2.32}$$

When applied to the vacuum state this produces the two-mode squeezed vacuum (TMSV) state, which is entangled, and given for real ζ by

$$|\mathrm{TMSV}\rangle = \frac{1}{\cosh(\zeta)} \sum_{n=0}^{\infty} [\tanh(\zeta)]^{2n} |n, n\rangle_{1,2}\,. \tag{2.33}$$

The strength of the entanglement increases with the squeezing parameter. As $\zeta \to \infty$, each state $|n, n\rangle$ becomes equally likely, and the state tends towards maximal entanglement.

2.1.5 State quality measures

For quantum mechanics to be more than simply a descriptive account of nature, and to instead be a functional and prescriptive tool, it needs to have elements within it that can *quantify* the states that we produce with different measures. By such measures, we are able to say which states are good or bad for a given purpose, and use this information to characterise the capabilities of the protocols that produced such states. The number of such measures is vast, and different ones are used throughout different branches of quantum physics. Here we will introduce a few such measures which are ubiquitous throughout quantum information theory in particular.

Fidelity

The fidelity is a measure of the closeness of two quantum states, and is defined for two mixed states ρ and σ by:

$$F(\rho, \sigma) = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]^2. \tag{2.34}$$

The square root of a matrix, $\sqrt{A}$, is defined as the matrix B such that $BB = A$, with sign chosen such that B is positive semi-definite. This may be operationally calculated by diagonalising A and replacing each element in the diagonal matrix with its square root. This formula for the fidelity was first found by Uhlmann [Uhl76], and was expanded upon by Jozsa [Joz94], who showed that it was equal to the maximum transition probability between purifications of ρ and σ. Note that many sources, including [NC02] and [BBP15], define F as the square root of Eq. 2.34.

Despite first appearances, this function is in fact symmetric. This may be more easily seen when the density matrices ρ and σ are given by the pure states $|\rho\rangle\langle\rho|$ and $|\sigma\rangle\langle\sigma|$ respectively. Then we have that

$$F(\rho,\sigma) = |\langle\rho|\sigma\rangle|^2 , \tag{2.35}$$

which is the transition probability between $|\rho\rangle$ and $|\sigma\rangle$. i.e., it is the probability of getting the measurement corresponding to some measurement projector $|\rho\rangle\langle\rho|$ when such a measurement is performed on σ.

For the symmetry of the general case, we can write the fidelity as

$$\begin{aligned} F(\rho,\sigma) &= \mathrm{Tr}\left[\sqrt{\left(\sqrt{\rho}\sqrt{\sigma}\right)\left(\sqrt{\rho}\sqrt{\sigma}\right)^\dagger}\right]^2 , \\ &\equiv \|\sqrt{\rho}\sqrt{\sigma}\|_1 \end{aligned} \tag{2.36}$$

where $\|\cdot\|_1$ is the trace norm. By noting that $\|x\|_1 = \|x^\dagger\|_1$, we may see that $F(\rho,\sigma) = F(\sigma,\rho)$. This matches with what we might expect, given Jozsa's interpretation of the fidelity as the transition probability in a larger space.

Trace distance

The fidelity may be related to another common function for measuring state distinction: the trace-distance [FVDG99], given by

$$D(\rho,\sigma) = \frac{1}{2}\mathrm{Tr}[|\rho-\sigma|] , \tag{2.37}$$

where $|A| = \sqrt{A^\dagger A}$.[17] This is related to the fidelity by [FVDG99]

$$1 - \sqrt{F(\rho,\sigma)} \leq D(\rho,\sigma) \leq \sqrt{1 - F(\rho,\sigma)}. \tag{2.38}$$

The trace distance also has a role of particular importance within state discrimination. It has been shown that $1 - D(p\rho,(1-p)\sigma)$ is the minimum probability of error when performing a measurement to discriminate the two states, when ρ is given the prior probability of p [Hel69].

2.1.6 Entropy

The concept of entropy harks back to the study of thermodynamics that pre-dates its quantum-theoretical interpretation, and even pre-dates its interpretation in terms of statistical mechanics. It fundamentally characterises the *disorder* in a system. To define the entropy of a system, we should first characterise the system as a probability distribution. We should say that the system, $\mathcal{X}$, can be in state 1 with probability p_1,

[17] Note that this is equivalent to A picking up a minus sign in the cases where doing so would make it positive semi-definite.

state 2 with probability p_2, and so on. The entropy of the system is then given by Boltzmann's famous equation:[18] [Bol12],

$$S = -\sum_i p_i \log(p_i). \tag{2.39}$$

There are various different definitions of entropy, varying by an overall factor, and different bases for the logarithm (which is equivalent to a change of constant of proportionality). We will typically use base 2 logarithms, due to working mainly with binary probabilities. This was extended to the case of general mixed quantum states by von Neumann [Neu32], who showed that a good characterisation of the entropy in a quantum state was given by

$$S(\rho) = -\operatorname{Tr}[\rho \log(\rho)]\,. \tag{2.40}$$

Following the seminal work of Claude Shannon in 1948 [Sha48], the quantity S is often referred to as the *information*. He was concerned with finding the maximum lossless compressibility of a typical random string of characters drawn from some finite alphabet, described by the random variable $\mathcal{X}$. He found that the minimum length that such a compressed message could be was equal to that of the original message multiplied by a factor that he called the information, $H(\mathcal{X})$, which turned out to be exactly equal to the entropy of the probabilities of the characters in the alphabet![19] The discovery of the identity $S(\mathcal{X}) = H(\mathcal{X})$ proved to be foundational to the field of information theory, and allowed for the study of cryptography to be given a firm mathematical basis. In particular, it allowed for the security of a protocol to be defined in terms of mutual information between a legitimate party and an eavesdropper. This is discussed in Box 2.1.6, Box 3, Box 3.1.3, and in more detail in Section 3.1.3.

By the form of Eq. 2.39, it can be seen that a string formed by only repetitions of one character contains zero information,[20] whilst a string that contains an even distribution of all characters maximises the information of the string. Note that this is to be expected: If, from an alphabet $\{a,b,c,d\}$, we told that we could only send messages that contained the letter b, then there would be no way to encode any useful information. Any message could be condensed down to the statement "only the letter b," regardless of the length of the message.

[18] The form presented here is formally the Gibbs entropy. However, this was in fact discovered first by Boltzmann in [Bol66], where he proposed $S = -\rho \log(\rho)$. There, ρ was a density in phase space, which is equivalent to probability density, although he didn't make this connection explicit until after Gibbs [Gib78].

[19] Note that this only applies when the characters in the message are chosen *randomly and independently* according to some probability distribution. When we analyse the information content in, for example, an actual language, then this does not strictly apply, since the letters are not independent of one another.

[20] Here, $0 \times \log(0)$ is defined to be 0

Information

Shannon's original proof that the minimum compressible length of a message is given by $H(\mathcal{X}) = -\sum_{p_i|i\in\mathcal{X}} p_i \log(p_i)$ is complex, and it is not necessary to recreate it here. However, we may give an intuitive explanation for the equivalent statement that this quantity is the information contained within a typical message an alphabet described by $\mathcal{X}$.

If we let $i \in \mathcal{X}$ be the unique letters which comprise a message, each with probability p_i, then we can first ask, what is the information content we get when we receive letter i? We may note that we expect less likely letters to convey more information. For example, imagine Alice is thinking of a word, and Bob is trying to guess it letter-by-letter. If Bob correctly guesses that "E" is in the word, then that does not tell him much. However, if Bob were to guess the far less common letter "Q," then he would be in a far better position to guess the final word. If he were to go on to guess the double-letter combination of "QQ," then according to the standard UNIX dictionary he would have immediately narrowed it down to the single word "zaqqum."[21] We therefore see how a single additional piece of information leads to far greater knowledge of the final word.

Codifying this, we can suppose that we get two letters, i and j. We want to find what function of their probabilities describes their joint information content. We can let $h(p_i)$ be the information content of letter i. When we receive both letters, we receive an amount of information from both of them, and have thus learned an amount $h(p_i) + h(p_j)$. However, we also know that the probability of receiving the letters i and j in sequence is $p_i p_j$. Therefore, the information in this pair should be given by $h(p_i p_j)$. If we assert that $h(p_i)$ should be a continuous, smooth function, then the equation $h(p_i) + h(p_j) = h(p_i p_j)$ is solved for all p_i, p_j only by $h(p_i) = \log(p_i)$. The information content of the whole typical message is then given by the average information in a letter, which is given by $H(\mathcal{X})$ as above.

[21] A tree that is said in the Quran to grow in Jahannam (hell), and to feed its inhabitants with fruit shaped like demonic heads [ZCY18].

2.2 Entanglement

In their simplest form, multipartite states may be *simply separable* (also known as *product states*), in which case they correspond to our natural notions of particles having a specific form that is independent of the states of other particles. In more concrete terms, this means that

$$S(\rho_{1,2,\cdots,n}) = \sum_{i=1}^{n} S(\rho_i), \tag{2.41}$$

where ρ_i is state $\rho_{1,2,\cdots,n}$ with a partial trace over every subspace except i. In other words, everything about state $\rho_{1,2,\cdots,n}$ may be discovered by measurements only on single particles, and the combined state may be written as

$$\rho_{1,2,\cdots,n} = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n. \tag{2.42}$$

These may be generalised to *separable states* which include some classical uncertainty, but where each component is still a product states. i.e. states of the form

$$\rho = \sum_j p_j \, \rho^j_{1,2,\cdots,n}, \tag{2.43}$$

where each $\rho^j_{1,2,\cdots,n}$ is a density matrix of the form of Eq. 2.42, and each p_j is a probability.

Far more interesting, however, is the class of particles for which, in a very real sense, the whole is greater than the sum of its parts. Multipartite states for which the above equations do not hold are known as *entangled*.

2.2.1 Bell pairs and entanglement quantification

A Bell pair [Bel64] is one of a family of states consisting of two entangled qubits. Let $|0\rangle$ and $|1\rangle$ be the computational basis states for a qubit. The four canonical Bell states are then as follows:

$$\begin{aligned}
|\Phi^+\rangle &= \frac{|0,0\rangle + |1,1\rangle}{\sqrt{2}}, \\
|\Phi^-\rangle &= \frac{|0,0\rangle - |1,1\rangle}{\sqrt{2}}, \\
|\Psi^+\rangle &= \frac{|0,1\rangle + |1,0\rangle}{\sqrt{2}}, \\
|\Psi^-\rangle &= \frac{|0,1\rangle - |1,0\rangle}{\sqrt{2}}.
\end{aligned} \tag{2.44}$$

However, any state that is equivalent to the above under the application of local (single-particle) unitary operations can be considered a Bell state.

The Bell states are important because they are *maximally entangled*. This means that the information contained within them is entirely non-local. We can see this by comparing the entropy of the local states with the global state. If, for example, we let $\rho = |\Phi^+\rangle\langle\Phi^+|$, then $S(\rho) = 0$, a statement which is true for all pure states. This can be interpreted as saying that there is no classical uncertainty in the composition of the state when its global expression is known. On the other hand, $S(\rho_A) = 1$.[22] This means that if Alice only holds her part, she does not know whether the global state is $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, or $|\Psi^-\rangle$. She does not even know whether her state is part of a bipartite state, is unentangled, or even part of a larger entangled structure. This local uncertainty of entangled states is of great importance in the security of many quantum communication protocols.

For a pure state on two modes, A, B, a common entropic measure of entanglement is the *entropy of entanglement*, given by $S(\rho_A)$ (or equivalently $S(\rho_B)$). For mixed states, one can define the *entanglement of formation* [BDSW96] as

$$E_F(\rho) = \min \sum_i p_i S(\rho_{i,A}), \tag{2.45}$$

where the minimisation is over all ensembles $\{(p_i, |\psi_i\rangle)\}_i$ which form $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, and where $\rho_{i,A}$ is the partial trace over mode B of $|\psi_i\rangle\langle\psi_i|$. As one might expect from an entropy-based measure, this is sub-additive under composition of independent subsystems, and is believed to be additive.[23] That is to say,

$$E_F(\rho_1 \otimes \rho_2) \leq E_F(\rho_1) + E_F(\rho_2), \tag{2.46}$$

and it is conjectured that the "$\leq$" may be replaced with a "$=$".

The Bell pairs have the important property of forming a basis over $\mathcal{H}_4$. However, clearly not every state in $\mathcal{H}_4$ is entangled. From this we can draw the conclusion that a superposition of entangled states is not necessarily entangled, nor is a classical mixture of them. An important class of mixed states in the Bell basis is the class known as Werner states. These are states defined as having a certain fidelity, F, with respect to a target Bell state. If our target state was $|\Phi^+\rangle$, then a Werner state would be one of the form

$$\rho_W(F) = F\,|\Phi^+\rangle\langle\Phi^+| + \frac{1-F}{3}\left[|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|\right]. \tag{2.47}$$

Werner states typically arise from a target state undergoing an error. Suppose we initially have a copy of $|\Phi^+\rangle$. If, with probability ϵ, it undergoes some error, then it will be mapped to some mixture $(1-\epsilon)\,|\Phi^+\rangle\langle\Phi^+| + \epsilon\rho$. If we are ignorant to the model or mechanism of the model, then we must assume that we lose all

[22] With logarithms taken to base 2.

[23] This is a long-held conjecture, but which has only been proved to be true for a number of special cases, [BN01].

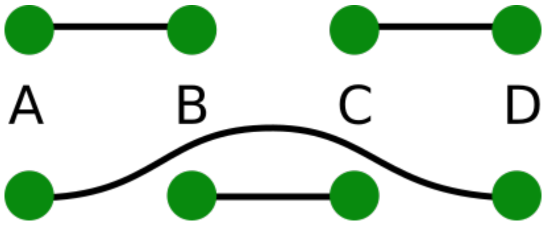


FIGURE 2.2: Entanglement swapping. Top shows A entangled with B, and C with D. A Bell-state measurement is performed on B and C, resulting in A entangled with D, and B with C.

information about the state after an error, letting $\rho = \mathbb{1}_4/4$, where $\mathbb{1}_n$ is the $n \times n$ identity operator.

2.2.2 Entanglement swapping and teleportation

Once entanglement has been generated, it may be transferred between different modes and particles. This is one of the ideas that has led to the conception of entanglement as a resource, since like many resources, we can more easily change the form of it than increase its quantity. This resource theory of quantum entanglement is particularly relevant when we wish to use entanglement to communicate securely — an idea that is investigated further in Section 3.1.

Consider 4 modes, A, B, C, D. We may think of these as, for example, the spin states of 4 electrons. Now suppose that A is entangled with B in the state $|\Phi^+\rangle$, and C with D, also in $|\Phi^+\rangle$. This is shown in the top part of Fig. 2.2. The global state is therefore

$$\begin{aligned} |\Psi\rangle_{A,B,C,D} &= |\Phi^+\rangle_{A,B} \otimes |\Phi^+\rangle_{C,D} \\ &= \frac{1}{2}\left(|00,00\rangle + |00,11\rangle + |11,00\rangle + |11,11\rangle\right)_{A,B,C,D} \end{aligned} \tag{2.48}$$

where $|0\rangle$ and $|1\rangle$ represent the spin-up and spin-down states respectively. This state may be written as

$$\begin{aligned} |\Psi\rangle_{A,B,C,D} = &|\Phi^+\rangle_{A,D} \otimes |\Phi^+\rangle_{B,C} + |\Phi^-\rangle_{A,D} \otimes |\Phi^-\rangle_{B,C} + \\ &|\Psi^+\rangle_{A,D} \otimes |\Psi^+\rangle_{B,C} + |\Psi^-\rangle_{A,D} \otimes |\Psi^-\rangle_{B,C}\,. \end{aligned} \tag{2.49}$$

Now suppose that we measure modes B and C in the Bell basis. That is, we perform a measurement such that each of the projectors is of the form $|\Bell\rangle\langle\Bell|$, where $|\Bell\rangle$ is one of the 4 Bell states. It is clear from Eq. 2.49 that if the measurement result corresponding to state $|\Bell\rangle$ is found, then this will project both the pair B, C *and* the pair A, D onto state $|\Bell\rangle$. This is shown in the lower part of Fig. 2.2. If the users at node A and D wish to use their entangled pair for any practical purpose, such as creating a shared secret bit (as explained in Section 3.1.1), then they must wait for the users at B and C to send a classical signal indicating which Bell state they measured. Before receiving this signal their shared state is an equal mixture of all four Bell states. This is the maximally mixed state, which, by any measure, contains no entanglement.

The process is sometimes referred to as *teleportation of entanglement*. The reason for this can be seen most clearly when we consider a comparable protocol that uses a single entangled Bell pair. Suppose that Alice has a quantum state,

$$|\psi\rangle_{A_1} = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{2.50}$$

and that she and Bob share the state $|\Phi^+\rangle_{A_2,B}$, with state labels unrelated to those in the previous section. The composite state may now be expressed as

$$\begin{aligned}|\psi,\Phi^+\rangle_{A_1,A_2,B} =& \frac{\alpha}{2}\left(|\Phi^+\rangle + |\Phi^-\rangle\right)_{A_1,A_2}|0\rangle_B + \frac{\alpha}{2}\left(|\Psi^+\rangle + |\Psi^-\rangle\right)_{A_1,A_2}|1\rangle_B + \\ & \frac{\beta}{2}\left(|\Psi^+\rangle - |\Psi^-\rangle\right)_{A_1,A_2}|0\rangle_B + \frac{\beta}{2}\left(|\Phi^+\rangle - |\Phi^-\rangle\right)_{A_1,A_2}|1\rangle_B\,.\end{aligned} \tag{2.51}$$

Alice now measures her two modes, A_1 and A_2, in the Bell basis. Leaving Bob with a state $|\psi'\rangle_B$. This is related to $|\psi\rangle$ by the application of a local unitary operation, consisting of Z, X, or ZX. The unitary that is to be applied is dependent on the result of Alice's Bell-state measurement. By doing this, we can see that Bob now holds the original state, meaning that the state has been teleported from Alice to Bob![24]

2.2.3 Graph-state entanglement

It may be said that entanglement represents the greatest departure from classical physics, and as such, many of the techniques available to us to study the classical world no longer hold. It is therefore of great benefit to have access to techniques that aid in the mental manipulation of large and complex states, allowing for a mathematical method more intuitive than performing arithmetic on an exponentially large number of coefficients. In particular, from our first conception of a ball rolling down a slope to performing scattering calculations using Feynman diagrams, our understanding of the physical world often progresses through use of visualisations. When studying a classical system, N particles, each with M degrees of freedom (e.g. standard positional space, where $M = 3$) may be visualised as N points in an M-dimensional space. With the exception of the particular case where the state across N modes is fully separable, that luxury is not available when we

[24] Of course, some will object that this has not teleported the object itself, but its current state of being. However, it is reasonable to say that an object is nothing *but* the set of physical properties which it holds. A corollary of this is the philosophical principle known as the *identity of indiscernibles*, which states that two objects cannot be said to be distinct if they share all properties. This is discussed in a dialogue in [Bla52], and in the context of quantum physics in [FR88]. A counter to this may be the accurate claim that Alice and Bob's particles may have other properties that render them distinguishable, such as energy, although in principle, such continuous variables may also be teleported [BK98].

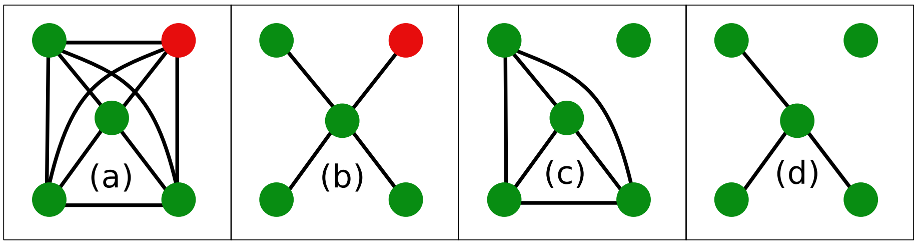


FIGURE 2.3: Graph states, showing measurements and local-unitary (LU) equivalence. Graphs (a) and (b) both represent states that are equivalent up to LU to the GHZ state: $|GHZ\rangle = \left[|0\rangle^{\otimes 5} + |1\rangle^{\otimes 5}\right] / \sqrt{2}$. Graphs (c) and (d) are LU-equivalent to each other, and are the result of applying either Z, Y, or X to the upper-right qubit shown in red in (a) and (b). We see here how graphical calculations that may be tedious for a given state may become more simple when we apply LUs, perform the measurements, and invert the LUs.

move to the quantum world, where the vast size of the Hilbert space implies that a state must be represented as a single point in an M^N-dimensional space.

Graph states represent a subset of the N-qubit Hilbert space that permits a nice visualisation in terms of graphs, as shown in Fig. 2.3 [HEB04]. Like all the best pictorial representations,[25] this allows us to quickly understand and analyse the details of a system, and even perform calculations without going through the full mathematical machinery of state vectors and matrix operators. Importantly, many of the quantum states involved in communication networks may be represented as graph states.

An undirected graph is a collection of nodes, and a set of undirected edges between them. An empty graph is a set of nodes with no edges between them. In terms of quantum states, this represents a collection of N qubits, initialised in the $|+\rangle$ state, where each qubit is identified with a node. A graph with a single edge between nodes i and j represents the state formed by applying a CZ gate to qubits i and j, defined as:

$$CZ_{i,j} \, |a,b\rangle_{i,j} = (-1)^{ab} \, |a,b\rangle_{i,j} \, . \tag{2.52}$$

This operator stands for *controlled* Z, and is so-called since it applies a Z gate to mode j, conditional on mode i being in state $|1\rangle$ (or vice-versa). We can similarly define CX (also known as CNOT) and CY gates.

The degree of a node is the number of edges connected to it. When nodes i and j are connected by an edge and they each have a degree of one, then the qubits represented by these nodes are in the state

$$|\psi\rangle_{i,j} = \frac{1}{2}\left[|0,0\rangle + |0,1\rangle + |1,0\rangle - |1,1\rangle\right]_{i,j} , \tag{2.53}$$

[25] Such as Feynman diagrams [Mat92], tensor networks [WBC11, Orú14], and the diagrams of Hurst [HK18].

which is a maximally-entangled state, equivalent up to single-qubit unitary operations to the Bell pairs.

While any graph can be drawn, and the corresponding state be written down, it is not the case that an edge between any pair of nodes always implies that they are perfectly quantumly correlated. This is because the entanglement structure in any state must obey *monogamy of entanglement*. Simply put, this says that if modes A and B are maximally entangled, then they can have no correlation whatsoever with a third mode C. Conversely, as the strength of entanglement between C and the (A, B) subsystem increases, the entanglement between A and B decreases, as measured by the concurrence [CKW00] or the entanglement of formation [dOCF14].

In many quantum systems, we will build up a graph-state by performing entanglement-generating operations between distant nodes. This is particularly the case for quantum repeaters (see Section 3.2 and Chapter 5), although these graphs are often of a fairly simple, linear type (exceptions do exist, including more complex graph-state repeaters [EKB15] and repeaters between more than two parties [WZM+16]). When entanglement is used in some quantum procedure, we typically want our correlations to be as strong as possible, whilst ensuring that the results of our measurements are unknowable to any third party that does not possess a part of the entangled state. We therefore wish to know, given some graph representing a state, what series of operations we can perform such that we are left with a maximally-entangled pair across the two qubits of our choosing.

This may be done by performing Pauli measurements on the qubits. Fortunately, the operation of these measurements may be given as transformations of the graph, and so allow us to work entirely within a graph formalism, instead of working directly with the state and keeping tracking of hundreds of amplitudes and phases. The Pauli measurements have the following graphical operational effects when applied to a qubit i [HEB04]:

- Z_i: Qubit i is removed from the graph, along with all edges connected to it.
- Y_i: The neighbourhood of i, $\mathcal{N}_i$, should be identified. This is the subset of nodes that are connected to i by an edge. The subgraph on this subset is then inverted. This means each edge between pairs of nodes in this subset is removed, and each pair of nodes becomes connected by an edge if formerly they were not. A Pauli Z measurement is then applied to qubit i.
- X: A secondary qubit $j \in \mathcal{N}_i$ should be chosen. The subgraph on $\mathcal{N}_j$ is inverted, and Y applied to qubit i. The subgraph on the new neighbourhood of j is then inverted again.

Note that all measurements end up removing qubit i from the graph. This is because the information in i is then fully known, and it can no longer have any quantum correlation with any other subsystem. Note that we are here only discussing

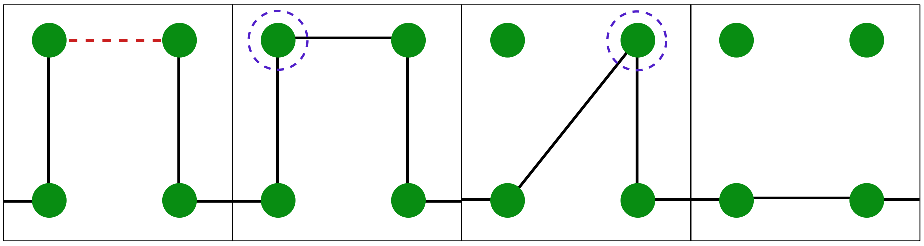


FIGURE 2.4: Entanglement-swapping by CZ gates. The red dashed line shows where a CZ unitary is to be applied, and blue dashed circles show qubits to be measured in the Pauli Y basis.

the operations up to local unitaries. This means that after performing such measurements, single-qubit operations, contingent on the results of the measurement, will typically need to be applied to the state to return it to its canonical graph-state form. Here we neglect the full discussion of such operations. This is because when using graph states for the transmittance of ultimately classical signals, as is the case in QKD, the effects of local unitaries can be applied after the completion of the protocol, by performing conditional flips on the bit values generated by the protocol.

In Fig. 2.4 we illustrate how conditional phase rotations (CZ gates) plus Y measurements can be used as an alternative schema to perform entanglement swapping without the necessity of performing direct two-qubit measurements.

2.2.4 Entanglement distillation

Entanglement distillation is the process by which we may take multiple entangled pairs, each with some low fidelity with respect to a specific Bell state, and produce from them a smaller number of pairs of higher fidelity. This is an essential feature of quantum technologies, since many entanglement-based quantum technologies, such as quantum key distribution (Section 3.1), set a minimum threshold fidelity for the entangled states that they use. If the shared states are of a fidelity below this threshold then no secret key can be generated, no matter how many states are shared. There is, therefore, a natural trade-off between the number of states in one's possession and the fidelity of those states. This trade-off is explored in more detail in Section 6.

Distillation was first introduced by Bennett et al. in [BBP+96], and refined by Deutsch et al. in [DEJ+96]. It is this latter scheme that we will use throughout this thesis when we discuss distillation, and we refer to it by the initials of its inventors: DEJMPS. This operates on two entangled pairs, which we will call $\rho_{1,2}$ and $\rho_{3,4}$ (on modes 1, 2, 3, and 4). Here we will discuss the mathematical structure of distillation and its effect on quantum state operators. For details on how this may be physically implemented, see Section 5.1.1 We may assume without loss of generality in this protocol that both pairs are diagonal in the Bell basis. The coefficients of

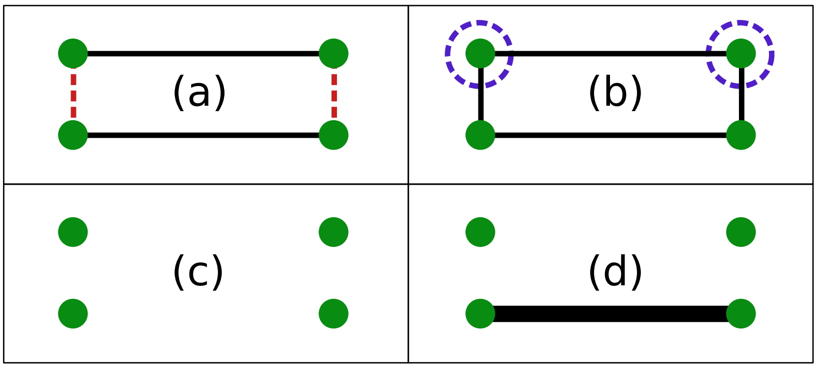


FIGURE 2.5: Distilling two low-fidelity Bell pairs to a higher-fidelity pair. Black lines indicate entanglement, with the width of the line indicating fidelity. In (a), Alice entangles her qubits on the left (shown with a red dashed line), and Bob on the right entangles his qubits. In (b), Alice and Bob then measure the qubits of one pair, shown by blue dashed circles. If these measurement disagree, Alice and Bob are left with (c), a system with no entanglement. If they agree, they are left with (d), which has a stronger entanglement between the remaining pair.

$|\Phi^+\rangle, |\Psi^-\rangle, |\Psi^+\rangle, |\Phi^-\rangle$ will be denoted a, b, c, d respectively for $\rho_{1,2}$, and a', b', c', d' respectively for $\rho_{3,4}$.

Now the two pairs are entangled with each other: Mode 1 is entangled with mode 3, and mode 2 with mode 4. This additional entanglement is meant in the sense of graph-state entanglement as described in Section 2.2.3, so is created by applying a CZ gate between the nodes to be entangled.

Now modes 1 and 2 are measured in the Z (computational) basis. If the results are different, then both pairs are discarded. If, however, the outcomes are both 0 or both 1, then modes 3 and 4 are kept, and are now in a state defined by the following Bell state coefficients:

$$\begin{bmatrix} (aa' + bb')/N \\ (c'd + cd')/N \\ (cc' + dd')/N \\ (a'b + ab')/N \end{bmatrix}, \tag{2.54}$$

where

$$N = (a + b)(a' + b') + (c + d)(c' + d') \tag{2.55}$$

is the probability that the distillation attempt will succeed (equal outcomes). This procedure is shown in Fig. 2.5.

The necessity for measurement results to agree may cause problems when using distillation in a practical context. For example, a common use of distillation is when modes 1 and 3 are held by one party (Alice), and 2 and 4 are held by another (Bob) who is spatially separated from Alice. They may wish to use distillation to produce

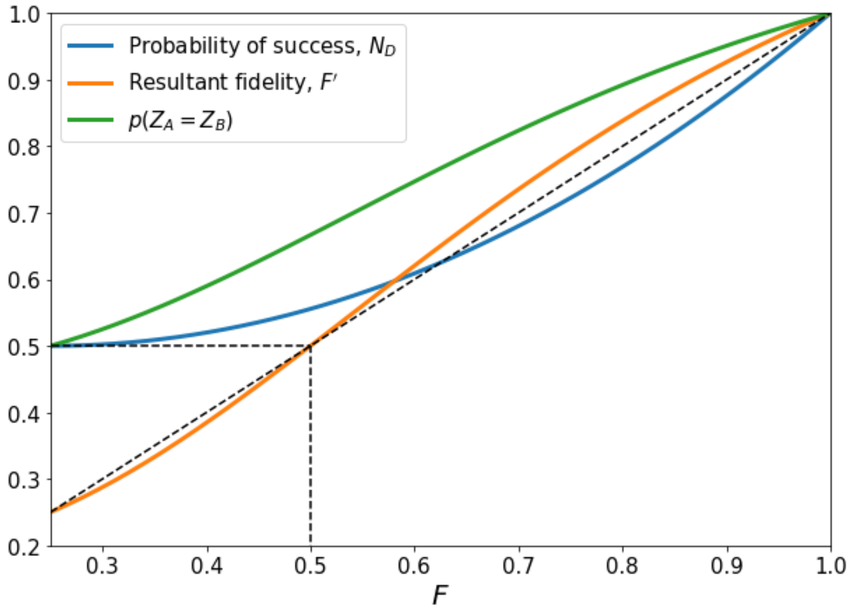


FIGURE 2.6: Probability of distillation success, N_D, resulting fidelity, F', and probability of bit-wise agreement for two distilled Werner states, each of initial fidelity F. Fidelity starts at 0.25, which represents a maximally-mixed (and so maximum entropy) state.

high-fidelity pairs in order to communicate. However, when the qubits are realised by quantum memories (for example, as up and down spins of a bound electron or nucleus), then the fidelity of these states will decay over time, due to the finiteness of the memory lifetimes. Therefore, after Alice and Bob have waited for the signals from each other to determine whether or not a distillation attempt was successful, their shared state will have dropped back down to a lower fidelity.

One technique that may be used to get around this is *blind distillation*. Here, instead of checking if the attempt was successful and repeating it if it was not, Alice and Bob assume that it *was* successful, and go on to use their state accordingly by, for example, performing a measurement on their remaining qubits. Only after this has been made do they then go back and compare the results of their first measurements to determine whether the distillation was successful (and so whether their remaining qubits were in fact entangled, which determines whether their second measurements should have correlated or not). Fortunately, distillation *usually* succeeds. If we are distilling two Werner states, each with an initial fidelity with respect to $|\Phi^+\rangle$ of F, then the probability of success is given by

$$N_D = \frac{8F^2 - 4F + 5}{9}, \tag{2.56}$$

which is shown in Fig. 2.6. We also show the fidelity of the resulting state with respect to $|\Phi^+\rangle$, given by

$$F' = (aa' + bb')/N_D = \frac{10F^2 - 2F + 1}{9N_D}, \tag{2.57}$$

as well as the probability that, if Alice and Bob measure each part of the resulting state in the Z basis, they will get the same measurement result, given by $p(Z_A = Z_B) = F' + (1 - F')/3 = (2F' - 1)/3$, where this comes from the fact that they will agree if they share $|\Phi^+\rangle$ or $|\Phi^-\rangle$. This final quantity is important when realising the quantum communication protocol E91, as explained in Section 3.1.2.

The decision of whether or not to use blind distillation becomes important when we consider multiple entanglement connections chained together in a quantum repeater (Sections 3.2 and 5.1.2)

2.2.5 The ontology of quantum states

The quantum state, as the primitive object of quantum theory, has been the subject of intense mathematical analysis for the last century or so, a small sample of which has been presented in the preceding sections. However, this does not answer the question: "what actually *is* a quantum state?" Is its nature something physical, independent of the questions asked of it? Is it a fancy calculational trick? Or is it a shadow of something deeper and more fundamental? To answer this, we should first dispel some notions as to what the state is *not*.

Firstly, the state does not represent a distribution of the matter of the system itself. It is a density of probability, but it is not a density of "stuff." If it were, then an electron that is passed at speed through a narrow slit should become spread-out on a screen ahead of it. For a fundamental unit charge of e, then in such a situation we might find $0.3e$ of charge over here, $0.5e$ over there, and $0.2e$ somewhere else. This is not the case; we always find electrons comprising exactly one unit charge, and always in a single location. No matter the broadness of the spread of its likely locations, it remains a point-like particle.

Equally, the state does not represent the expression of a preordained set of local variables. This is a conception of the quantum state that we may have when first encountering the mysterious correlatory properties of the Bell state. When it is learned that Bob will always get the same measurement result as Alice, we may ask whether that was not determined from the start. After all, if I were to put either two red balls or two blue balls into a bag, and ask Alice and Bob to each withdraw one, the situation would seem superficially similar. Alice has no knowledge of which ball she will withdraw, but once she does, she has perfect knowledge of what Bob will draw. However, the key difference is that such a system is classical. In a quantum system, she would have the possibility to draw in the {red + blue, red - blue} basis. The fact that strong correlations exist even when Alice and Bob are free to choose the basis of measurement implies that quantum theory cannot be described by local hidden variables or the statistics of classical objects. That is to say, there does not exist a variable describing the system that has a value determined before measurement and only transmits information slower than the speed of light.[26] Quantum theory breaks either or both of these suppositions, depending on one's preferred interpretation. The degree to which a system deviates from a classical statistical description is quantified by the CHSH inequality, which is discussed in Section 3.1.2

The Copenhagen interpretation [Sta72] (amongst many others) rejects the apriori objective reality of physical properties such as momentum. According to this theory, a measurement causes a collapse of the state, whereupon one out of the many eigenstate possibilities is "chosen," and the rest vanish. This is the interpretation most commonly taught in universities, and is likely how many theorists visualise their state dynamics in day-to-day calculations. However, many take issue with the fact that it seems to assign special meaning of a measurement, in addition to, as Einstein said, "playing dice with the universe."

The de Broglie-Bohm (dBB) theory [Hol95], on the other hand, proposes to reject the notion that physics must rely only on slower-than-light signals. It suggest the existence of faster-than light carrier waves that guide particles through space, like a bouncy ball carried on a turbulent sea. This gained in popularity for a while, when experiments by the team of Couter on a model where a single droplet bouncing on the surface of a bath of liquid seemed to reproduce the results of the double-slit experiment [CF06]. However, more recent experiments seem to show that the connection was only surface-deep, and such droplet-based models do not recreate the quantum case [PHFB18]. Of course, such a model does not in and of itself conclusively disprove the existence of a non-local objective-reality theory. However, notwithstanding experimental objections, dBB does not seem to give a convincing answer as to how entanglement may be incorporated, nor explain the source of the power of quantum computers.

Whereas dBB theory tries to put the physical world back on a quasi-classical, objective grounding, some interpretations of quantum mechanics go in the other direction. In particular, the theories of Relational Quantum Mechanics (RQM) [Rov96], by Rovelli, and Quantum Bayesianism (QBism) by Fuchs and Mermin [FMS14, FS16, Fuc17], take a fully Bayesian approach to the problem. In QBism, it is claimed that the probabilities assigned within quantum theory are degrees of confidence, and not representative of any absolute, objective claim about the world. This may be seen as an extension of the classical Bayesian statistical theory. Seen in context of the

[26] The speed of light restriction is necessary because Alice and Bob are free to choose their measurement settings immediately before their halves of the state reaches them, which would prevent any signals being sent back to the other part of the state. It has been noticed that there are a class of interpretations, known as *superdeterministic theories* [Lal01], for which this is not necessary. One could argue that the physical universe is fully deterministic. Therefore, it is erroneous to suggest that Alice and Bob had free will in choosing their measurement settings, and that they were in fact predetermined at the inception of the universe. Unsurprisingly, such theories are not popular. In part, this is because they are unfalsifiable. Of course, this in itself does not sound the death knell of a theory if it is sufficiently elegant, such as string theory. Superdeterminism, however, necessitates that the universe was constructed in just such a way at the very moment of the Big Bang, such that every particle movement collaborated to fix the behaviours of all quantum physicists to the specific set of actions that would render the theory consistent. This may be regarded as the strongest and ugliest possible form of the anthropic principle, and as such is not taken seriously by many.

question of the failure of absolute non-realism or absolute non-locality, Fuchs says of both RQM and QBism that “rather than relinquishing the idea of reality (as they are often accused of), they are saying that reality is more than any third-person perspective can capture,” [Fuc17] by introducing a realism that is contingent upon the participation of the experimenter.

There may be said to be as almost as many quantum interpretations as there are quantum physicists [SN16]. However, except in the most extreme of philosophical cases[27] one’s choice of interpretation makes no difference to one’s choice of action, or the results of calculations. The rest of this thesis will therefore largely focus on the calculational aspects of various quantum problems and their implications for real-world security, without needing to dwell on the philosophical interpretations.

[27] For example, an ardent believer of Everett’s many-worlds theory [EI57, DG15] may take more risks, in the belief that he will by necessity always experience a world where he is alive (see *quantum suicide*, [Teg98]).

Chapter 3

Quantum communication preliminaries

In this section we will move from quantum mechanics generally, to the specific field of quantum key distribution. When it was realised that classical cryptography methods would no longer be sufficient to protect transmitted data in the age of quantum computers, a number of new protocols were developed that promised information-theoretic security. We discuss a few of these in this Section 3.1, and a few within the context of long-range cryptography in Section 3.2. It is here that we also discuss some of the mathematical fundamentals of QKD. Of particular importance is Subsection 3.1.3, in which we examine how security may be guaranteed within the context of a noisy protocol, and the rates at which secure communication can be guaranteed.

Information-theoretic security

An example of a protocol that is information-theoretically secure is the *one-time pad*. Here, Alice might have some message, m, selected from the message set $\mathcal{M} = [0,1]^n$ for some $n < \infty$, and a key $k \in \mathcal{M}$. The secret is then simply produced by $s = k \oplus_2 m$. The secret is then sent on to Bob, who should also possess k. Bob then reproduces the message by $m = k \oplus_2 s$. It may be seen by intuition, and was proved by Shannon [Sha48], that there is no way to reconstruct m from s without k, since the two have no *mutual information*. Let m and s be considered as specific values of random variables M, S, drawn with probabilities $p(m), p(s)$. Then:

$$I(M;S) = \sum_{m,s} p(m,s) \log \left(\frac{p(m,s)}{p(m)p(s)} \right) = 0. \tag{3.1}$$

This may be easily seen to be the case by noting that if k is randomly selected, then $p(m,s) = p(m)p(s)$. The problem with this scheme is that in order to use it to share a secret message, it requires that both parties already possess a copy of k. Therefore Alice and Bob would need to posses a separate secure channel, over which to distribute k, which means we have only moved the problem around but not solved it. This problem is addressed by public-key cryptography, at the expense of guaranteed security.

3.1 Quantum cryptography protocols

Here we will discuss two of the most ubiquitous QKD protocols which serve as a basis for many of the more complex protocols mentioned in Section 3.2. In Subsection 3.1.2, we show that this ubiquity is greater than it may initially appear, with a deep equivalence between the two protocols.

3.1.1 Bennet and Brassard 1984

The first quantum cryptography protocol to be developed was that of Charles Bennet and Giles Brassard in 1984 [BB84], which is usually referred to as BB84 for short. The fundamental idea which underlies the security of this protocol is the Heisenberg Uncertainty Principle, where the uncertainty inherent in certain measurements is shown to lead to a corresponding uncertainty in Eve's knowledge of the key.

Protocol description

The protocol works as follows:

- For each qubit, Alice chooses randomly whether to send a state in the Z basis $(|0\rangle, |1\rangle)$ or the X basis $(|+\rangle, |-\rangle)$.

- Alice then chooses randomly whether to send a 0 bit (represented by $|0\rangle$ or $|+\rangle$) or a 1 bit (represented by $|1\rangle$ or $|-\rangle$).

- This qubit is then encoded in a photon. As discussed in Section 2.1.3, there are multiple ways of doing this. One such way is by encoding it in the polarisation of the photon, which presents itself as a natural option due to the two orthogonal polarisation modes. Another common encoding is in the phase of the photon. However, since quantum measurements are invariant under the global phase of a state, we require a reference mode for the phase to be defined *with respect to*. One might suggest sending a second photon along with the first to act as a reference, but this would reduce the efficiency of the qubit transfer from Alice to Bob, since we would then require *both* photons to travel through the optical fibre without being attenuated. Instead the photon is split into two modes - an early mode and a late mode. The early mode is then phase-shifted by θ - a single parameter that entirely determines the qubit as follows:

  $$\begin{aligned} |0\rangle &\to \theta = 0 \qquad & |+\rangle &\to \theta = \pi/2 \\ |1\rangle &\to \theta = \pi \qquad & |-\rangle &\to \theta = 3\pi/2 \end{aligned} \tag{3.2}$$

  This is shown in Fig. 3.1. In this thesis (in particular Chapter 4) we shall assume that any BB84 protocol is run with this phase-based encoding. This is due to the fact that θ has a natural interpretation in the physical Hamiltonian when viewed in this way. We do not lose generality by neglecting the polarisation encoding, since we may transform between the two encodings by a passive linear-optical system, as shown in Fig. 3.2.

- This photon is then sent through an optical fibre towards Bob, who randomly decides on a basis Z or X and measures the photon in that basis. Suppose Bob happens to choose the same basis that Alice chose. If the photon was transmitted faithfully, with no interference of either natural or malicious origin, then Bob will measure the same bit as the one decided on by Alice. That is, if Alice chose base X and bit 0, thus sending state $|+\rangle$, then Bob would also measure state $|+\rangle$, and interpret this as a 0 bit. On the other hand, if Alice sends in base X and Bob measures in Z, then he will get the results $|0\rangle$ and $|1\rangle$ 50% of the time each, with no correlation with Alice's choice of bit.

- After sending all qubits, Alice and Bob then publicly announce which basis they used for each qubit. They discard the results of all measurements for

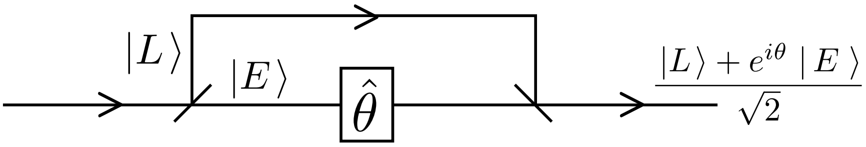


FIGURE 3.1: Phase-encoding of a photonic qubit.

which their basis choices disagreed. Then, assuming no errors occurred on the qubits in flight, the string of bits corresponding to the measurement results on the remaining qubits will be shared and identical for both Alice and Bob.

This, in the ideal case, will generate a perfectly shared key, which can be used to encrypt a message (see *Information-theoretic security* in Section 3). However, we cannot guarantee that the optical fibre has not been tapped. Clearly if Eve also possesses the key then she can read the encrypted message that will follow, so we need to be sure that this is not the case. Consider what will happen if Eve tries to measure the photonic qubits herself to try to learn the key. Like Bob, she will have no way of knowing which basis to measure in, so will have to make a random selection. Half of the time she will choose the incorrect basis. Half of the times that she measures in the wrong basis, Bob's measurement of the state when he receives it will give the wrong result (i.e. not the state that Alice sent). Therefore there will be a 25% error rate in Alice and Bob's shared key.

Security check

To ensure that no eavesdropping has occurred, Alice and Bob should perform a *security check*. Here, they choose some fraction of the key, say 5%, to be a used in the test. They take this fraction of the key bits, chosen randomly, and publicly compare them. If too many of the compared bits disagree, then the protocol is abandoned — we cannot guarantee that no one is eavesdropping on the line. The question that naturally presents itself here is "how many disagreeing bits is *too many*?" — a matter that is addressed in Section 3.1.3. An important extension of BB84 is the variation known as *efficient BB84*. Note that in the standard version of BB84, half of the photons that are sent will *not* be used to distil a key, since Alice's and Bob's bases will not match. Lo et al. showed in [LCA05] that Alice and Bob do not have to use each basis half of the time. They can, for example, agree to use the Z basis 90% of the time, and the X basis 10% of the time. They will therefore agree on a basis 82% $(= 0.9^2 + 0.1^2)$ of the time. In order to ensure that Eve cannot take advantage of this by measuring every incoming photon in the Z basis, Alice and Bob must then check the error rate in the Z basis and X basis separately. Only if *both* error rates pass the security check will the protocol proceed. Note that they cannot use, for example, just the X basis for a security check. If this was the case, then Eve could simply measure all states in the X basis, and learn 50% of the key whilst causing no errors.

An alternative efficient form of BB84 involves Alice and Bob again predominantly sending and measuring states in the Z basis. However, in this version, they

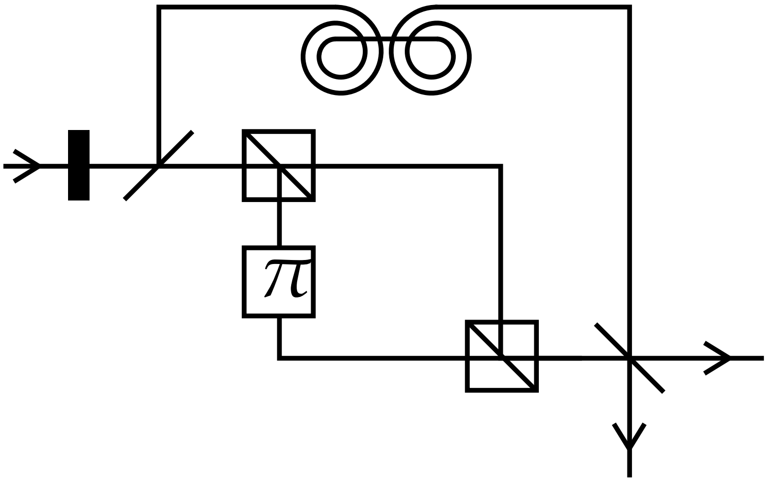


FIGURE 3.2: A linear optical system to convert between a polarisation-encoded photonic qubit (incoming on the left) and a qubit encoded in a phase shift between the early- and late-arriving modes of a photon (outgoing on right). Diagonal lines are beam splitters, and diagonal lines in boxes are polarising beam splitters. The black box is a quarter-wave plate, that maps $|V\rangle \mapsto i\,|V\rangle$. The marked box applies a constant phase shift of π. Implicit in the linearity is the fact that this system can be reversed, and phase-encoded photons send right-to-left to encode their qubit in a polarisation basis.

would *only* use the Z states for generating the key. The X basis states would then be used to check for the error-rate. This is the scheme used in experimental implementations such as [RLY+17].

Non-standard attacks

At this point one may notice that there are attacks which Eve may carry out that do not involve covertly intercepting the photons. These may be divided into two categories. The first is known as *man-in-the-middle attacks* [WWLH09]. This is where Eve impersonates both Alice and Bob. That is to say that Eve sits between the legitimate parties, and runs BB84 with both of them. In order to avoid this threat, it is necessary that the classical channel over which they communicate is *authenticated*. Unfortunately, quantum authentication systems require Alice and Bob to initially share some secret! It is therefore often said that QKD is more accurately called a *key expansion procedure*, rather than a key generation procedure. Those interested in this topic will find a great body of literature available, with theories and experiments often developed in concordance with QKD. These may involve protocols that acts as a guarantee to existing QKD protocols [BCG+02, CS01], as well as quantum direct-communication protocols, that allow for the direct sending of an authenticated encrypted message, and not just a key [DLL03, LLY06].[1] The second non-standard type of attack is known as a side-channel; a category which includes a wide variety of different types of attacks. We discuss this in length in Section 3.1.5.

3.1.2 Ekert 1991

In 1991, Arthur Ekert devised a new QKD protocol [Eke91], often referred to as E91, that uses principles of entanglement in a rather elegant manner to guarantee its security. The process of sharing a single secret bit begins by Alice generating a photonic two-qubit state in a single, publicly known Bell state, say $|\Phi^+\rangle$. As with BB84, this may be encoded in the phase or polarisation degrees of freedom of the photon. One photon from this pair is retained, and one is sent on to Bob. Alice and Bob both then choose a random basis in which to measure. In Ekert's original paper, 3 bases were used. However, the principle works equally well with 2 bases (Z and X), which allows for a more direct comparison with BB84. If (and only if) Alice and Bob share a perfect $|\Phi^+\rangle$ state, then they will get the same outcome whenever they both measure in the same basis. This shared result will then form a single, shared key bit.

A key difference between E91 and BB84 is in the nature of the security check. Here, Alice and Bob perform a *Bell test* on a subset of the pairs. This is a statistical test that aims to show that the states *cannot* be accurately described by local real variables, and so must be quantum-mechanically entangled.

[1] Whilst the idea of direct secure communication may seem promising, note that QKD has a very prominent advantage. We can take advantage of the fact that no *individual* bit corresponds initially to any bit in the message. We may arbitrarily shuffle the key with no loss of message meaning or security. Similarly, we may shorten the key to boost security, as explained in Section 3.1.3. However, this is clearly not possible with direct communication, where every bit is important.

CHSH test

The CHSH test (named after John Clauser, Michael Horne, Abner Shimony, and Richard Holt) is a measurement of the CHSH statistic, S_{CHSH}, for a set of identical pairs of qubits. This is given by

$$S_{\text{CHSH}} = \mathbb{E}(Z_A, Z_B) - \mathbb{E}(Z_A, X_B) + \mathbb{E}(X_A, Z_B) + \mathbb{E}(X_A, X_B), \tag{3.3}$$

where $\mathbb{E}(A, B)$ is the expected product of outcomes when Alice measures in basis A and Bob measures in basis B, and the outcome results are mapped to $\{+1, -1\}$. Here, we also define Z_B, X_B to be rotated in the Z, X plane by an angle of $\pi/4$ with respect to Z_A, X_A. For a classical state reproducible by local hidden variables (such as a product state) this is bounded by $|S_{\text{CHSH}}| \leq 2$. For entangled states, this bound is "violated," and for maximally entangled qubit pairs it saturates the quantum bound of $|S_{\text{CHSH}}| \leq 2\sqrt{2}$. Like the entanglement of formation, the Bell statistic therefore gives a measure of how entangled a pair of modes are.

In a way similar to the measurement of error rates in BB84, this ensures that no measurement has been made of Bob's particles before they reached him. However, the CHSH test makes it clear that the security test is in fact far *stronger* than simply excluding the possibility of malicious measurements. We also want to be sure that Eve has not entangled Bob's photon with her own, which she will then measure at a later time. Here, Alice and Bob are protected by the concept of the *monogamy of entanglement* (discussed in Section 2.2.3). Therefore, even if the entanglement is not perfect (due, perhaps, to errors accrued in transmission or detection) Alice and Bob can use distillation to increase the security of their communication, at the cost of reducing the number of bits they share. This trade-off between security and bit rates is fundamental to cryptography, and is something that we will encounter throughout this thesis in different guises.

Equivalence with BB84

Upon first inspection, it may seem that BB84 and entanglement-based systems (of which E91 is one example) are two completely separate protocols that require distinct analyses. However, it was noticed by Shor and Preskill [SP00] that one may be reduced to the other, allowing them to prove the security of BB84 by considering the behaviour of entanglement distillation.[2] To see this duality, one should first note that the local density matrix of a photon in flight from Alice to Bob is the same in both protocols. That is to say, an eavesdropper who intercepts a photon and does not make a measurement will only be able to say that the state of the qubit encoded by it is $\mathbb{1}_2/2$. Even if she measures it and concludes that its post-measurement state is, for example, $|0\rangle$, this will tell her nothing about whether Alice and Bob are using states with pre-determined orientations, as in BB84, or entangled pairs, as in E91.

Therefore, if Eve does not know this, then the physical process that is carried out is also the same from Bob's perspective. In either case he chooses a random basis and measures the photon. As for Alice, we can assume that for both BB84 and entanglement-based QKD, she produces a $|\Phi^+\rangle$ state. If she is using BB84, her next step will be to decide on a basis. She will then measure her part of the state, which will make the "decision" for her as to whether to send a 0 or a 1. Unlike the basis choice, which may be heavily skewed one way by implementing efficient BB84, the distribution of 0s and 1s should follow a 50/50 split, in order to maximise the entropy (and therefore security) of the key. It therefore does not matter whether she actually makes the decision of sending a 0 or 1, or has it made for her by measuring a Bell state. The key difference between BB84 and entanglement-based QKD can then be seen as whether Alice measures her part of the state before or after sending Bob's part to him. Since the non-local effects of entanglement are of a correlatory rather than causal nature,[3] this forms a loose equivalency between the two protocols.

We say a *loose* equivalency, since the analogy is incomplete — it does not deal with any equivalency between the security tests, error corrections or privacy amplifications. These, however, were dealt with by Shor and Preskill, who showed how quantum error-correcting codes can be used to give a comprehensive proof of the security of BB84 when identified with an entanglement-based protocol.

3.1.3 Security and secret key rates

Here, we want to examine the effects of different results from the security check on the ability to securely communicate. In both this section, and for the rest of the thesis, we will assume the BB84 security check is used, whereupon the number of errors, ϵ, in the distributed key is measured, and used to determine whether or not to proceed. By the argument at the end of Section 3.1.1, if Alice and Bob find an error rate of 25% or greater, then this would imply they might have an eavesdropper. However, a more complete analysis reveals both a more pessimistic and a more optimistic side to the equation. Pessimistically — the argument in Section 3.1.1 assumed that Eve simply used a naïve strategy of measuring every single photon that she encountered. On one hand, she may attempt a more sophisticated measurement. She may, for example, entangle many incoming photons together and then perform a

[2] Note that they formally prove an equivalence to the Lo and Chau protocol, [LC99].

[3] i.e. One cannot use entanglement to send information-carrying messages faster than the speed of light, which would enable information to be sent backwards in time.

joint measurement on them. On the other hand, she may not even measure all of the photons. She may decide that only measuring *half* of the photonic qubits is enough to satisfactorily inform herself of the secret key.

However, if Alice and Bob *do* believe that Eve has learned some of the key, then they do not necessarily have to abandon the protocol. It is here that we return to the trade-off between security and bit rates, for there are classical *privacy amplification* protocols that Alice and Bob may carry out on their bit-string, b, returning a shorter but more secure bit-string, b'. By "more secure," we here mean has a lower mutual information with Eve's best estimation of the string.

Consider the following simple example protocol. Suppose Alice and Bob share the bit-string, b. They also believe that Eve knows 50% of the bits, but they do not know which bits. Let e be Eve's best estimate for b. i.e. a string which is equal to b on the marked bits, and random on the unmarked bits. An example of this is:

$$\begin{aligned} b &= \underset{\bullet}{0}\underset{\bullet}{0}\underset{\bullet}{1}000\underset{\bullet}{1}0\underset{\bullet}{0}1\underset{\bullet}{1}0\underset{\bullet}{0}\underset{\bullet}{1}00\underset{\bullet}{1}0\underset{\bullet}{1}0\underset{\bullet}{1}\underset{\bullet}{0}\underset{\bullet}{0}1\underset{\bullet}{0}\underset{\bullet}{0}\underset{\bullet}{0}1\underset{\bullet}{1}\underset{\bullet}{1}\underset{\bullet}{0}0\underset{\bullet}{1}11\underset{\bullet}{0}0\underset{\bullet}{1}\underset{\bullet}{1}0\underset{\bullet}{0}0\underset{\bullet}{1}\underset{\bullet}{1}\underset{\bullet}{1}, \\ e &= \underset{\bullet}{0}\underset{\bullet}{0}\underset{\bullet}{1}101\underset{\bullet}{1}0\underset{\bullet}{0}0\underset{\bullet}{1}1\underset{\bullet}{0}\underset{\bullet}{1}10\underset{\bullet}{1}0\underset{\bullet}{1}1\underset{\bullet}{1}\underset{\bullet}{0}\underset{\bullet}{0}0\underset{\bullet}{0}\underset{\bullet}{0}\underset{\bullet}{0}0\underset{\bullet}{1}\underset{\bullet}{1}\underset{\bullet}{0}0\underset{\bullet}{1}11\underset{\bullet}{0}0\underset{\bullet}{1}\underset{\bullet}{1}0\underset{\bullet}{0}0\underset{\bullet}{1}\underset{\bullet}{1}\underset{\bullet}{1}, \end{aligned} \tag{3.4}$$

where a bullet mark indicates a bit known by Eve. Alice and Bob then do the following. They pair up the bits in their string; the first bit with the second bit, the third with the fourth and so on. From each pair, $\{b_i, b_{i+1}\}$, they produce the new bit: $b'_{(i+1)/2} = b_i \oplus_2 b_{i+1}$ (equivalent to b_i XOR b_{i+1}). Eve will only know the value of this new bit if she knows *both* b_i and b_{i+1}. Only for 1 in 4 pairs will both bits be known by Eve. Therefore although the length of b' is half of b, the proportion of bits known by Eve has also been halved. In terms of mutual information, we can say that if p is the proportion of the key bits that Eve knows for sure, then

$$I(b;e) = \frac{1}{2}\left[(1+p)\log_2(1+p) + (1-p)\log_2(1-p)\right]. \tag{3.5}$$

The privacy amplification effected by the pairwise bit addition will therefore have reduced Eve's mutual information with the key from 0.189 bits to 0.046 bits.

We can therefore characterise the behaviour of a QKD system by the *secret key rate*, K, which is the rate at which bits of a secure bit string are produced from operations on a possibly insecure string. Here, security is defined in the following box.

Security definition

We quote here directly Gottesman, Lo, Lütkenhaus and Preskill: [GLLP04]

"A quantum key distribution protocol is secure if for any attack by Eve that passes the verification test with a probability that is not exponentially small, with high probability Alice and Bob agree on a final key that is nearly uniformly distributed and Eve's information about the final key is exponentially small. Here "exponentially small" means bounded above by e^{-CN} where N is the number of signals transmitted in the protocol and C is a positive constant, "high probability" means exponentially close to 1, and "nearly uniformly distributed" means with a probability distribution exponentially close to the uniform distribution."

If we let R be the rate at which Alice sends photonic qubits to Bob, known as the raw rate, then we can say that $K \leq R$, where the inequality is saturated only in the case where the channel between them has perfect transmission, with neither attenuation nor errors occurring to photons in transit. Note that some authors may use slightly different definitions. These may include defining R as the rate at which Bob receives photons, or defining the key rate as K/R. The shorter, secure key is produced by performing some privacy amplification, either classical or quantum, upon the received bits. An example of a classical algorithm is that which we have just seen, whereby the information in multiple bits can be combined to form a more secure bit. Privacy amplification may also be done at the level of the quantum state, before measurements are made to determine their bit. If two entangled pairs are distilled to produce a higher fidelity pair, then by the monogamy of entanglement, this reduces the mutual information that Eve may have with any future bit string produced from measurements made on the resulting state. The first proof that BB84 was secure against general attacks was provided by Lo and Chau [LC99]. This was simplified by Shor and Preskill [SP00], in a paper in which they present the first general expression for K for the BB84 protocol (or comparable entanglement-based protocols) as function only of the overall error rate, ϵ. Here we present the results of Gottesmann et al. [GLLP04],[4] who extended the above work and analysed the behaviour for realisations of BB84 with basis-dependent imperfections — a distinction that will be relevant for Chapter 4. Suppose that Alice and Bob have distributed between them a series of noisy $|\Phi^+\rangle$ Bell pairs. They want to check the error rates to

[4] Note, this paper is an excellent and comprehensive resource for those wishing to understand the security of the BB84 protocol.

ensure that they actually hold perfectly entangled pairs, or equivalently, that no interference has occurred in their transmission, whether malicious or accidental. They choose some pairs, and they each measure their part in the Z basis, and calculate $Z_A \times Z_B$ for these pairs. They say that the fraction of pairs measured in this basis that give the result -1 is ϵ'_{bit}, since this implies a bit-flip error has occurred and they actually held $|\Psi^+\rangle$ or $|\Psi^-\rangle$. Similarly, the fraction for which $X_A \times X_B = -1$ is called ϵ'_{phase}, since a phase error implies the presence of $|\Phi^-\rangle$ or $|\Psi^-\rangle$. It can then be shown that, for any $\delta > 0$, there exists a number of bits n such that an error-correcting code exists that enables secure communication at the rate

$$K = R\left[1 - h_2(\epsilon_{\text{bit}}) - h_2(\epsilon_{\text{phase}})\right], \tag{3.6}$$

where $\epsilon_{\text{bit}} = \epsilon'_{\text{bit}} + \delta$, and similar for ϵ_{phase}. This is then the highest rate at which Alice and Bob can communicate without sacrificing security and allowing Eve a non-negligible mutual information with the key.

Eq. 3.6 is an incredibly important equation for the field of QKD, and we will use it heavily throughout this thesis. The function $h_2(x)$ is known as the binary information. It characterises the amount of information in a string of 1s and 0s, where a proportion x of them are 0 (or equivalently, a proportion x are equal to 1), and is therefore equal to

$$h_2(x) = -x\log_2(x) - (1-x)\log_2(1-x). \tag{3.7}$$

Let us examine the terms in Eq. 3.6. The pre-factor of R is expected and simply understood; when Alice sends photons at twice the rate, she and Bob will generate a secret key at twice the rate. The interesting part is the amount by which the key has to be reduced in order ensure that Eve does not know it. To understand this, we should make note of another interpretation of $h_2(\epsilon)$. It may be seen as a measure of *mutual information*. If Eve tries to estimate a bit string, but makes an error when estimating a fraction ϵ of the bits, then her mutual information with the string (as may be calculated by the box in Section 3) is $1 - h_2(\epsilon)$. The two negative terms therefore represent knowledge that has been lost to the environment, or equivalently, to Eve (since we must make the pessimistic assumption that Eve is in control of the whole universe outside of the QKD system). Specifically, it is the case that one term (ϵ_{bit}) is the amount that has to be sacrificed to perform error correction on the bits on which Alice and Bob disagree, and one term $(\epsilon_{\text{phase}})$ is the amount that has to be sacrificed to perform privacy-amplification.

The form of K given in Eq. 3.6 is relevant when Eve has some prior knowledge about the choice of basis used by Alice and Bob, which makes it particularly relevant for analysing *side-channel attacks* (Section 3.1.5 and Chapter 4). However, in many cases we might make the usual assumption that Eve has no special knowledge about the basis, and that the bit and phase error rates are equivalent. In such a case, Eq. 3.6 reduces to

$$K = R[1 - 2h_2(\epsilon)], \tag{3.8}$$

where $\epsilon = \epsilon_{\text{bit}} = \epsilon_{\text{phase}}$.

The secret key rate is the key figure of merit for any QKD protocol or form of analysis. Discovering tight bounds on K is the aim of analytical theorists, increasing the bounds is the aim of all protocol designers, and realising high K values is the aim of all experimentalists and component manufacturers. It is a great convenience often not found in other disciplines that a single variable is ubiquitous as a measure of performance across the field, and as such its usage may extend beyond the strict application of QKD security measurement for which it was created. It is useful for any situation in which one wishes to generate entangled pairs as fast as possible. Quantum key distribution is one such application, but this is also important for quantum secure direct communication [LLY06], quantum secret sharing [HBB99], quantum authentication [BCG+02] and quantum computation [BBD+09]. The secret key rate provides a natural measure for judging the performance of entanglement-distributing protocols that encapsulates both the rate of generation and the fidelity of the states.

A commonly considered case is one where Alice and Bob share a set of entangled Werner states, each with fidelity F. Then, $\langle\epsilon\rangle = 2(1-F)/3$. However, this leads us on to a natural question: "what if $\epsilon \neq \langle\epsilon\rangle$?" This may be the case, since ϵ is the *fraction* of errors that actually occur, and not the *probability* that such an error will occur. Therefore, ϵ is actually a random variable. Fortunately, this does not greatly affect the analysis as long as the following condition holds: For any $\delta > 0$, there exists a number of bits, n, such that the probability that the inequality $\epsilon < \langle\epsilon\rangle + \delta$ is *broken* is exponentially small in n. This condition holds in almost every conceivable circumstance. Typically, QKD protocols are analysed in the regime where $n \to \infty$, in which limit $\epsilon' = \epsilon = \langle\epsilon\rangle$. Therefore, for the remainder of this thesis, we shall not make the distinction between these subtly different quantities. We shall simply use ϵ to mean the calculated probability that an error occurs.

3.1.4 Double-heralding

We have learned in the preceding subsections that entanglement states, and in particular Bell pairs, are of foundational importance to QKD. Once these entangled pairs between communicating parties have been established, there are many routes that we can investigate regarding ways to improve a QKD system, such as analysing non-standard attacks (Chapter 4), extending QKD to more than 2 parties [WZM+16], chaining together QKD systems into a repeater network for long-range communication (Chapters 5 and 6), or even reaching long ranges *without* the use of repeaters [LYDS18]. However, the necessity for reliable, *high-quality* entangled pairs underlies every other facet of QKD, and indeed quantum communication generally.

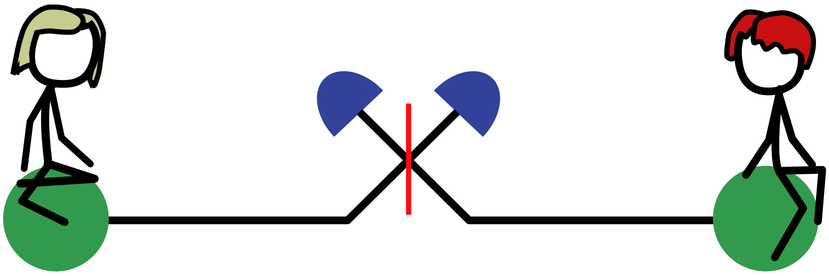


FIGURE 3.3: The physical set-up for both Cabrillo et al.'s protocol, and Barrett and Kok's (double-heralding) protocol. Green dots are photon emitters with level structures described in the main text, black lines are optical channels, the red line is a beam-splitter, and blue shapes are photon detectors.

Numerous methods for the creation of Bell pairs have been devised, although many of these face significant issues. For example, the technique of Cabrillo et al. [CCGFZ99] is a prototypical example of emission-based entanglement procedures. Two atoms are prepared with 3 relevant electronic levels, which we will denote $|0\rangle, |1\rangle$ and $|e\rangle$. States $|0\rangle$ and $|1\rangle$ are ground-state levels with slightly different energy levels, and $|e\rangle$ is an excited level. Both emitters are initialised in state $|0\rangle$, and then stimulated with a weak laser pulse that has a very small chance to cause the transition $|0\rangle \leftrightarrow |e\rangle$. An excited system will then relax back to a ground state over a time-scale τ_q, causing the emission of a photon of either wavelength $\lambda_{0,q}$ or $\lambda_{1,q}$. The emitted photon(s) are sent through optical channels, through a beam-splitter, and into two detectors tuned to detect photons of wavelength $\lambda_{1,q}$. If a single photon is detected, and the users are *sure* that there are no photons that were emitted and not detected, then the resultant state of the system will be $|\Psi^+\rangle$ or $|\Psi^-\rangle$, dependent on which detector detected the photon. The entanglement is generated by the fact that the beam splitter quantumly-erases the information indicating the emitter from which the photon originated. This set-up is illustrated in Fig. 3.3.

The drawback of this state is the fact that it relies on having a high confidence in the claim that no other photons were emitted but lost. This confidence may be increased by decreasing the laser power, which decreases the probability that a photon is emitted, ϵ_{emit}. The ratio of the probability for one photon to be emitted to the probability for two to be emitted goes with $1/\epsilon_{\text{emit}}$. However, as ϵ_{emit} is decreased, so does the success rate of entanglement. The aforementioned confidence may also be increased by ensuring that the photo-detectors used have extremely high efficiency (probability that a single incident photon is registered) as well as perfect multi-photon resolution — a set of requirements beyond current experimental capabilities.

Shortly after the publication of this work, Bose et al. [BKPV99] showed the versatility of emission-based quantum-optical protocols, by showing how this same technique can be used in order to teleport an arbitrary qubit state. However, this protocol suffers from the same drawbacks as the former. A protocol by Duan and Kimble [DK03] improves upon this. It uses emissions into two different polarisation modes. After passing these through a polarising beam splitter, it gives a Bell state when a single photon is detected of each polarisation. This is not sensitive to detector inefficiencies, however it does rely on a very carefully constructed level structure. In particular, it requires that there be four transition energies (two for each polarisation, and two for each emitter) that are perfectly degenerate. This also presents an experimental challenge.

The process known as *double-heralding* [BK05] aims to address such shortcomings. It is a protocol that allows for a Bell pair to be generated in a way such that the fidelity does not rely on the detector efficiency, η_D or optical transmissivity of the fibres, η_T (which will henceforth be referred to by single efficiency variable, $\eta = \eta_D \eta_T$), although as with the Duan-Kimble protocol, the success probability does drop off quadratically with efficiency. Additionally, there is no trade-off between the success probability and the fidelity. The fidelity's independence of these quantities allows double-heralding to produce Bell pairs that greatly exceed the fidelity threshold of 83%[5] required for secure communication through BB84. The high fidelity of the states also allows this protocol be *scalable*. That is, many of them can be used in conjunction and the error rates will not quickly compound to high levels, making the protocol ideal for large-scale quantum technologies. As a result of its favourable fidelity, double-heralding was used to generate entanglement for the first loop-hole free test of Bell's inequality in Delft [HBD+15], which was the first experiment to prove *conclusively* that the correlations resultant from measurements of entangled states cannot be explained with classical physics.

The protocol works as follows. Consider a photon-emitting system with two low-energy states, $|0\rangle$ and $|1\rangle$, and an excited state $|e\rangle$. Examples of such systems that exhibit such a structure are NV centres in diamond ([JGP+04] and Section 5.1.1) and quantum dots with excess electrons [PBC+03]. These are constructed such that an optical π pulse causes the transition of $|1\rangle \to |e\rangle$, whilst $|0\rangle \to |e\rangle$ is forbidden. The energy gap to $|e\rangle$ should also be sufficiently large that the system can be considered to be a qubit when unexcited. After excitation the system relaxes back to $|1\rangle$ over a decay time-scale τ_q, emitting a photon into an optical cavity in the process. This system is described by the Hamiltonian

$$\hat{H} = \sum_{j=A,B} \left[\frac{g_j}{2} \left(|1\rangle_j \langle e| \, c_j^\dagger + \text{H.c.} \right) - i\kappa_j c_j^\dagger c_j \right], \tag{3.9}$$

[5] At which point $1 - 2h_2[2(1-F)/3] = 0$. The quantity $2(1-F)/3$ represents the probability of making a bit error, averaged across Z and X basis measurements, for a Werner state of fidelity F with respect to $|\Phi^+\rangle$.

where g_j is the Jaynes-Cummings coupling strength [SK93] between the $|1\rangle_i \leftrightarrow |e\rangle_i$ transition and corresponding mode of the optical cavity and κ_j is the energy of the photons. We may express the decay constant in terms of these parameters as $\tau_q = (\kappa - \sqrt{\kappa^2 - g^2})^{-1}$ (where we have set $g_A = g_B = g$) [BK05].

In order to use this to establish a Bell pair, consider two emitters, each initialized in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, giving an overall system state of

$$\begin{aligned} |\psi\rangle &= |+\rangle \otimes |+\rangle\,, \\ &= \frac{1}{2}\left(|0,0\rangle + |0,1\rangle + |0,0\rangle + |1,1\rangle\right). \end{aligned} \tag{3.10}$$

These emitters should both be stimulated by a π pulse as described above, and any outputs sent through an optical channel and through a beam-splitter. This round of the protocol is considered a success if we detect exactly one photon. The experimental set-up for this is the same as for the Cabrillo protocol, and is also shown in Fig. 3.3.

If it were somehow the case that $\eta = 1$, and we were sure that the detection of one and only one photon implied the emission of one and only one photon, then our system would now be in state $|\Psi^+\rangle$ or $|\Psi^-\rangle$ (dependent on the detector that clicked). Similar to the case of Cabrillo's protocol, the beam-splitter would have erased the information as to which emitter was the source of the photon, and the resultant state would be a coherent superposition of both possibilities.

However, since in any real-world setting it is always the case that $\eta < 1$, especially when considering multiple-photon-counting applications, there exists the probability that the state was mapped to $|1,1\rangle$, and one photon was simply lost. The detection of a single photon therefore implies the system is in the mixed state

$$\begin{aligned} \rho &= \frac{\eta\frac{1}{2}\,|\Psi^+\rangle\langle\Psi^+| + 2\eta(1-\eta)\frac{1}{4}\,|1,1\rangle\langle 1,1|}{\mathrm{Tr}\left[\eta\frac{1}{2}\,|\Psi^+\rangle\langle\Psi^+| + 2\eta(1-\eta)\frac{1}{4}\,|1,1\rangle\langle 1,1|\right]} \\ &= \frac{1}{2-\eta}\left(|\Psi^+\rangle\langle\Psi^+| + (1-\eta)\,|1,1\rangle\langle 1,1|\right). \end{aligned} \tag{3.11}$$

Therefore, after waiting a suitable length of time for any excitations to relax, we apply a Pauli X gate to each emitter, mapping $|0\rangle \leftrightarrow |1\rangle$, and again excite both of them. If, after this second excitation and emission, we again measure a click in one of the detectors, we know that we could not have had $|1,1\rangle$ in the first round, because that would map to $|0,0\rangle$ in the second round, which will result in no photoemission. Therefore, we will have projected onto the *pure* Bell state $|\Psi^+\rangle$ (or $|\Psi^-\rangle$, depending on which detector clicked). If we fail to detect a photon in either step, both emitters are reinitialized to $|+\rangle$ and the process is repeated until a connection is formed.

The greatest benefit of this method is that, unlike other schemes for remote entanglement generation, the fidelity of the final pair is not affected by attenuation in the channel or imperfections in the detectors. The fidelity is also completely unaffected by the presence of multi-photon components (although the success probability is affected). However, there are still ways in which the fidelity may be affected. Dark counts may cause us to falsely believe that we had measured a signal emission, reducing the fidelity. This is analysed in Section 5.3.1. Mismatching of the parameters κ and g between the two emitters will mean that the photons retain an element of distinguishability after the beam splitter, and so detection will result in imperfect path erasure and again a reduced fidelity. Additionally there is the problem of decoherence of the emitters, which will be analysed further in Section 5.3 and Chapter 6.

3.1.5 Side-channels

In Section 3.1.3 we described how, by measuring the average error rate in a sample of bits, Alice and Bob can ensure that their communication is exactly secure. If the error rate is less than 11%,[6] this is done by producing, from the raw key, a smaller key that is completely secure. If the error rate is above 11% then the communication is abandoned.[7] Does this mean then that Eve is forever thwarted? It does not! As we briefly alluded to earlier, the standard security proofs only hold for standard attacks, where Eve intercepts the flying qubits and operates upon them. We must assume that Eve is wily and cunning, and will seek alternative avenues of attack. . .

In general, a *side-channel* is any method of covertly learning secret information that does not primarily involve directly attempting to read the transmitted signal. This can apply to communication side channels (as we will focus on here), but also to computational side-channels, where, for example, the information processed by CMOS devices can be read by measuring their noise outputs [DV13, MPG05], or even their temperatures [HS13]. The first modern cryptographical side-channel that was known to be exploitable was the timing attack, demonstrated by Kocher [Koc96]. Here, the runtime of key distribution algorithm (such as RSA) could be measured and used to infer the content that was being transmitted.

[6] This is the value of ϵ for which K, given by Eq. 3.8, equals 0.

[7] The absence of a communication attempt is trivially secure.

Historical side-channels

It is interesting to note here that just because a side-channel has been discovered and protected against in one circumstance, does not mean that future cryptographers will remember to do so. We say this because the fundamental notion behind Kocher's attack (or rather, its reciprocal — that the frequency of a process or message may disclose its content) was known as early as World War II. When messages encoded by the enigma machine were sent by the Axis powers, it was mandated that the same number of messages be sent each day. Otherwise, an Allied force that intercepted them may be able to discern the coming of a large offensive by an increase in the frequency of messages, even if they could not decode them. This tale, however, highlights an important maxim that should act as a warning for all modern cryptographers: *active defences against side-channels may themselves become new side-channels*. Indeed it was the case that the Axis defence against the variable-length-message side-channel itself became a source of information leakage when it was combined with the weak point that underlies every cryptographic system: human laziness. In one case, an Axis soldier who was sending a message to fill the daily quota simply sent a message that, when decrypted, consisted only of the letter "L" many times. Attentive cryptographers at Bletchley Park noticed that the coded form of this message did not contain the letter "L," since the enigma machine never mapped a letter to itself. This oversight allowed the code-breakers to decode other messages from that day [Thi16]. This also highlights how an encoding system that misses out on cryptographic security by only a small amount, where the entropy is slightly less than maximal, can still be exploited and broken.

Moving specifically to quantum cryptography, side-channel attacks (SCAs) are an important and active area of investigation. It is within this field that the discovery of a side-channel could have the most devastating consequences, simply because QKD is the only type of cryptography for which the protocols have information-theoretic security against standard attacks, making SCAs necessarily the only avenue of attack for Eve. Mirroring the classical case, one of the first QKD SCAs to be demonstrated was the timing attack [LLK07]. Shortly after this, Nauerth et al. [NFSM$^+$09] analysed information leakage in free-space QKD, noting that the spacial and spectral distributions of pulses are possible avenues of attack for Eve.

One class of SCAs is known as photon-number-splitting attacks. These exploit the fact that the signals sent by Alice may often not be true single-photon states, but instead weak coherent pulses with a low but non-zero probability to contain more than one photon. By passing the flying qubits through a beam splitter, Eve could measure one photon, and send the rest on to Bob who would receive them unaware of the attack. This SCA has been comprehensively dealt with by the concept of *decoy states*, introduced by Hwang [Hwa03] and analysed by Wang [Wan05], amongst others.

Progress has also been made in devising communication protocols that completely remove side-channels that target the physical devices involved. These fall under the labels of measurement-device-independent QKD (MDI-QKD) [POS$^+$15, LCQ12, TYC$^+$14] and fully device-independent QKD (DI-QKD) [MPA11, BP12, Cur12]. While these eliminate a large class of side-channels, they typically do so at the expense of a much lower secret key rate and greater complexity.

In Chapter 4 we will examine the specific case of a side-channel attack known as the Trojan-Horse Attack. It is there (particularly Section 4.1.2) that we will investigate the effect that the presence of a side-channel has upon the key rate that may be achieved.

3.2 Quantum repeaters

By now you should be convinced that quantum key distribution has great potential to enable communication of unprecedented security. However, it has, in its current form, a significant hurdle to overcome. The qubits of discrete-variable QKD are encoded in single photons, and I'm sure the reader will appreciate that single photons are easily lost. When photons are sent down an optical fibre, they typically suffer an attenuation that is exponential with distance. That is to say, the probability that the photon will *not* have been lost after travelling a distance L is given by $\exp(-L/L_{\text{att}})$, where L_{att} is the attenuation length of the fibre, typically around 25km [KLH$^+$15]. Therefore, when Alice and Bob wish to communicate over a distance of more than a few tens of kilometres, their raw key rate will soon drop to zero if the photons are sent directly. More importantly, in many quantum communication and entanglement generation protocols, the probability that an error will accumulate will increase with distance, especially if imperfect local operations are used. We therefore require a system that can fulfil two roles: to boost the raw key generation rate, and to ensure that the error rate stays below the communication threshold of 11%, or some other cut-off that the experimenters desire.

In classical communication, this is overcome with the use of a repeater. This is simply a station positioned between Alice and Bob that amplifies the incoming signal. When we are using quantum signals, however, this cannot be done. This is due to the *no-cloning theorem* (fundamental concept conceived in [Par70], described more explicitly in [Die82, WZ82]), which states that it is impossible to create a perfect copy of an arbitrary and unknown quantum state.[8] Instead, the channel is divided up into multiple sections, with each section bookended by a *quantum repeater*. These quantum repeaters should extend the range of the signal, without directly cloning and amplifying it.

[8] The *teletransportation paradox* is an old problem in the philosophy of identity, to which we may consider this fact a partial solution. It considers a future world where fully-functional teleporters have been developed. These act by deconstructing a person at one location and reconstructing them atom-by-atom at another from a reserve bank of carbon, nitrogen, oxygen atoms etc. If I am identified with my material form, then I should have no problem with this. From my perspective it should be no different to going to sleep and waking up — the clone will experience a continuation of consciousness, and will in all senses, be *me*. Derek Parfit [Par84] asked what should happen if the machine scanned the location of every atom in my body, but then malfunctioned: it might fail in deconstructing me, but succeed in building my clone, leading to two simultaneous copies of me. We should then ask where my sense of *self* resides. If I am my material form, then the clone is just as much me as the original. By this argument, the original should have no particular qualms with being murdered after stepping out of the faulty teleporter. After all, he was going to be deconstructed anyway! However, we can imagine that most people will feel uncomfortable with this step. The no-cloning theorem could address this quandry. If the human brain relies on coherent quantum processes and superpositions (a contentious, but plausible and seriously considered hypothesis, [Pen91]), then any fully faithful teleporter system would have to teleport these coherently. While it is well known that an arbitrary unknown state may be teleported, it can not be cloned. Therefore, the paradox of two perfect copies of myself (and so two copies of my *self*) can never occur, if the teleportation is faithful down to the quantum level.

No-cloning theorem

This proof is found in Nielsen and Chuang [NC02], page 24.
Suppose there existed a "state-cloning unitary operation," $\hat{U}_{\text{SC}}$, that mapped

$$\hat{U}_{\text{SC}}\left(|\psi,0\rangle\right) \rightarrow |\psi,\psi\rangle$$

for all $|\psi\rangle$ in the qubit Hilbert space. Consider generally that $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Since unitary operations act linearly on inputs,

$$\begin{aligned}\hat{U}_{\text{SC}}\left(|\psi,0\rangle\right) &= \alpha\hat{U}_{\text{SC}}\left(|0,0\rangle\right) + \beta\hat{U}_{\text{SC}}\left(|1,0\rangle\right)\\ &= \alpha\,|0,0\rangle + \beta\,|1,1\rangle\,.\end{aligned}$$

However, the above expression is not the same as the cloned state $|\psi,\psi\rangle = \alpha^2\,|0,0\rangle + \beta^2\,|1,1\rangle + \sqrt{2}\alpha\beta\,|\Psi^+\rangle$. Therefore, there cannot exist a unitary that clones an arbitrary quantum state. The more general proof that there exists no quantum circuit (whether unitary or not) that achieves this task is given on page 530 of Nielsen and Chung [NC02].
Note that the unitary can always clone classical information. If the qubits are in fact bits, and $\alpha, \beta \in \{0,1\}$, then the unitary clones the state perfectly. It is the quantum nature of the qubit that prevents cloning, not the nature of the task itself, as evidenced by the fact that classical repeaters are far simpler to create than quantum repeaters.

It is critical that the quantum repeater (or hereafter simply a *repeater* for short) is able to increase the range of the communication without learning the state or the value of the bit. This is because we do not want to have to trust the intermediaries — the protocol is most secure when the amount of trust required of external parties is at a minimum. A long quantum network for intercontinental communication may require dozens of repeater stations, each one of which would become a target of attack if they had access to the key.

The most basic form of repeater is based on entanglement-swapping. The distance between Alice and Bob is divided into N_S sections, where each section consists of a set of optical fibres linking Alice to a repeater station, a repeater to another repeater, or a repeater to Bob. The distance between each pair of stations is assumed to be a constant L_0 for simplicity. Alice and Bob will use this network, along with the E91, or entanglement-based, protocol (Section 3.1.2) in order to generate a key. Each section consists in fact of a number of parallel optical channels, linked to a quantum memory at either end. Some elementary entanglement-generation protocol should then be chosen, and used to attempt to generate entanglement between each pair of memories, using quantum signals transmitted by photonic states sent along the fibres. At some point after at least one of the entanglements on either side of a station have been established, entanglement-swapping operations at the station then connect together these states to form entanglement over a longer distance. When such swapping operations have completed at all stations, then Alice and Bob should share a Bell pair, which can be measured to produce a Bell pair. From this general model, a number of different repeater proposals have been developed. These may vary depending on the specifics of their implementation. For example, different proposals may implement different measures to reduce errors, such as distillation [BDCZ98] or fault-tolerant encodings [LBSB13]. They may rely on different forms of quantum memories, such as atomic ensembles [SSDRG11] or NV centres [NTD+16], or different forms of initial entanglement generation [NTD+16].

3.2.1 Functionality

The way in which quantum repeaters can decrease Alice and Bob's measured error rate requires us to examine more closely the specifics of a repeater protocol, so we will leave that to Section 3.2.2. However, this basic model of an entanglement-swapping repeater is sufficient to explain how they can be used to bolster the raw key rate. The first way is by including redundancy in the number of connections between each pair of stations. Suppose that each section is spanned by q pairs of memories. The probability that any one section will produce at least one Bell pair that spans it in after one round of sending photons[9] is given by $1-(1-e^{-L_0/L_{\text{att}}})^q$. If at least one Bell pair is generated on each side of a repeater, then that repeater can simply apply entanglement-swapping to the appropriate memories, and form a longer connection. Therefore the probability that an end-to-end connection is established in the first round of photon-sending is

$$p(\text{conn with repeaters}) = \left[1-(1-e^{-L_0/L_{\text{att}}})^q\right]^{N_S} \tag{3.12}$$

If we were not using repeaters, but instead had q optical channels spanning directly from Alice to Bob, then the one-shot connection probability becomes

$$p(\text{conn without repeaters}) = 1-(1-e^{-N_S L_0/L_{\text{att}}})^q. \tag{3.13}$$

[9] This process takes a time $t = 2L_0/c$. The photonic qubits must be sent from one station to the next, taking a time L_0/c. Those stations must then send messages back the other way to indicate whether the qubit was received or lost, so the "sender" station knows whether to send another one.

Since $p(\text{conn with repeaters}) \geq p(\text{conn without repeaters})$,[10] a network which is divided up into sections which are entangled individually has a higher probability of producing an end-to-end connection than a network which only involves end-to-end channels, even when those channels also include redundancy. We can see that the redundancy plays a part here by setting $q = 1$, for which we recover $p(\text{conn with repeaters}) = p(\text{conn without repeaters}) = e^{-N_S L_0 / L_{\text{att}}}$.

The second way in repeaters may bolster the raw key rate is by their use of memories to store existing entanglement while other parts complete. We can consider a repeater network with $q = 1$, and ask, what is the probability that an end-to-end Bell pair will have been established after t time-steps, where one time-step is the time taken to attempt an entanglement creation operation over one elementary section. We can assume here that entanglement-swapping takes no time, which is a good approximation when considering the time-scales involved. We should also assume as a first approximation that the fidelities of quantum states held in the quantum memories do not decay over time.

If the channel is again divided into N_S sections with no redundancy, then at time t Alice and Bob will share a Bell pair at time t with probability

$$p(\text{conn with repeaters}) = \left[1 - (1 - e^{-L_0/L_{\text{att}}})^t\right]^{N_S}. \tag{3.14}$$

The comparable channel without repeater stations will produce entanglement in the same time with probability

$$p(\text{conn without repeaters}) = 1 - (1 - e^{-N_S L_0 / L_{\text{att}}})^t. \tag{3.15}$$

We can show that $p(\text{conn with repeaters}) \geq p(\text{conn without repeaters})$ by direct comparison with Eqs. 3.12 and 3.13, but with q replaced by t. The two cases may then be seen as expressing the spatial and temporal faces of a single conceptual advantage.

Of course, this duality only truly holds in a theoretical sense, and real-world considerations will break the symmetry. For example, relying on a large spatial redundancy requires us to be in possession of a large number of qubits, which may be expensive to buy or produce. There is also the non-trivial issue that atomic qubits at a repeater must be arranged in some kind of physical spatial formation. Some of these qubits will therefore be adjacent, and some may be far apart, which may cause issues as q grows very large.

There is also the issue of the compounding infidelity resulting from local operations. The quantum operation that a repeater applies to a pair of qubits to cause a swapping of entanglement would have some probability to fail, and thus using

[10] This is difficult to prove by a manipulation of the equations, but easy to prove when you consider their interpretations. We may consider the repeater-less system as a special case of the repeater network, but one where the entanglement swaps that may be made at a given station are *fixed*. i.e., the repeater may only connect the topmost quantum memory on its left with the topmost memory on its right, and the second memory on its left to the second memory on its right, and so on. Clearly when we remove this restriction the connection probability can only increase.

more repeater stations will lower the fidelity of the final state. However, even if the local operations are perfect, the entanglement-swapping operations between two states that are approximately Bell-pairs can amplify any imperfections in them. For example, suppose a repeater section wanted to connect a Werner state on its left to a Werner state on its right, each with fidelity F. If this is done by the standard method of applying a CNOT gate to the middle states, then measuring them in the X basis (as discussed in Section 2.2.3), then this would result in a state with fidelity $(4F^2 - 2F + 1)/3$, which is less than F when $F > 0.25$. [11]

These issues that affect the final fidelity of the quantum states shared by a repeater network, or equivalently the security of the bit strings distributed using them, are exacerbated by the probabilistic nature of the initial entanglement-creating operations between the stations. We have already seen how a repeater can help overcome this, by the use of quantum memories, although such memories are not perfect. Typically, the probability that they will *not* have accrued an error decays exponentially with time, meaning the states that they store will have a fidelity of

$$F = \frac{1 + 3e^{-t/\tau}}{4}, \tag{3.16}$$

for some coherence time τ. Therefore, while the probability of establishing entanglement increases with additional sections, the fidelity of the resulting states may not. We therefore must use the secret key rate which acts to combine both metrics, to determine what set-up is best for any given set of experimental parameters. This is particularly true when each elementary entanglement-creation operation succeeds with a low probability, since the memories will have a high probability of accumulating an error before the protocol completes.

This presents a difficulty in our analysis. We cannot say that some protocol completes in some exact time, with some exact error rate. Instead we will have a distribution of completion times, and at each completion time we will have some distribution of error rates. Due to the highly non-linear nature of many of the equations involved in proving the security of BB84, as well as the need to give lower-bounded guarantees on secret key rates and not simply averaged values, we cannot ignore such statistical contributions to the behaviour of the protocols. In Chapter 6 we shall develop the statistical tools required to describe this, and use them to analyse this interaction between memory lifetimes and entanglement success probabilities in detail.

[11] If the Werner state is caused by a stochastic loss of information from an ideal version of a target state, then the fidelity will not drop below 0.25. This is because the completely mixed state has a fidelity of 0.25 with respect to each Bell state.

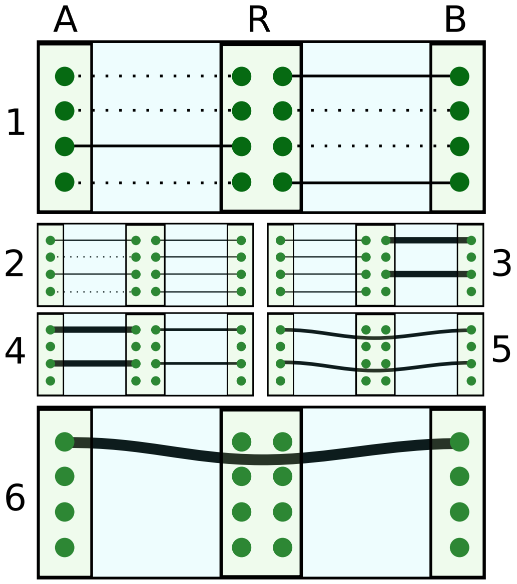


FIGURE 3.4: Innsbruck quantum repeater scheme with 2 sections. Part 1 shows a situation where some pairs of memories have established entanglement, where A is Alice, R is a repeater and B is Bob. Circles show quantum memories dotted lines show no entanglement and solid lines are successful entanglement connections. In parts 2–3, the right-hand section has completed, and distilled down to 2 pairs. The thickness of the lines between memories shows the fidelity of that entangled state. In parts 4–5, the left-hand section is also distilled and these states are connected to the right-hand pairs by entanglement swapping. In part 6 the two pairs between Alice and Bob are distilled again to form a single high-fidelity pair.

3.2.2 Innsbruck protocol

The first quantum repeater protocol to be devised was that of Briegel et al. in 1998, [BDCZ98], which we shall refer to throughout as the *Innsbruck protocol*. The fundamental idea is the same as with many repeater proposals. This is illustrated in part 1 of Fig. 3.4.

When the entangled states that bridge a given section have all established, then these are distilled down to a smaller number of higher-fidelity states (parts 3 and 4 of Fig. 3.4). The sections are then arranged into pairs, which are connected together by entanglement swapping, to form entangled pairs over a distance $2L_0$. We may consider each pair of sections as then a higher-level "section" in its own right, since it is now pairs of sections that are bridged by entangled pairs. These states are then again distilled, and the new higher-levels are paired up and connected by entanglement swapping, forming new "sections" of length $4L_0$. The distance that is spanned by any one Bell pair therefore doubles with each round of distillation and entanglement connection. Note that the inclusion of distillation is critical here. Without it, the repeated use of entanglement-swapping operations (whether perfectly faithful or not) would soon result in the fidelity of Alice and Bob's final state to drop to useless levels.

3.2.3 Other repeaters

The system described above, where quantum memories stationary in a repeater node are used to interface with flying qubits before applying entanglement-swapping, is possibly the simplest and most well-known general model of a quantum repeater. Our own protocol, described in Chapter 5, falls into this category. In addition to other protocols already mentioned, one noteworthy pioneering example is the DCLZ protocol [DLCZ01]. Here, they give one of the first full operational descriptions of a quantum repeater, involving atomic ensembles to interface with the photonic qubits. They also highlight another explanation for the ability of a quantum repeater to extend the range of communication: It is noted that a repeater network between Alice and Bob may be thought of as a *single* quantum channel, but one that features measurements of the qubit states along the way. In a way analogous with the quantum Zeno effect [IHBW90], these repeated measurements act to reduce the uncertainty in the state and extend the range of its information.

The full gamut of quantum repeater protocols extends to ideas outside the basic model we have described here. Some of these aim to overcome certain current experimental shortcomings, while others go in the opposite direction, aiming to deliver higher secret key rates, but at some point in the future when experimental capabilities have caught up with the demands of the protocol. However all of them broaden our understanding of what is and what may be possible with quantum technologies. We briefly discuss two of these here, while others are discussed in Section 5.5.

Coherent light repeater

In Section 3.1.5 we highlighted how photonic pulses containing more than one photon could be exploited by Eve to learn the key without alerting Alice and Bob. As such, the signal pulses in such a protocol are typically limited to having an average photon number of almost exactly one. Unfortunately, this means that the signals are easily lost. The team of van Loock et al. [VLLS$^+$06, vLLMN08] went instead in the other direction, and suggested a quantum repeater protocol that uses bright coherent pulses of light. They call it a hybrid repeater, since it in fact makes use of both discrete and continuous quantum variables.

Here, the flying qubits are not encoded in a 2-dimensional Hilbert space at all. Instead, they are encoded in the phase of a coherent state, that is rotated conditional upon the qubit state of a memory. Given a qubit state at one station $|+\rangle$ and a coherent state $|\alpha\rangle$, the resultant state would be transformed by a phase shift of $-\theta$ as

$$U_\theta \left[\frac{1}{\sqrt{2}} \left(|0\rangle + |1\rangle \right) |\alpha\rangle \right] = \frac{1}{\sqrt{2}} \left(|0\rangle |\alpha\rangle + |1\rangle \left| e^{-i\theta}\alpha \right\rangle \right). \tag{3.17}$$

This goes on to another station where it interacts with another qubit in a similar way. After applying a final phase shift of $+\theta$, the resultant state will then be

$$|\psi\rangle = \frac{|\Psi^+\rangle |\alpha\rangle}{\sqrt{2}} + \frac{|0,0\rangle |e^{i\theta}\alpha\rangle + |1,1\rangle |e^{-i\theta}\alpha\rangle}{2}. \tag{3.18}$$

Therefore, if one can perform a measurement (such as a homodyne detection) that can distinguish between an un-rotated coherent state, and one rotated by $\pm\theta$, then a measurement outcome indicating no rotation will project onto the maximally-entangled Bell pair.

This protocol has the advantage that it does not rely on single photons, therefore has the potential to extend to further ranges than single-photon-based networks. However, its communication efficiency suffers due to the fact that coherent states are not perfectly distinguishable, and large quantum-conditional rotations are difficult to achieve in practice. This leads to a situation where the maximum achievable fidelity for the generated Bell pairs is only around 80% (and depends on the detector efficiency), and the maximum success probability is just over 50%. To help abate these difficulties, the authors suggest using long-lived nuclear spins as quantum memories, which we discuss in Section 5.1.1.

We note here that the primary issue with this protocol is the low fidelities it creates, which are just on or below the minimum fidelities required by BB84 ($F = 0.8$ gives an error rate of 13% for Werner states). It is therefore likely that an increase in the fidelity of the generated Bell pair would result in a greater secret key rate, even if it came at the expense of a lower success probability. If we first put aside issues of technical capability, one way in which this could be achieved is by unambiguous state discrimination. Since we are only looking for the $\theta = 0$ result, we could construct a POVM with two measurement operators:

$$\begin{aligned} \hat{E}_{\checkmark} &= |\perp\rangle\langle\perp|, \\ \hat{E}_{\times} &= \mathbb{1} - |\perp\rangle\langle\perp|. \end{aligned} \tag{3.19}$$

Here, $|\perp\rangle$ is a vector that is constructed to be orthogonal to both $\left|e^{+i\theta}\right\rangle$ and $\left|e^{-i\theta}\right\rangle$. If the optical state is projected onto this projector, we know that the qubit state cannot be either $|0,0\rangle$ or $|1,1\rangle$, and so must be $|\Psi^+\rangle$. This will have a success probability of $|\langle\alpha|\perp\rangle|^2$. However, the difficulty of this method lies in constructing the state $|\perp\rangle$, both mathematically and physically.

All-optical repeater

This is a proposal by Azuma, Tamaki and Lo [ATL15] that aims to remove the need for quantum repeaters altogether. Here, they introduce the innovative idea of performing some of the entanglement-swapping *before* the long-range entanglement generation. Here, the repeaters are divided up into two types — *senders* and *receivers*. The senders prepare $2m$ loss-tolerant qubits,[12] an act which effectively acts

[12] A loss-tolerant qubit is one encoded in many photons, in such a way that the information can be recovered if some of them are lost. See [VBR06] for an example.

as a pre-prepared entanglement-swapping. Each of these is entangled with all other such qubits in the station, and each one is also entangled with another qubit known as a *second-leaf qubit*.[13] They then choose m of the loss-tolerant qubits and their associated second-leaf qubits, and send them to the receiving station on their left, with the remaining being sent to the station on their rights. Upon receiving these, the receiver stations perform entanglement-swapping on the second-leaf nodes, and then apply Y measurements in the appropriate place to remove extraneous qubits from the network.

This has the advantage of not relying on quantum memories. Any protocol which removes some component that was previously thought to be fundamental is one that holds the potential to enable a great advance in the field, since it eliminates an entire set of experimental design requirements. However, in this case, it seems that such a reduction in memory requirements has been met on the other side by an increase in entanglement-construction requirements, since the protocol involves very complex entangled photonic structures. Whilst significant progress has been in producing complex entangled states, with 6-photon GHZ states[14] having been made [LZG+07], and well as atomic entangled structures of up to 3000 atoms [MZH+15], it remains to be seen whether the photonic structures required here can be produced with a reliability that gives them an advantage over memory-based approaches, especially given the advances in nuclear-spin based memories (Section 5.1.1).

[13] This multi-particle entanglement is meant in the sense of graph-state entanglement as described in Section 2.2.3, and does not conflict with monogamy of entanglement.

[14] Of a form proportional to $|0\cdots 0\rangle + |1\cdots 1\rangle$, and equivalent by the application of local unitaries to the complete graph state.

Chapter 4

Defending against the Trojan-Horse side-channel attack

In Section 3.1.5, we introduced the idea of a cryptographical side-channel attack (SCA), and highlighted the danger that they can pose to the ability to securely communicate. One such SCA that has attracted recent theoretical [GFK+06, DLZZ05, LCW+15] and experimental [JSK+15, JAK+14, SRK+15] attention is the so-called *Trojan Horse Attack* (THA). Here, Eve will tap into the optical channel that Alice and Bob use to communicate. She will then send her own optical state into Alice's system, whereupon it will reflect off the same apparatus used to encode the legitimate photonic qubits. Having picked up some information on the encoding of the latest quantum state that Alice sent, it will return out and be measured by Eve. Eve will then use the result of this measurement, possibly combined with some operation on the legitimate qubits, to make a best estimate of the state that Alice sent to Bob, thus giving her some non-negligible information mutual with the key.

This attack has previously been analysed by Lucamarini et al. [LCW+15] They assume that Eve uses a coherent state (Section 2.1.4) to probe the system, and describe using a one-way attenuating filter at the entry-point of Alice's apparatus as a defence. The effect of this is to absorb the majority of light that is sent into the system, such that Eve receives far less than one photon back per attempt, reducing her ability to estimate the key bit. They make use of the theoretical framework of Gottesman et al. [GLLP04] to get an expression for the rate at which Alice and Bob can generate a secret key in the presence of such an attack. In this paper we make use of this same framework, but extend the range of powers of both Eve and Alice. The chapter is organised as follows. After discussing a few of the necessary preliminary notions, we begin in Section 4.1 by describing and analysing the effect of Eve performing a THA on the system, allowing her to use any Gaussian state including multimode entangled states. We prove that the (separable) coherent states are optimal amongst this class. As part of this analysis, we do *not* make the assumption that Alice's system is noise-free. In particular, we assume that there may be a thermal noise component to Alice's emitted signal. We analyse the effect that this can have upon the secret key rate, but we do so from the position of the strong assumption that Alice can control the average photon number of the thermal noise in the appropriate mode. We will see that this can be used as a defence in itself, and can more than double the range over which a QKD system can securely communicate. We do note, however, that this assumption may be a little strong for some readers' tastes[1] In this case, we are careful to highlight the results that follow from the first assumption but not the second. i.e. the case where Alice's system and her signal mode are not guaranteed to be noise-free, but she has no control over this noise.

Motivated by the revelation that entanglement does not assist Eve when using Gaussian states to attack the system, in Section 4.2 we restrict Eve to separable states. We derive a bound on the information that Eve may learn about the key when we allow her to use *any* separable state. Finally, in Section 4.3 we describe an active defence system that may be used in place of an attenuating filter. This relies on the causing Eve's THA light pulse to undergo many reflections before being returned.

In this work we will be assuming that Alice and Bob communicate via BB84, with qubits encoded with a phase-shift of the early mode of a time-bin-split photon, where the qubit is encoded by the parameter θ (as shown in Fig. 3.1). It is this parameter that Eve wishes to estimate. In order to do this, she prepares her own state, ρ. This is assumed to exist in the photonic Fock space of a single mode. The single mode assumption is justified since we may say that Alice will filter out all frequencies that are not equal to the one sent to Bob. It may also be assumed without loss of power or generality to be pure.[2] This state is sent into Alice's system. Here it passes through a filter which allows a fraction $\eta \ll 1$ of the light to be transmitted, resulting in a state ρ_η. It then reaches the polarizing filter, where it evolves according to the same Hamiltonian that encoded θ into the photon that was sent to Bob. That is to say, it is transformed as follows:

$$\rho_\eta \to \rho_\eta^\theta \equiv \hat{R}(\theta)\,\rho_\eta\,\hat{R}(\theta)^\dagger, \tag{4.1}$$

where $\hat{R}(\theta)$ is the rotation operator on the Fock space of Eve's photons, defined in Eq. 2.25. After Eve's state has picked up the phase information it returns to her. She then performs some operation to try to make an estimate of θ.

[1] In particular, we must choose precisely the right frequency and time modes to fill with thermal noise. We cannot fill every frequency in some range with a finite amount of thermal noise, since this would require a total of an uncountably infinite number of photons, which would surely destroy the system. Of course, Alice also cannot create a wave-packet that is infinitesimally narrow in frequency or time, so she could probably get away with adding thermal noise that has a frequency that lies within the spectrum of her signal if she wanted to ensure that Eve would not be able to perfectly separate her signal and noise. However, a detailed analysis of this is beyond the scope of this work.

[2] Note that a mixed state cannot improve Eve's predictive power. Suppose we have a state that is the classical mixture of two states given by $\rho = \alpha\,|\psi_1\rangle\langle\psi_1| + (1-\alpha)\,|\psi_2\rangle\langle\psi_2|$. Her resultant key rate will be $K = \alpha K_1 + (1-\alpha)K_2$, where K_i is the key rate of the pure state $|\psi_i\rangle$. This is simply optimised by $\alpha = 0$ or 1, depending on whether K_1 or K_2 is larger. By induction, we can see that no mixing may improve Eve's key estimation.

4.1 Gaussian-state attack

Attenuation-based defence systems, which aim to muddy the phase information on Eve's state by blocking most of the incoming attack state, have been previously analysed by Lucamarini et al. [LCW+15]. They show that in order to realise any appreciable level of security, a very high level of attenuation is required. Specifically it requires that Eve get back far less than one photon per attempt. We want to investigate whether it is possible to relax this requirement by implementing complementary security measures.

To this end, we will consider a system where Alice adds a small amount of thermal noise with average photon number μ_T into the signal she sends out to Bob. Since Eve taps into this same channel, some of this noise will also be picked up by Eve. It will be combined with her returned state ρ_η^θ to produce the state ρ_{η,μ_T}^θ.

When running a protocol such as BB84, Alice and Bob can only try to generate secret key bits from the attempts where Bob successfully received a signal. So after post-selecting on these bits Bob must receive at least one photon per bit generation attempt. On the other hand, we may choose the strength of the attenuator such that Eve receives much less than one photon per attempt. Therefore, if μ_T is less than one, but comparable to the average photon number of ρ_η, then the addition of this thermal noise is likely to affect Eve more than it affects Bob. Sections 4.1.2 and 4.1.3 quantify this.

4.1.1 State description

Here we will describe specifically how we construct ρ_{η,μ_T}^θ from ρ, and how Eve should choose ρ to maximise her knowledge of the key.

Firstly, it is clear that the choice of initial state ρ will have a significant effect on Eve's ability to discern θ. There are certain properties of this state that we can identify that we expect to affect this in varying degrees. The property that may be most apparent in its effect is that of the average photon number of the state. If Eve sends in a single photon, then given a high amount of attenuation, she is not likely to get much back and will not be able to reliably learn θ. On the other hand, if she is allowed to send in an arbitrarily bright state with unbounded average photon number it is clear that she will always be able to distinguish the different settings of θ perfectly. Therefore, to be able to implement any QKD protocol, the first step in protecting against a THA is putting some upper bound on the average number of photons that may pass into the system. This may be done by way of some defence such as an optical fuse [DCL91], which melts when sufficiently many photons pass through it, or by identifying some other component which will be irreversibly damaged when subject to a bright enough light [CRD03]. A more detailed examination of the numbers and figures behind such defences may be found in [LCW+15], but for our purposes we may simply assume that there does exist some bound N such that $\langle \hat{n} \rangle_\rho \equiv \mathrm{Tr}\left[\hat{n}\rho\right] < N$.

Another relevant property may be the purity of the state ρ_η *after* passing through the attenuator. Most states will become mixed after undergoing loss, however coherent states (as used in [LCW+15]) will not. They are instead mapped to coherent states with lower photon numbers. As a result of this, the loss does not introduce any *classical* uncertainty into the estimation of the phase. It may also be the case that entanglement assists the estimation, as is the case with entanglement-assisted illumination [Llo08]. It is important that we search for the most powerful possible attack that Eve may make, taking all of these factors into consideration. It is only then that we may have confidence in our security proofs against the THA or other SCAs.

We will consider the case where Eve may use any multi-mode entangled Gaussian state. Since only one mode enters Alice's apparatus, Eve needs only to use at most one idler mode, which she retains as a reference [NC02]. This state is created by applying a two-mode squeezer to the vacuum followed by a displacement on the mode that enters Alice's system (applying a displacement to the idler mode turns out to have no effect on the amount of information that Eve may learn about the key). Up to a change of variables in the squeezing and displacement parameters, this setup is equivalent to all other combinations of Gaussian operations [Bra05], such as applying single-mode squeezers and displacing before squeezing. This, therefore, represents the most general Gaussian-state attack that Eve may make.

Eve's initial state is then

$$\rho = \hat{D}(\alpha)\ \hat{S}_2(\zeta_E)\ |0\rangle\langle 0|\ \hat{S}_2^\dagger(\zeta_E)\ \hat{D}^\dagger(\alpha), \tag{4.2}$$

where $\hat{D}(\alpha)$ is the displacement operator (defined in Eq. 2.24), $\hat{S}_2(\zeta_E)$ is the two-mode squeezing operator (defined in Eq. 2.32) that operates on Eve's probe and idler modes, and $|0\rangle$ is the vacuum state. Without loss of generality we will let ζ_E be real.

As is typical, we will model the loss due to the attenuator as a beam splitter. A fraction η is allowed to pass through to reach Alice's apparatus, and $1-\eta$ is diverted into an auxiliary environment mode.

The final ingredient to be included is the thermal noise. This may be produced by heating up a portion of the optical fibre, so that Eve receives both her own photons that she sent in, as well as the thermal noise photons added by Alice. Here we need some careful thought as to how exactly we will *mathematically* combine these two states. In other works [LFU17], thermal noise has been added to a signal by passing both the signal and the noise through a beam splitter. However, this does not seem to us to be appropriate in this situation for the following reason. Suppose the combined state is produced by passing these two states through a beam splitter with transmissivity η_{Th}, so that $\eta_{\text{Th}} = 1$ means that the resulting state is entirely a thermal state, and $\eta_{\text{Th}} = 0$ means it is all signal. This introduces a new variable into the situation, which implies some degree of coupling between the thermal source and Eve's returned state. We want Eve to be oblivious as to the actual source of the thermal noise, and simply consider it as a simultaneously arriving light source. In particular, if we let $\eta_{\mathrm{Th}} = 1$ and $\mu_T = 0$, we arrive at the rather paradoxical conclusion that the signal has been completely overwhelmed by a thermal state containing no photons. For a similar reason we cannot combine ρ_η^θ with a thermal density matrix ρ_{Th} by way of a classical mixture such as $p\,\rho_\eta^\theta + (1-p)\,\rho_{\mathrm{Th}}$. As such, we expect that the strength of the thermal noise should depend *only* on the single parameter μ_T. In case the reader is not yet fully convinced that we might yet recreate these same dynamics with an appropriately chosen beam-splitter, we show in Appendix A that this cannot be the case.

A method for the proper treatment of constructing a combined state from multiple simultaneously arriving photonic states was described by Glauber in his original treatment of the coherent states [Gla63]. However, that method involved expressing the states in a diagonal coherent basis (the so-called P-representation). Whilst this is a powerful method, it results in an expression for the state that is no longer easily analytically tractable (although it *is* possible to use this to *numerically* analyse the effects of adding non-thermal noise). Since we are dealing here with Gaussian states we shall take advantage of a nice property of thermal states: that they may be produced be taking the partial trace over one mode of a two-mode squeezed vacuum with squeezing parameter $\zeta_T = \operatorname{arcsinh}(\sqrt{\mu_T})$. Therefore, we shall model the addition of the thermal noise as the action of Alice passing Eve's returning signal through a two-mode squeezer with the vacuum, and discarding one of the resulting modes. Note that she does not *physically* do this, it is only used to find the mathematical form of the state. Within this framework, Eve should choose α and ζ_E to maximise her mutual information with the secret key. The full set-up for the construction of Eve's state is illustrated in Fig. 4.1.

A great advantage of working with Gaussian states is that they may be completely characterised by their first and second moments. For an n-mode Gaussian state let $\underline{\hat{u}}$ be the vector of operators $[\hat{x}_1, \hat{p}_1, \ldots, \hat{x}_n, \hat{p}_n]^T$.[3] Then to each Gaussian state, ρ, we may uniquely assign a pair $(\underline{u}, V)$ which we call the *mean vector* and *covariance matrix* respectively, with elements defined by

$$\begin{aligned} u_i &= \operatorname{Tr}\left[\rho\hat{u}_i\right], \\ V_{i,j} &= \frac{\operatorname{Tr}\left[\rho\hat{u}_i\hat{u}_j\right] + \operatorname{Tr}\left[\rho\hat{u}_j\hat{u}_i\right]}{2} - \operatorname{Tr}\left[\rho\hat{u}_i\right]\operatorname{Tr}\left[\rho\hat{u}_j\right]. \end{aligned} \tag{4.3}$$

Let ϕ be the relative angle between the displacement and the squeezing parameters in the complex plane, $\mu_D = \eta\,|\alpha|^2$ be the average photon number due to displacement after loss, and $\omega = \cosh(2\zeta_E)$ be the normalised quadrature variance for a squeezed vacuum state. It then follows from Eq. 4.3 and Eq. 4.2 that the mean vector and covariance matrix for Eve's returned states corresponding to $\theta = 0$ and $\theta = \frac{\pi}{2}$ are as follows:

[3] Note, in many texts a different convention is used, where all position operators are listed first before all of the momentum operators.

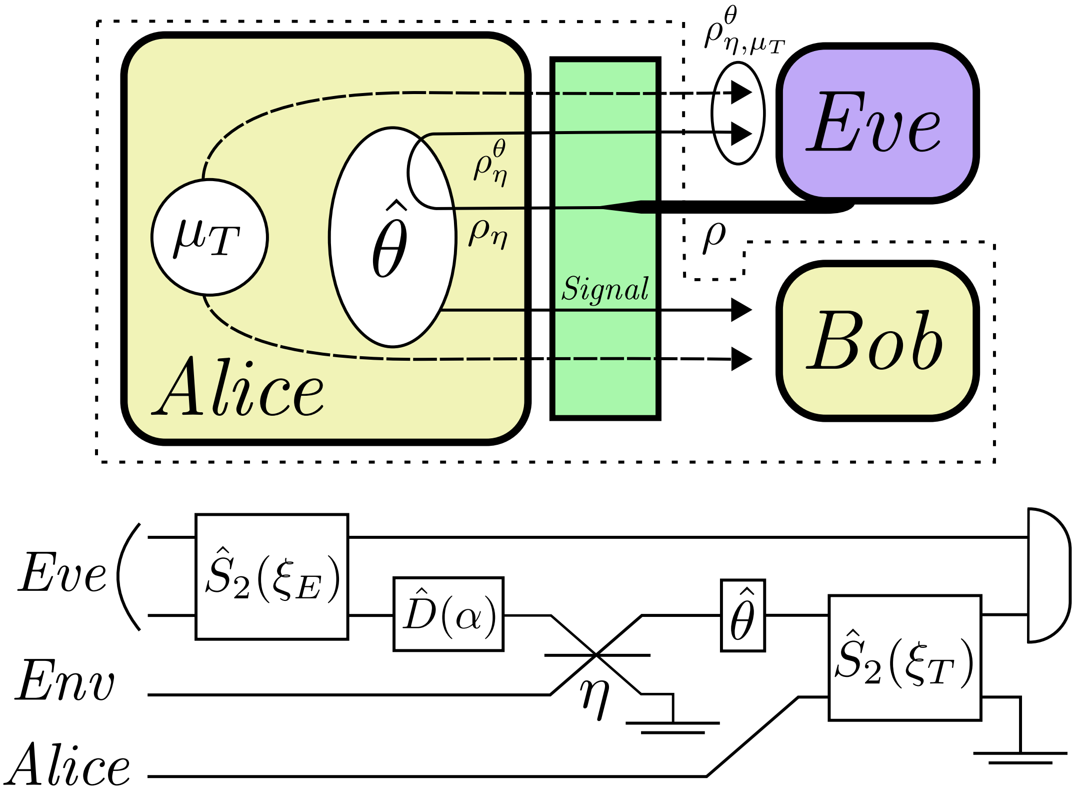


FIGURE 4.1: Top: Schematic diagram illustrating the *physical* mechanisms that produce Bob and Eve's states. The dotted line outlines the "legitimate" part of the protocol, comprising the signal and the thermal noise (dashed line), with the green box representing an attenuator. Bottom: Circuit diagram showing the *mathematical* mechanisms that produce Eve's final state. Shown from left to right are the effects of Eve's squeezing, Eve's state displacement, Alice's attenuator, picking up the phase information, and adding the thermal noise. Double horizontal lines represent taking a partial trace over the relevant mode.

$$
\begin{aligned}
u_{\theta=0} &= \begin{bmatrix} (\sin\phi + \cos\phi)\,\sqrt{2\mu_D} \\ (\sin\phi - \cos\phi)\,\sqrt{2\mu_D} \\ 0 \\ 0 \end{bmatrix}, \\
V_{\theta=0} &= \frac{1}{2}\begin{bmatrix} \left[(1+\mu_T)\,\omega\eta + \mu_T\right]\mathbb{1}_2 & A\,Z \\ A\,Z & \omega\mathbb{1} \end{bmatrix}, \\
u_{\theta=\frac{\pi}{2}} &= \begin{bmatrix} (\cos\phi - \sin\phi)\,\sqrt{2\mu_D} \\ (\cos\phi + \sin\phi)\,\sqrt{2\mu_D} \\ 0 \\ 0 \end{bmatrix}, \\
V_{\theta=\frac{\pi}{2}} &= \frac{1}{2}\begin{bmatrix} \left[(1+\mu_T)\,\omega\eta + \mu_T\right]\mathbb{1}_2 & A\,X \\ A\,X & \omega\mathbb{1} \end{bmatrix},
\end{aligned}
\tag{4.4}
$$

where $A = \sqrt{(1+\mu_T)(\omega^2-1)\,\eta}$ and Z and X are the respective Pauli matrices.

4.1.2 Secret key rate

Recall from Eq. 3.6 that in vanilla BB84 with no threat of THA, the secret key rate is given by $K = R\left[1 - h_2(\epsilon_{\text{bit}}) - h_2(\epsilon_{\text{phase}})\right]$, where one of these terms of $h_2(\epsilon)$ is due to Alice and Bob sacrificing key bits to perform error correction, and one factor is due to them applying classical privacy amplification algorithms.

Due to the nature of the THA being a SCA, Eve's attack will not affect the bit-error rate measured by Alice and Bob. However, it will still clearly compromise the security, so Eq. 3.6 must be modified. In particular we expect that the $h_2(\epsilon_{\text{bit}})$ term representing the error correction should remain unchanged, since a properly implemented SCA will not induce additional errors. However, Alice and Bob *will* have to do additional privacy-amplification, so the $h_2(\epsilon_{\text{phase}})$ term will be modified.

The key rate for BB84 in the presence of an SCA was found by [GLLP04, Koa09, TKI03]. They show that the effect of the SCA may be summarised by a quantity known as the *distinguishability*, Δ. This is used to modify the error rate, ϵ, in the privacy-amplification term to become an *effective error rate*, $\tilde{\epsilon}$ given by the following (which is proved in [LP07], Appendix A.):

$$\begin{aligned} \tilde{\epsilon}(\epsilon,\Delta) = {} & \epsilon + 4\Delta(1-\Delta)(1-2\epsilon) \\ & + 4(1-2\Delta)\sqrt{\Delta(1-\Delta)\epsilon(1-\epsilon)}. \end{aligned} \tag{4.5}$$

This means that we do not have to know exactly what Eve does with the states and the information available to her. For example, she may perform a THA to try to learn θ directly. Or, she might tailor her THA such that the measurement on the returned state only reveals information about the basis that Alice has chosen. After estimating this basis, she might then measure the *flying* qubit in that basis to learn θ without disturbing the state. She might do some combination of these approaches, or something else entirely. As such, it is of foundational importance to our analysis that we have some way of quantifying the strength of a THA that only makes reference to the state she sends *out*, not to *what she does* to the state she gets back, including any measurement or series of measurements on any combination of the returned state and flying qubits.

The distinguishability varies from 0 when all choices of θ are indistinguishable from the point-of-view of Eve, to $\frac{1}{2}$ when she can distinguish all settings with certainty. In practice, a value of Δ much greater than 0 will result in a secret key rate of 0, since it would require Alice and Bob to be sacrificing raw key bits for error correction and privacy amplification at a rate faster than they are being generated. This formulation of the strength of a THA in terms of Δ puts a lower bound on the secret key rate that Alice and Bob can hope to achieve. The distinguishability is given by[4]

$$\Delta \leq \frac{1 - \sqrt{F\left(\rho^{0}_{\eta,\mu_T}, \rho^{\pi/2}_{\eta,\mu_T}\right)}}{2}, \tag{4.6}$$

where F is the quantum fidelity function, given in Eq. 2.34. Note that this is different from the form given in [LCW+15]. There, they reduce Eq. 4.6 to a form involving the optimal purifications of the two output states. Since they are using pure coherent states, such optimal purifications are easily found. However, there exists no prescriptive formula to find these for a pair of general mixed states, so we must use the fidelity form of the distinguishability.

The rest of this section is dedicated to calculating an exact expression for Δ for the set of thermalised Gaussian states described above, and section 4.2 is focused on calculating a bound on Δ for the set of general separable states. It should be noted that, unlike ϵ, $\tilde{\epsilon}$ (or equivalently Δ) cannot be directly measured in the process of running the QKD protocol. Therefore Alice should be able to perform some local action to be able to determine Δ to some high precision, and then use this value to determine how much privacy amplification they should perform.

The problem of calculating the fidelity between two multimode Gaussian states was solved by [BBP15][5]. There, they show that, for any Gaussian states ρ_1, ρ_2, we have:

$$\begin{aligned}\sqrt{F\left(\rho_1, \rho_2\right)} &= \mathcal{F}\left(\tilde{V}_1, \tilde{V}_2\right) e^{-\frac{1}{4}\left(\underline{u}_1 - \underline{u}_2\right)^T \left(\tilde{V}_1 + \tilde{V}_2\right)^{-1} \left(\underline{u}_1 - \underline{u}_2\right)} \\ \mathcal{F}\left(\tilde{V}_1, \tilde{V}_2\right) &= \frac{\prod_{k=1}^{n} \sqrt{w_k + \sqrt{w_k^2 - 1}}}{\sqrt[4]{\det\left(\tilde{V}_1 + \tilde{V}_2\right)}},\end{aligned} \tag{4.7}$$

where $\tilde{V}$ is equivalent to V, but expressed in the basis $\left[\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n, \hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n\right]^T$ and w_k are the eigenvalues of the auxiliary matrix W, defined as

$$\begin{aligned} W &= -2i\Omega^T \left(\tilde{V}_1 + \tilde{V}_2\right)^{-1} \left(\frac{\Omega}{4} + \tilde{V}_2 \Omega \tilde{V}_1\right) \Omega, \\ \Omega &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \mathbb{1}_n. \end{aligned} \tag{4.8}$$

[4] This may be seen by considering Ref. [GLLP04], section VIII. The purified states corresponding to each basis are as defined in Ref. [LCW+15], Appendix B. From these it may be seen that $1 - 2\Delta$ is equal to the *average* square-rooted fidelity between a state being emitted in the X basis and one in the Y basis. By the symmetry and unitary invariance of the fidelity function, this reduces to finding the root fidelity between only the states corresponding to $\theta = 0$ and $\theta = \pi/2$.

[5] Note, their definition of F is the square root of the definition used in Eq. 2.34. We modify the result of the cited work to have a definition consistent with our usage.

When we combine the fidelity given in Eq. 4.7 with the mean vectors and covariance matrices given in Eq. 4.4, we find that the fidelity between two of Eve's returned states is given by

$$\begin{aligned} &\sqrt{F\left(\rho^{0}_{\eta,\mu_T},\rho^{\pi/2}_{\eta,\mu_T}\right)} = \\ &\frac{1}{4B}e^{-2\mu_D\omega/B}\left(\sqrt{C}+\left|4\mu_T\omega+4\eta(1+\mu_T)-1\right|\right), \end{aligned} \tag{4.9}$$

where

$$\begin{aligned} B &= 2\mu_T\omega+(1+\mu_T)(\omega^2+1)\eta, \\ C &= 16\eta^2(1+\mu_T)^2+8\eta(1+\mu_T)\omega(4\mu_T+\omega) \\ &\quad +(1+4\mu_T\omega)^2. \end{aligned} \tag{4.10}$$

Eve wants to choose her parameters ζ_E and μ_D in order to minimise the fidelity (and so maximise the distinguishability) between her returned states. Whilst increasing either of these parameters decreases F, she is not necessarily free to do both simultaneously. Both squeezing and displacement increase the average number of photons in each mode, which is limited by some number N. Overall, a state that is displaced by α and squeezed by ζ_E has average photon number

$$\langle n\rangle = |\alpha|^2+\sin^2(|\zeta_E|). \tag{4.11}$$

We may see that the average number of photons in such a state has two parts: a displacement-contingent part and a squeezing-contingent part. Since these parts are independent, we may imagine that the situation is as follows: Eve takes her N photons, and distributes them into displacement and squeezing. She uses pN of her available photons to contribute towards squeezing and $(1-p)N$ towards displacement. Since a squeezing parameter of ζ_E gives an average photon number per mode of $\sinh^2(\zeta_E)$, and a displacement parameter of α contributes $|\alpha|^2$ photons, we find that we can do no better that setting the parameters such that $\omega = \cosh\left[\operatorname{arcsinh}(2\sqrt{pN})\right]$, $\mu_D = (1-p)N\eta$ for some p. When we insert these values into Eq. 4.9, we can investigate the behaviour as a function of p and η for various values of N and μ_T. We find that F is always minimised when $p = 0$. This means that Eve is best served by using *all* of her photons to contribute to the displacement of her state. As such, we can now simplify the fidelity, which may be written as

$$\sqrt{F\left(\rho^{0}_{\eta,\mu_T},\rho^{\pi/2}_{\eta,\mu_T}\right)} = \exp\left(-\frac{\mu_D}{1+2\mu_T}\right). \tag{4.12}$$

This means that Eve's optimal Gaussian-state attack is one involving coherent states only. This provides a rigorous footing for earlier works which analyse the results of coherent-state attacks with an attenuating defence [LCW+15]. Note also, that this is true whether or not Alice has fine control over the thermal noise level.

4.1.3 Effect of thermal noise on Bob

We can see from Eq. 4.12 that Eve's knowledge of θ is minimised when μ_T becomes very large. However, when Alice sends a lot of thermal photons into the system, some of these photons are also picked up by Bob. If Bob measures these instead of the signal photons then he is likely to pick up a bit error. Clearly when $\mu_T \to \infty$ Alice and Bob will not be able to securely communicate, so we need to find an optimal level of μ_T that clouds Eve's estimation of the state without affecting Bob too much. In this section we analyse and quantify this.

Consider the case where Alice and Bob are sending and receiving in the same basis. By implementing asymmetric BB84 [LCA05], this can be the case for almost all qubits. We now say that at the end of the optical channel between Alice and Bob there is a polarising beam splitter which sends the incoming photons into one of two detectors which we label the *correct* and *wrong* detectors. These represent Bob measuring the bits that Alice did and did not send respectively. That is to say, if Alice sends a 0 bit and Bob measures a 0 bit, or if Alice sent a 1 and Bob measured a 1, then we say that the detector labelled *correct* clicked. Otherwise, we say that the *wrong* detector clicked. This is simply a shorthand way of encapsulating the symmetry of the system in a way that does not overly-complicate things.

Alice and Bob will try to distill a secret key from the key bits where Bob believes he detected only a single photon (which he must assume to be the signal photon). The probability for this to occur for a given signal qubit we will call p_{succ}. Here we will assume that Bob uses bucket detectors. That is to say they have two measurement outcomes: either no photons were detected, or one or more photons were detected. We show in the box below that, perhaps counter-intuitively, this actually gives a *better* secret key rate than using number-resolving detectors, in agreement with previous results [LFU17].

Let $p_{\checkmark}, p_{\times}$ be the probabilities that the signal photon was detected in the correct and wrong detector respectively, and $p_{\bullet}$ be the probability that the signal photon is not detected at all. Let $q(c,w)$ be the probability that c noise photons are detected in the correct detector and w in the wrong detector. Bob will register a "valid" qubit if exactly one of the detectors clicks. We can then say that

$$\begin{aligned} p_{\text{succ}} = p_{\checkmark} \sum_{c=0}^{\infty} q(c,0) + p_{\times} \sum_{w=0}^{\infty} q(0,w) \\ + p_{\bullet} \sum_{c=1}^{\infty} q(c,0) + p_{\bullet} \sum_{w=1}^{\infty} q(0,w). \end{aligned} \tag{4.13}$$

We can identify the error rate, ϵ, as the probability that a photon gives a click in the wrong detector, and is therefore equal to

$$\epsilon = \frac{p_\times \sum_{w=0}^{\infty} q(0,w) + p_\bullet \sum_{w=1}^{\infty} q(0,w)}{p_{\text{succ}}}. \tag{4.14}$$

To calculate these quantities, we identify the following:

$$\begin{aligned} p_\checkmark &= T(1-Q), \\ p_\times &= TQ \\ p_\bullet &= (1-T) \\ q(i,j) &= \frac{\tilde{\mu}_T^i}{(\tilde{\mu}_T+1)^{i+1}} \frac{\tilde{\mu}_T^j}{(\tilde{\mu}_T+1)^{j+1}} \\ \sum_{c=0}^{\infty} q(c,0) = \sum_{w=0}^{\infty} q(0,w) &= \frac{1}{\tilde{\mu}_T+1} \\ \sum_{c=1}^{\infty} q(c,0) = \sum_{w=1}^{\infty} q(0,w) &= \frac{\mu_T}{(\tilde{\mu}_T+1)^2}. \end{aligned} \tag{4.15}$$

Here, Q is the probability for Bob to register a bit-flip error in the absence of thermal noise, T is the transmissivity of the channel including the efficiency of Bob's detectors,[6] and $\tilde{\mu}_T = T\mu_T/2$ is the average number of thermal photons arriving at each detector. In practice, $\tilde{\mu}_T$ may be lower than this, since Eve will have inadvertently intercepted some of them. However, we give here the worst-case scenario. We can then say that

$$\begin{aligned} p_{\text{succ}} &= \frac{2T(2+2\mu_T - T\mu_T)}{(2+T\mu_T)^2}, \\ \epsilon &= \frac{2Q + \mu_T(1-T+QT)}{2+\mu_T(2-T)}. \end{aligned} \tag{4.16}$$

In order to find an expression for T we assume that the photons face an exponential drop-off with distance, and set $T = e^{-L/L_{\text{att}}}$ where L_{att} is the attenuation length. To model Q, we may assume that, as is usual for QKD protocols, Alice and Bob are equipped with quantum memories, and the flying qubits are used as a process by which they establish entanglement between these memories [SDRA+07, NTD+16, ZDB12, VK17]. This is necessary for all but the most primitive protocols, since some storing of entangled qubits is required in order to perform entanglement distillation and privacy amplification algorithms such as DJEMPS [DEJ+96], which are needed in order to prove that Eve has not entangled Bob's state with an ancilla. As such, we set $Q = \frac{1}{2}\left[1 - e^{L/c\tau}\right]$, where τ is the lifetime of the memory (typically on the order of microseconds) and c is the speed of light. Finally, when $p_{\text{succ}} < 1$, we must replace Δ with Δ/p_{succ}, since the lost signals may have been selectively eliminated by Eve to improve her mutual information with the key (Ref. [GLLP04], Eq. 32).

[6] In Chapter 5 we will use η for the transmissivity, as is commonly the case in communication literature. However, here we use T to avoid confusion with the transmissivity of Eve's channel.

Use of photon-number resolving detectors

Our calculation of the secret-key rate is based on the fact that Bob uses bucket detectors. One might naturally ask whether using more state-of-the-art technology such as photon-number-resolving detectors (PNRDs) would improve the situation for Bob.
When this is the case, Eq. 4.13 becomes

$$\begin{aligned} p_{\text{succ}} &= p_{\checkmark}\, q(0,0) + p_{\times}\, q(0,0) \\ &\quad + p_{\bullet}\, q(1,0) + p_{\bullet}\, q(0,1), \\ &= T\,\frac{1}{(\tilde{\mu}_T+1)^2} + 2\,(1-T)\,\frac{1}{\tilde{\mu}_T+1}\frac{\tilde{\mu}_T}{(\tilde{\mu}_T+1)^2}, \end{aligned} \tag{4.17}$$

since a click is only registered when *exactly* one photon enters a detector.
When we calculate ϵ by taking the sum of the contributions to p_{succ} that cause the *wrong* detector to click conditioned on the probability of getting a click in the first place, we find that ϵ is actually the same whether we use PNRDs or bucket detectors! This interesting congruence is a result of the fact that the noise obeys thermal statistics, and will not generally be true for other noise distributions. However, whilst the relative error is not affected, p_{succ} is actually *lower* in the case of PNRDs! This means that the secret key rate can not be improved by using PNRDs, and is in fact worsened in almost all cases. This is because the restriction to PNRDs means that Bob is more likely to reject legitimate signals than noise photons.

4.1.4 Results of Gaussian-state analysis

By combining these elements, which measure the effects of the thermal noise on Eve and on Bob, as well as the result that coherent states are optimal, we find that the final secret key rate for a general multi-mode Gaussian state attack in the presence of an attenuating filter and a thermal noise defence with error rate ϵ may given by

$$K = p_{\text{succ}}\left[1 - h_2\left(\epsilon\right) - h_2\left(\tilde{\epsilon}\left(\epsilon, \Delta'\right)\right)\right], \tag{4.18}$$

where $\Delta' = \left[1 - \exp\left(-\frac{\mu_D}{1+2\mu_T}\right)\right] / \left(2\, p_{\text{succ}}\right)$ and $\tilde{\epsilon}(\epsilon, \Delta')$ is defined as above.

Since Eq. 4.18 is highly dependent on μ_T, we optimise K over μ_T to find a true measure of the utility of the thermal noise defence. We consider the case where

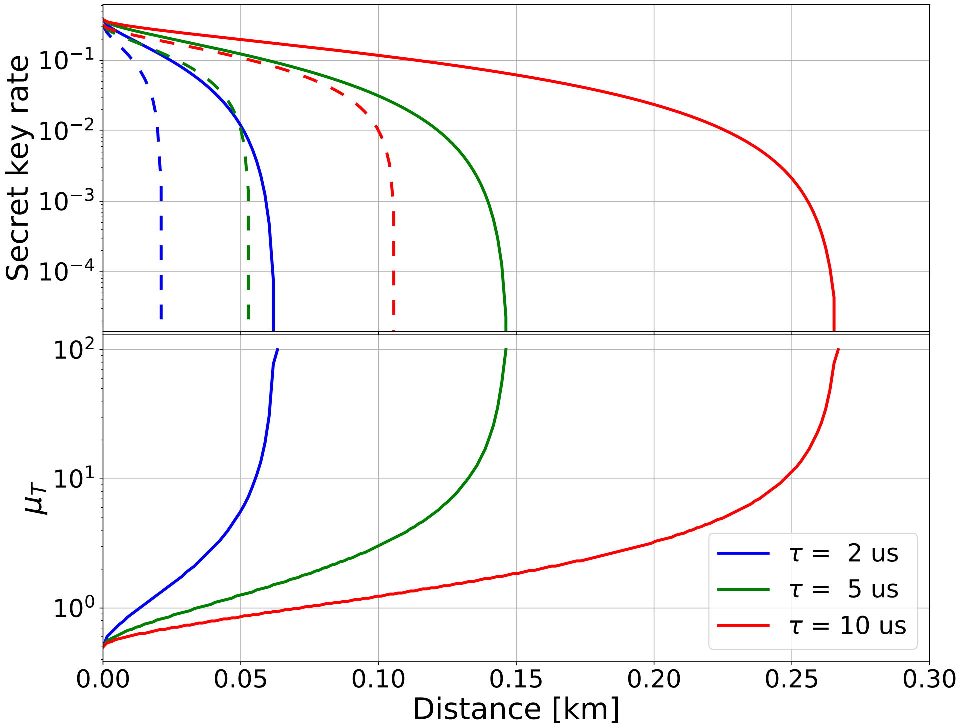


FIGURE 4.2: Minimum achievable secret key rate under the presence of a coherent state attack (the most powerful Gaussian-state attack), with a thermal noise defence involving μ_T photons being added by Alice to cloud Eve's attack signal. Top graph shows the key rate for different values of quantum memory time, τ. Dashed lines show the key rate with no thermal noise added, and solid lines are key rates maximised over μ_T. From left to right for both dashed and solid lines, we have $\tau = 2, 5, 10\,\mu s$ respectively. Lower graph shows the optimal values of μ_T at each distance. In all plots $\mu_D = 0.1$ and $L_{\text{att}} = 25\,\text{km}$.

$\mu_D = 0.1$ and $L_{\text{att}} = 25$ km. The ability of the thermal noise defence to protect the secret key rate is shown in Fig. 4.2.

We can see from this that employing a thermal noise defence has the capacity to more than *double* the effective range of a QKD system. However, when the key rates get very small, very high temperatures are required in order to retain any security. This is because the mechanism by which increasing μ_T decreases the key rate is by decreasing the pre-factor p_{succ}, which may go arbitrarily low, but never reach 0. On the other hand, there is a small *critical* effective error-rate not equal to 1 that Eve can induce that will result in the key rate dropping to zero. Therefore, when Alice calculates that the effective error that could result from Eve launching a THA is above the critical value, Alice is forced to increase the thermal noise to very high values to bring the effective error down again. This is another expression of the general principle that raw key rates may be sacrificed to boost security.

One may notice that the ranges shown in Fig. 4.2 are well below those shown in many other QKD proposals. One reason for this is that other proposals in general do not consider the effects of THAs, or SCAs in general, and the key-rate boon provided by the thermal noise defence cannot exceed the calculated key rate of a system that does not recognise SCAs in the first place. Secondarily, the key-rate plots shown here are intended to show the *relative advantage* of adding thermal noise to the system. They are not intended to show the state-of-the-art ranges that might be achieved by more advanced QKD protocols, which may take advantage of technologies such as quantum repeaters [VK17, DLCZ01, PR15, EKB16], long-lived quantum memories [PRM17, NTD+16] or post-selective entanglement generation [KMN+07, VK17].

4.2 General separable attacks

We have shown that, amongst multi-mode collective Gaussian attack states that Eve might use, the separable coherent state is optimal. Whilst it may seem initially surprising that entanglement does not assist her, note that entanglement between any two modes will drop off as more of the signal is attenuated. We are left with a distinguishability that depends only on the average output photon number, μ_D, so Δ does not *explicitly* depend on the transmissivity η.

It may be argued that coherent states are likely to be optimal amongst the separable states, since under loss, photon-counting statistics will tend to be Poissonian [HPLG07]. Therefore the best state one can hope to use is a Poissonian state that retains coherence, i.e. a coherent state. However, a state that is initially highly non-Poissonian in its statistics may require a very high attenuation before it approximates a Poissonian distribution, and there is no guarantee that the expression for Δ derived from coherent states will still hold.

In this section, we consider the set-up where Alice defends against a THA by use of an attenuator but uses no thermal defence, and that Eve attacks the system using *any* separable state, but gets back a state with only a few photons. Whilst this seems to be a special case for Eve, note that it is more general that the situation considered in Section 4.1 since this approach considers a set of states which includes, yet is larger than, the set of coherent states that are optimal within the Gaussian states.

Here, we will consider Eve's input state in its density matrix form instead of its covariance matrix form. We will consider the effect of the attenuator on ρ as a quantum channel, which we express in Kraus operator form:

$$\begin{aligned} \mathcal{E}\left(\rho\right) &= \sum_{k=0}^{\infty} \mathcal{E}_k\left(\rho\right) \equiv \sum_{k=0}^{\infty} \hat{A}_k \rho \hat{A}_k^{\dagger} \\ \hat{A}_k &= \sum_{j=k}^{\infty} \sqrt{\binom{j}{k}} \sqrt{\eta}^{\,j-k} \sqrt{1-\eta}^{\,k} \left|j-k\right\rangle\left\langle j\right| . \end{aligned} \tag{4.19}$$

Each term $\hat{A}_k \rho \hat{A}_k^{\dagger}$ represents k photons being lost from the state ρ, each with independent probability $1-\eta$.

We express each term in the map as follows

$$\begin{aligned}\mathcal{E}_k(\rho) &= \sum_{i,j=k}^{\infty} B_{i,j,k}(\eta)\,\langle i|\,\rho\,|j\rangle \cdot |i-k\rangle\langle j-k| \\ B_{i,j,k}(\eta) &= \sqrt{\binom{i}{k}\binom{j}{k}}\,\eta^{\frac{i+j}{2}-k}(1-\eta)^k.\end{aligned} \tag{4.20}$$

Since we require a very high level of attenuation to achieve any kind of useful secrecy, we may assume that η is very close to 0. Therefore we may expand the factor $B_{i,j,k}(\eta)$ as

$$B_{i,j,k}(\eta) \approx B_{i,j,k}(0) + \frac{\mathrm{d}B}{\mathrm{d}\eta}\Big|_0 \eta + \frac{1}{2}\frac{\mathrm{d}^2B}{\mathrm{d}\eta^2}\Big|_0 \eta^2, \tag{4.21}$$

leading to an expansion of each term in the map as

$$\mathcal{E}_k \approx \mathcal{E}_k^{(0)} + \mathcal{E}_k^{(1)} + \mathcal{E}_k^{(2)}. \tag{4.22}$$

Using the fact that $\lim_{x\to 0} x^p = \delta_{p,0}$ for $p > 0$, we can see that $B_{i,j,k}(0) = \sqrt{\binom{i}{k}\binom{j}{k}}\delta_{(i+j)/2,k}$. Performing the sums over i, j, k we get $\mathcal{E}^{(0)} = |0\rangle\langle 0| \sum_{k=0}^{\infty} \langle k|\,\rho\,|k\rangle = |0\rangle\langle 0|$.[7]

In the same way, we find that

$$\begin{aligned}\mathcal{E}_k^{(1)} &= \sum_{i,j=k}^{\infty} \sqrt{\binom{i}{k}\binom{j}{k}}\left[\left(\frac{i+j}{2}-k\right)\delta_{\frac{i+j}{2},k+1} - k\delta_{\frac{i+j}{2},k}\right]\eta \\ \to\ \ \mathcal{E}^{(1)} &= -\mu\,|0\rangle\langle 0| + \\ &\quad \mu\,|1\rangle\langle 1| + \\ &\quad \sum_{k=0}^{\infty}\sqrt{\binom{(k+2)}{k}}\,|0\rangle\langle 2|\,\langle k|\,\rho\,|k+2\rangle + \\ &\quad \sum_{k=0}^{\infty}\sqrt{\binom{(k+2)}{k}}\,|2\rangle\langle 0|\,\langle k+2|\,\rho\,|k\rangle\,,\end{aligned} \tag{4.23}$$

where $\mu = \eta \sum_{k=0}^{\infty} k\,\langle k|\,\rho\,|k\rangle$ is the average number of photons that Eve receives back after attenuation.[8]

Similarly,

[7] Where we have used the fact that $\mathrm{Tr}[\rho] = 1$.

[8] Note that $\delta_{(i+j)/2,k+1} = \delta_{i,k}\delta_{j,k+2} + \delta_{i,k+2}\delta_{j,k} + \delta_{i,k+1}\delta_{j,k+1}$

$$\begin{aligned}\mathcal{E}^{(2)} =& \frac{\eta^2}{2}\left(v + \langle\hat{n}\rangle_\rho^2 + \langle\hat{n}\rangle_\rho\right)|0\rangle\langle 0| - \\ & \eta^2\left(v + \langle\hat{n}\rangle_\rho^2 + \langle\hat{n}\rangle_\rho\right)|1\rangle\langle 1| + \\ & \frac{\eta^2}{2}\left(v + \langle\hat{n}\rangle_\rho^2 + \langle\hat{n}\rangle_\rho\right)|2\rangle\langle 2| + \\ & \text{off-diagonals on } \; |0\rangle\langle 2| \; \text{ and } |2\rangle\langle 0| + \\ & \text{terms on } \; |i\rangle\langle j| \; \text{ where } i \text{ or } j \geq 3.\end{aligned} \tag{4.24}$$

where $v = \langle\hat{n}^2\rangle_\rho - \langle\hat{n}\rangle_\rho$ is the variance in the initial state. Whilst the diagonal terms can be expressed in terms of the macroscopic observables of ρ, the off-diagonal terms have no such simple expression.

Firstly we should bound the effects of higher-order terms. We do this by supposing that Eve performs a measurement on her returned state to determine whether or not the state contains 2 or fewer photons. That is to say, her initial measurement of $\mathcal{E}(\rho)$ has 2 outcomes corresponding to operators $\hat{E}_\checkmark = |0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|$ and $\hat{E}_\times = \sum_{k=3}^\infty |k\rangle\langle k|$.

In order to ensure that this measurement does not reduce the information that Eve learns about the state, we say that if she gets the result corresponding to $\hat{E}_\times$ then we assume that she learns the key bit θ perfectly. That is to say, instead of receiving $\mathcal{E}\left(\rho\right)$ she can be said to receive some state $|\theta\rangle\langle\theta|$, where $\langle\theta_1|\theta_2\rangle = \delta_{\theta_1,\theta_2}$.

If the measurement is successful, then Eve's state is projected onto the two-or-fewer-photon subspace, giving $\rho_\text{sub} = \hat{E}_\checkmark\mathcal{E}(\rho)\hat{E}_\checkmark / \operatorname{Tr}\left[\hat{E}_\checkmark\mathcal{E}(\rho)\right]$. In the case where $\theta = 0$, this may be expressed in the basis of $\{|0\rangle, |1\rangle, |2\rangle\}$ by

$$\rho_\text{sub}^{\theta=0} = \begin{bmatrix} 1-\mu+\frac{\eta^2}{2}\alpha & 0 & \beta \\ 0 & \mu-\eta^2\alpha & 0 \\ \beta & 0 & \frac{\eta^2}{2}\alpha \end{bmatrix}, \tag{4.25}$$

where $\alpha = v + \langle\hat{n}\rangle_\rho^2 + \langle\hat{n}\rangle_\rho$ and β is some coefficient that cannot be easily expressed in terms of macroscopic properties of the state. In the case where $\theta = \pi/2$ we simply pick up a factor of -1 on the coefficient β. The overall state Eve gets back is then

$$\rho_\text{returned}^\theta = \operatorname{Tr}\left[\hat{E}_\checkmark\mathcal{E}(\rho)\right]\rho_\text{sub}^\theta + \operatorname{Tr}\left[\hat{E}_\times\mathcal{E}(\rho)\right]|\theta\rangle\langle\theta|. \tag{4.26}$$

In order to bound the contribution of the second term, we show in Appendix B that

$$\operatorname{Tr}\left[\hat{E}_\checkmark\mathcal{E}(\rho)\right] \geq e^{-\mu}, \tag{4.27}$$

and so $\operatorname{Tr}\left[\hat{E}_\times\mathcal{E}(\rho)\right] \leq 1 - e^{-\mu}$.

Since we want to find an upper limit to the information that Eve can learn, we say that she can receive any state that is consistent with both Eq. 4.25 and the laws

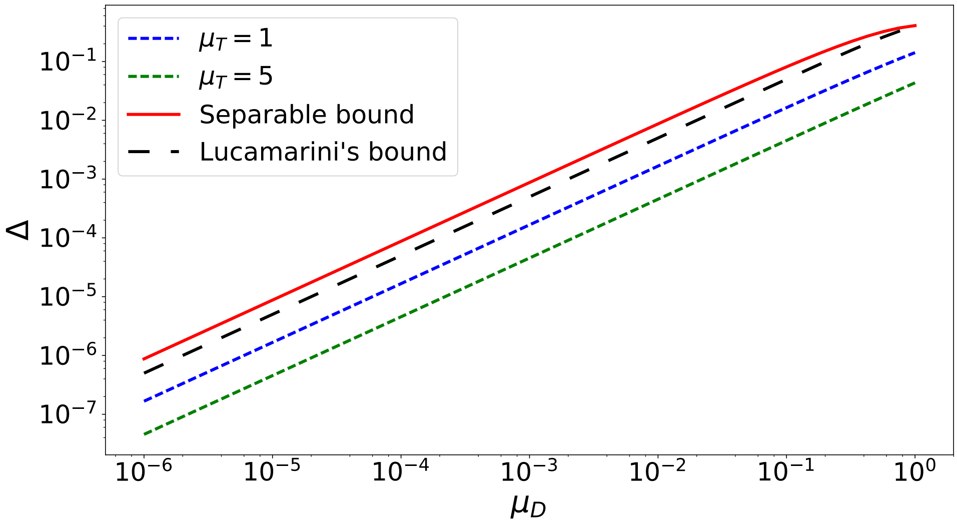


FIGURE 4.3: Upper bounds on the distinguishability, with average returned photon number μ. Dotted lines show the distinguishability for a coherent-state attack with a thermal noise defence, of thermal photon number μ_T. Upper dotted line (blue) shows $\mu_T = 1$, lower dotted line (green) shows $\mu_T = 5$. The black dashed line is the distinguishability for coherent-state attacks found by Lucamarini et. al. Solid line is the bound for separable states, which as expected is always greater that the other bounds.

of physics. We find that the fidelity between two such density matrices is minimised when the variance v is chosen such that the $|1\rangle\langle 1|$ component is 0 and the off-diagonal terms are maximised. Because of this, it turns out that we do not need to be able to express β in a way relating to macroscopic properties such as average photon number and variance. We simply choose β to be the largest value such that $\rho^{\theta}_{\text{sub}}$ remains positive semi-definite, which is $\beta = \frac{1}{2}\sqrt{\mu(2-\mu)}$.

This, along with Eq. 4.26 and Eq. 4.6 gives an ultimate distinguishability bound for the separable attack-state case of

$$\Delta \leq \frac{1 - e^{-\mu}\sqrt{1 - 3\mu(2-\mu)/4}}{2}. \tag{4.28}$$

Importantly, this is a function of a single variable that is measurable by Alice; the average output photon number. By bounding quantities of Eve's state that cannot be measured, we have ensured that Alice can make an accurate assessment of how secure her QKD system is, whilst not knowing anything about the microscopic details of Eve's state.

Fig. 4.3 shows that the value of Δ is higher for our separable bound than for the case of a coherent state, whether one diluted by thermal noise or not. We also show it to be higher than the bound on Δ for a noiseless coherent-state attack found by Lucamarini et al. of $\Delta = [1 - e^{\mu}\cos(\mu)]/2$. Whilst our bound on Δ is not absolutely tight (since the 3-photon contributions surely will not convey perfect information of θ), we can see that is not too generous, since it tracks the known achievable bounds quite closely.

4.3 Shutter defence

In both sections 4.1 and 4.2, it was necessary that Alice use a strong one-way attenuator, also known as an optical isolator. This was able to let almost all of the light through in a forward direction, whilst blocking all but perhaps a single photon going in the reverse.[9] Whilst such devices certainly exist [JPE+13, FLO+00], we should ask if such security is possible without such a high level of attenuation, and if there is any other device we can insert into the optical channel that will result in a large *effective* attenuation factor.

We show here that this is possible by considering the effect of replacing the attenuator with a *shutter*, such as an electro-optic modulator or a chopper. This runs contrary to the claim of [GRTZ02] Sec. VI K, where it is claimed that a shutter cannot defend against a THA (although they consider only a shutter directly adjacent to the apparatus). This lets light through for a short duration of t_S, at a period of t_P. Once Eve's pulse passes the shutter, it must travel the remaining distance to the apparatus that encodes θ, reflect off it, and return to the shutter. If the shutter is closed when the pulse arrives, it will reflect off the rear-side of the shutter with some coefficient of reflectivity η_R, return to the apparatus, and then reflect again. This will continue until the pulse of light arrives back at the shutter whilst it is open, and it will then pass through and be detected by Eve. This is schematically illustrated in Fig. 4.4. Note that whilst the pulse will pick up an additional phase factor of θ on each reflection, we will hold by the principle of assuming that Eve's computational and measurement power is the maximal allowed by the laws of physics. As such, it is plausible that she will be able to know by the time taken for the pulse to return exactly how many times the light reflected off the apparatus, and so calculate an estimate for the actual value of θ from her measured value.

Let t_L be the time period that a pulse of light takes to make the return journey from the shutter to the encoding apparatus and back. If the light makes R return journeys, then it will have reflected off the rear-side of the shutter $R-1$ times. One may see that R can be found to be the smallest integer such that

$$0 \leq (R \times t_L \bmod t_P) \leq t_S. \tag{4.29}$$

Note that if the light travel time is known to arbitrary precision and t_S can be made arbitrarily short, then R can be made arbitrarily high, by letting $t_L = t_P - \delta$ for some arbitrarily small δ. This, however, is physically unrealistic. We should instead model the shutter as being open for some finite fraction of the light travel time. Suppose initially that the shutter is open for a tenth of the period, i.e. $t_S = t_P/10$. We

[9] An example of this is a *Faraday isolator*. The first component of this is a vertical polarising filter (filter A). After passing through this filter, the light passes through an optical rotator, which rotates the polarisation by 45° clockwise, after which it meets another filter (filter B), aligned at an angle that is rotated 45° clockwise with respect to the first, letting all light through. However, when light tries to pass in the reverse direction, it will pass through filter B, rotate 45° clockwise (with respect to its new direction of travel), and subsequently meet filter A at an angle perpendicular to the direction of its filtering, allowing no light to pass.

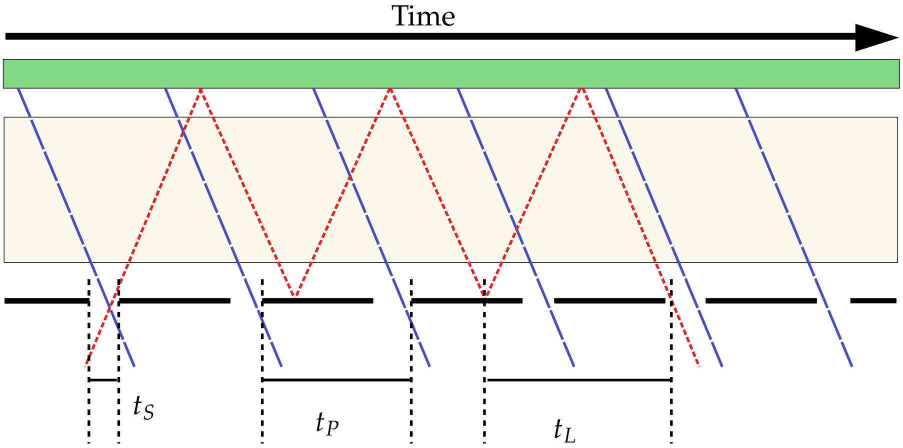


FIGURE 4.4: Schematic showing the operation of the shutter defence. The green box is the apparatus that encodes the phase θ (and so the key bit). The beige box is a one-way attenuating filter. Long blue dashes show the signals, sent out by Alice. These are sent at times such that they always pass out of the filter. The short red dashes show Eve's probe, which must undergo many reflections before being able to leave the filter.

may say that the light travel time may be varied by implementing various lengths of coiled optical fibre between the shutter and the apparatus.

The upper part of Fig. 4.5 shows the values of R that result from varying light travel times, where the light travel time is measured in units of t_P. In the middle part of Fig. 4.5 we convert this value of R into a secret key rate. We do this by using Eq. 4.18 to find K (with $p_{succ} = 1$) and use the separable bound Eq. 4.28 for Δ. We may say that the average output photon number is given by $\mu = N\eta_R^{R-1}$, where N is the input photon number.

The lower part of Fig. 4.5 is the achievable key rate after the application of what we call a *minimising convolution*. This is a functional which takes a function $f(\cdot)$, and maps each point x to the minimum of $\{f(x+x') \mid x' \in [-\delta, \delta]\}$ for some convolution width δ. This is necessary because there may be some experimental uncertainty in the light travel time. So, for example, whilst a value of t_L infinitesimally close to, but less than t_P may seem to give the highest value of R, and so the highest key rate, if t_L was even slightly underestimated, this would result in a value of $R = 1$ and so a far lower security would be achieved. For Fig. 4.5 we have used $\delta = 0.01$. This means that if we have a 95% confidence interval of knowing the light travel time to within 1%, then we can have a confidence of 95% of being able to achieve the key rate shown in the lower graph.

Note that if we fix the light travel time appropriately (to approximately about $0.9 \times t_P$) then we can achieve a secret key rate of almost unity from a co-efficient of reflectivity of $\eta_R = 0.5$. One thing that one should bear in mind is that this is strongly dependent on the width of the confidence interval on t_L. Since we take the minimum of the interval, it is clear that if the interval is too wide, then no secret key

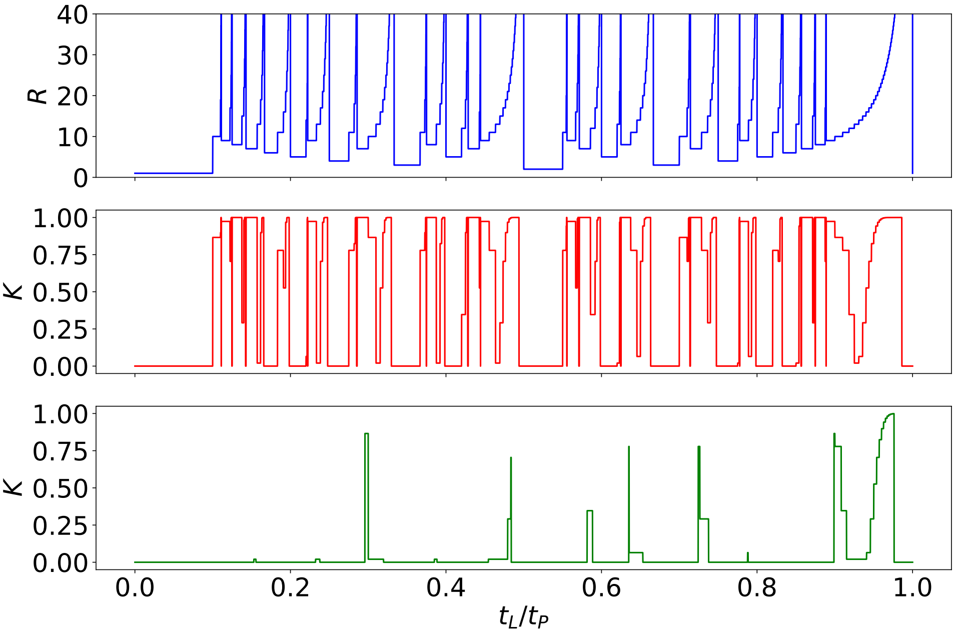


FIGURE 4.5: Top: The number of reflections a pulse of light will make, as a function of the light travel time between the shutter and the apparatus. Middle: The resulting secret key rate. Here, we have used $\eta_R = 0.5$, meaning half of the light is lost upon each reflection. Bottom: The minimum secret key rate that can be guaranteed if there is a 1% margin of error in the knowledge of the light travel time, found using a minimising convolution functional.

rate will be able to be guaranteed.

If one wishes to halve the width of the relative confidence interval, this can be done by doubling both the light travel time and the shutter period. However, this will result in a longer time between raw key bit attempts, which will have a lowering effect on the key rate. Similarly, if one tries to increase the key rate whilst using a shutter defence, one should be mindful of the effect it has on the minimum achievable key rate. For example, if the raw key rate was doubled, the uncertainty in t_L could increase from 1% to 2%. We find that this results in a maximum key rate after the application of a minimising convolution of around 0.75, and so increasing the overall secret key rate to $2 \times 0.75 = 1.5$ of the original value. Similar to the case with adding thermal noise, in any experimental realisation one will have to adjust the light travel time to find the right balance between increasing raw key rate and increasing secrecy. In any case, this brief analysis shows that the use of a shutter as a defence against the THA is one worthy of consideration, and may provide effective attenuations comparable to those of directional attenuators. The choice as to which to use in any given implementation will depend on the details of the experimental set-up.

4.4 Summary

The discovery and implementation of the Trojan Horse Attack once threatened to eliminate the security so famously promised by quantum key distribution. Early seminal works have shown that the situation is not hopeless, and have indicated ways to quantify and abate this threat.

In this work we have fully characterised and quantified the effect of the THA on the key rate under two general attack vectors. We have shown that if Eve uses a multimode Gaussian attack state, her best bet is to use a coherent state. We have also quantified the maximum damage on the secrecy that could be caused by Eve using an arbitrary separable state. We hope that this may be extended to the general entangled case in the future, but we have provided heuristic arguments for why we do not expect much of an improvement for Eve by doing this.

We have described two novel ways of counteracting the THA; a passive defence, enabled by adding thermal noise into the system, and an active defence with an optical modulator. These complement the attenuation-based defence discussed in earlier work.

This all shows that side-channel attacks cannot be considered only as an afterthought in QKD systems. Even a relatively rudimentary SCA can, if not protected against, hugely reduce the security of a protocol. If we try to improve the security by privacy amplification alone we find that the secret key rate soon drops to zero. This highlights the importance of proper and specific defences against SCAs that are easily quantified in terms of experimentally accessible quantities.

Chapter 5

A practical repeater for ultra-long distance quantum communication

In Section 3.2 we discussed various proposals for quantum repeater protocols, designed to extend the range of secure quantum communication beyond the limit imposed by the exponential attenuation that single photons suffer when transmitted through optical fibres. However, the question as to how we should generate the initial Bell pairs between the repeaters in a way that retains a high fidelity in situations of non-negligible photon loss and decoherence, remains an open one. This is a crucial element of any proposal for a repeater network, and long distance quantum communication and distributed quantum computing will never be achieved without a strong solution to this problem. Additionally, while many photonic methods for generating Bell pairs exist, it is critical that we examine these in the context of a quantum repeater. The difficulty of constructing a repeater is not simply equal to the difficulty of constructing a single Bell pair multiplied by the number of repeater sections, since issues of technological feasibility and concurrent timing may render otherwise good methods to be unworkable.

In this chapter we address both of these issues by proposing a complete quantum repeater protocol based on doubled-heralding and *brokered Bell-state measurements*. Critically, these only makes use of existing technology which has been shown to work reliably in practice. We describe how the same equipment naturally provides a loss-tolerant way to perform all three parts of the protocol: high-fidelity entanglement generation, loss-tolerant indirect Bell measurements and state distillation. We consider specifically the application of distributing a secret key for secure communication, and an analysis of the relevant errors shows exceptional performance compared to similar protocols, which carries over to other applications which require shared entangled states. We demonstrate this using an in-depth analysis of the errors of the protocol. As such, this work may be constituted as forming a kind of “threshold theorem," such that if the stated parameters are met, one may be confident that the claimed rates will be practically achievable.

5.1 Techniques

Here we shall discuss in more depth some of the techniques that are used in our protocol. We have already encountered the idea of double-heralding in Section 3.1.4, which will be used for generating the elementary entanglement between the stations. We will cover the concepts that underpin our protocol's quantum memories and entanglement-swapping.

5.1.1 NV centres

A Nitrogen-Vacancy (NV) centre is a point defect within a diamond that has great use within quantum optics (see [DH76, HHM84, VOMG88] for original research and [NTD+14] for pioneering their use in scalable quantum information processing). It crucially contains two components which may each exhibit a two-dimensional (qubit) Hilbert space. The first of these is the spin state of the spin-1/2 nucleus of a ^{15}N isotope [NTD+14]. This is typically used to *store* information due to their long coherence times, which may be on the order of seconds [MKL+12]. The nuclear spins may also be easily initialised, and be measured by non-demolition measurements [NTD+14]. The defect also features an electronic spin. While this has a shorter coherence time, on the order of tenths of a millisecond [NTD+14], it can be used to easily read and write qubit information [JGP+04]. That is to say, it can serve as an emitter and a receiver of photonic qubits. Importantly, it also has the correct level structure [CDT+06] to serve as an emitter for the double-heralding protocol.

5.1.2 Brokering and distillation

Brokering, shown in Fig. 5.1 (a)-(e), is a procedure whereby two NV centres may be entangled in the graph state sense without destroying any existing entanglement that they might have with other centres. This is done by projecting existing entanglement relations on the electron spins onto the nuclear spin qubits. We then try and entangle the electron spin qubits by double-heralding as described above. While a failed attempt at entanglement generation requires us to reset the qubits involved, the fact that the existing entanglement is supported on a separate physical part of the NV centre means that the existing entanglement is not disturbed. When the electron-spin qubits are entangled, a microwave π pulse applied to each centre applies a controlled-not gate between the electron qubit to the nuclear qubit, entangling them [JGP+04]. Measuring the electron qubits then teleports the entanglement between the electron qubits down onto the nuclear spins, and measurement of the nuclear spins removes them from the chain of entanglement, which is equivalent to a Bell state measurement. While normal optical Bell state measurements by passive linear gates and no ancilla have a maximum efficiency of 50%, this procedure has an efficiency limited only by the fidelity of the gates involved and the decoherence

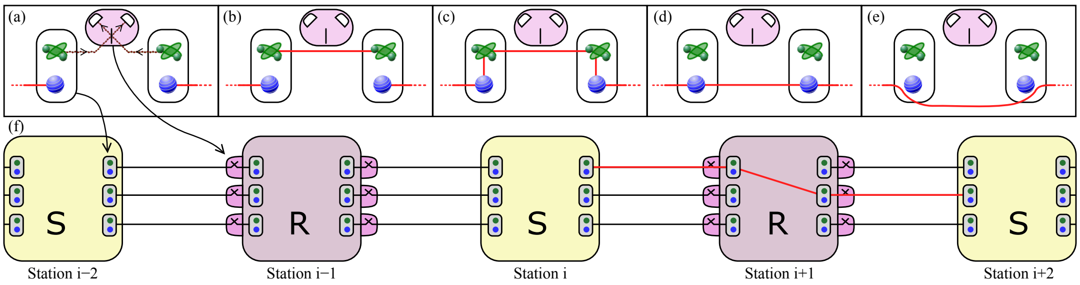


FIGURE 5.1: (a)-(e): Illustration of loss-tolerant indirect Bell measurements by brokering. Nuclear spin qubits are in blue and electron spin qubits in green. Red lines represent entanglement connections, which is meant in the sense of an "edge" in a graph state. In particular, the electron qubits are connected together by double-heralding in (a)-(b). In (c) these are entangled with the nuclear spins by microwave pulses, and the electron states are measured to give (d). Nuclear spins are then measured to give (e). (f): Four sections of the full repeater protocol, showing S and R type repeater stations. Type S stations send photons to the type R stations, which entangle one qubit from each side. The R stations then send classical signals back to the S stations, which create their own local entanglement. Local measurements then result in long-range Bell pairs between Alice and Bob.

times of the nuclear spins. This gate fidelity turns out to be the most important factor in determining the ultimate rate of generation of Bell pairs between Alice and Bob.

Due to the high fidelity of the Bell pairs that are generated between adjacent stations, this protocol creates high fidelity pairs between Alice and Bob even before any use of distillation. Nevertheless, distillation is a crucial ingredient in extending the reachable range. Here we propose to use the DEJMPS protocol ([DEJ$^+$96], Section 2.2.4).

In previous works on repeaters, it was suggested that this protocol may be unsuitable for use in a repeater network, since we require two-way communication to know which attempts have been successful [NTD$^+$16]. This requires waiting for a time equal to the travel time between distant stations, which we want to avoid since it leads to large decoherences. Alternative suggestions have involved using quantum computers and CSS codes [LBSB13, EKB15], but this goes against the philosophy of this work of constructing a simple system which only uses existing technology. The DEJMPS protocol is well suited to our system, since the CNOT gates involved can be implemented by a combination of brokered double-heralding and local rotations. We can avoid the necessity for long waiting times which may lead to decoherence by implementing *blind DEJMPS*, described in Section 2.2.4. We note that one might still want to use non-blind DEJMPS if gate errors outweigh memory errors.

5.2 Protocol summary

We now have the three essential elements to build the repeater network: the creation of long-range Bell pairs, the connection of these pairs within the repeater stations, and the distillation of states, all using the same system of NV centres and microwave pulses. The repeater stations are to be built in two types, type S and type R (for *sender* and *receiver*, shown in Fig. 5.1 (f)). Each station contains multiple qubits on each side (to connect to the stations before it and after it respectively). The presence of multiple qubits per station decreases the average time that it takes to make at least one entanglement connection between two adjacent repeaters, and so increases the rate of generation of Bell pairs between Alice and Bob even before applying distillation. The full protocol is then implemented as follows.

Type S stations send photons from their qubits to the type R stations before teleporting the state of electron spin qubit onto the nuclear spin. The type R stations use these to try to establish an entanglement connection by double-heralding. Once a type R station has established at least one entanglement connection to the type S stations on each side, it may deterministically entangle them together by using brokering to make a linear graph-like state [HEB04]. Classical signals are sent back to the type S repeaters bringing the information of which connections were successful. Once a type S repeater has received such a signal from either side, it may similarly perform a deterministic connection between these NV centres, leaving the final quantum state as a linear chain of entanglement from Alice to Bob via nuclear spins. These nuclear spins may then be removed from the chain by measuring in the computational basis (via projecting back up to the electron spin qubit) leaving Alice and Bob in possession of a pure Bell state. The manner in which distillation is carried out in this protocol is described in Section 5.3.3. The set-up outlined here could be used for any of the purposes for which we might want to have long-range entangled states. However, our analysis will focus on the particular case of QKD.

5.3 Analysis

We wish now to derive lower bounds on the secret key rates for both the cases with and without distillation. The main error sources which we identify in affecting the fidelity of the final state are dark counts in the detectors, mismatching the parameters of the NV centre cavities, failed gate operations when performing the indirect Bell measurements, and decoherence on the nuclear spins.

In considering the error analysis we may assume that all measurement results give the $+1$ result, so if all operations are successful Alice and Bob would expect to share $|\Psi^+\rangle\langle\Psi^+|$ as a final state (measurement results not equal to +1 can be accounted for in classical post-processing). We consider the worst case scenario where a single failed operation maps to the maximally-mixed state. The state shared between any two qubits can therefore be described by a Werner state. By "successful operation" we mean the quantum gates act as expected, the nuclear spins have not decohered, and we have not mistaken a dark count detection for a true detection from a double-heralding round. Let the product of these probabilities for an error *not* to occur be x, which is a function of the number of sections, N_S. The quantity that we want to maximise is the secret key rate, given by Eq. 3.8, where $\epsilon = (1 - x)/2$, which is the error rate averaged across bit and phase errors.

5.3.1 Dark counts

In assessing the effects of dark counts, the key parameter of interest is t_W, the *waiting time*. This is the time after the excitation of the electrons in the NV centres that we should wait in order to receive the emitted photons. If this is too small, we will miss the emitted photons, though if it is too great we will certainly measure a dark count, decreasing the fidelity of our states. It should be chosen to maximize K.

We model the dark counts (DCs) as a Poissonian process, so we say that the probability of measuring k dark counts in a time period t is

$$P(k \ \text{DCs in} \ t) = \frac{(t\Gamma)^k e^{-t\Gamma}}{k!}, \tag{5.1}$$

where Γ is the average dark count rate. Also recall from Section 2.1 that the excited electrons decay with time constant τ_q, so the probability that they will have *not* decayed and emitted a photon after time t is e^{-t/τ_q}.

Therefore, if we wait for a time t_W after each photon emission, the probability, P_1, that, after both rounds of emission and detection used for double-heralding we will have detected one photon in each round is

$$\begin{aligned} 2\,P_1 = {} & (1 - e^{-t_W/\tau_q})\,\bar{\eta}^2\,e^{-2t_W\Gamma} \\ & + 2(1 - e^{-t_W/\tau_q})\,\bar{\eta}\,e^{-t_W\Gamma} \cdot t_W\Gamma\,e^{-t_W\Gamma}\left[1 - (1 - e^{-t_W/\tau_q})\bar{\eta}\right] \\ & + \left[1 - (1 - e^{-t_W/\tau_q})\bar{\eta}\right]^2 \cdot t_W^2\Gamma^2\,e^{-2t_W\Gamma} \\ & + t_W^2\Gamma^2\,e^{-2t_W\Gamma}\left[1 - (1 - e^{-t_W/\tau_q})\bar{\eta}\right]^2 \\ & + t_W\Gamma\,e^{-2t_W\Gamma}\left[1 - (1 - e^{-t_W/\tau_q})^2\bar{\eta}^2\right], \end{aligned} \tag{5.2}$$

where $\bar{\eta} = e^{-L/L_{\text{att}}}\eta$ is the combined efficiency of transmission, photoemission, and photodetection (i.e. the probability that all were carried out successfully). The first three lines are the contributions from the states $|0,1\rangle$ and $|1,0\rangle$, representing two real detections, one real detection and one DC, and two DCs respectively. The last two lines give contributions from $|0,0\rangle$ and $|1,1\rangle$, representing two DCs, and one DC and one real count respectively. The overall factor of 2 is due to the fact that the initial state is equally weighted between $|0,1\rangle, |1,0\rangle$ and $|0,0\rangle, |1,1\rangle$.

From this we may calculate the contribution to x as

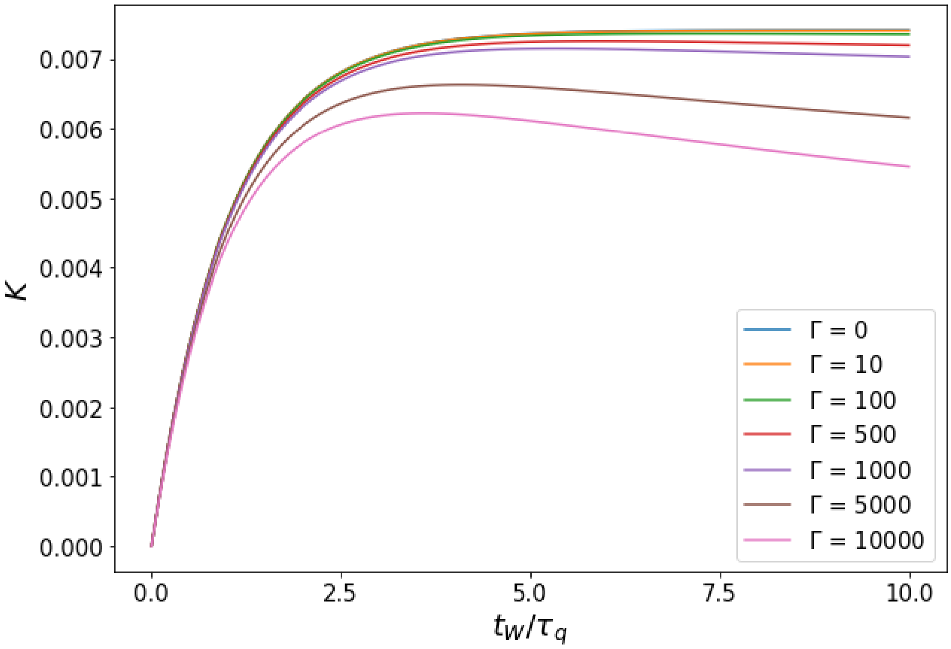


FIGURE 5.2: Secret key rate against waiting time. K here is the average number of secret bits generated per raw bit (limited by P_1). Here we have $L = 2L_{\text{att}}$, $\eta = 0.9$, and $\tau_q = 20$ ns.

$$x_{\text{d}} = \frac{1}{2}(1 - e^{-t_W/\tau_q})\,\bar{\eta}^2\, e^{-2t_W\Gamma}/P_1, \tag{5.3}$$

resulting in a fidelity of $F = (1 + 3\,x_{\text{d}})/4$. Substituting into Eq. 3.8 we get

$$K = R_0 P_1 \left[1 - 2h_2\left(\frac{1 - x_{\text{d}}}{2}\right)\right], \tag{5.4}$$

where R_0 is the rate of attempts at entanglement generation (so $R = R_0 P_1$). Fig. 5.2 plots this secret key rate against the ratio of waiting time to electron decay time-scale. It can be seen that the effects of dark counts are negligible, even at rates multiple orders of magnitude greater than what might be expected. It is therefore almost always better to err on the side of caution and make t_W larger than the optimal value of around $3\tau_q$. Note that the figure shows the results for the key rate of a single-section repeater over a distance of twice the attenuation length. As we shall find later (Fig. 5.3), the optimal inter-repeater distance never actually exceeds around 28 km, or 1.12 times the attenuation length. For a single-section repeater over this distance, and a realistic dark count rate of $\Gamma = 500$, the key rate is affected by dark counts by an amount of less than 0.3%. At $\Gamma = 100$, this becomes 0.08%. We shall therefore neglect the effect of dark counts in the remainder of this analysis.

It was found that for more a more realistic value of $\Gamma = 500$, the fidelity is affected by less than one part in 10^5.

5.3.2 Decoherence

In addition to dark counts the main sources of noise are mismatching of modes in adjacent cavities, gate fidelities, and decoherence of the nuclear spins. Mode mismatching has been shown to contribute to an error probability of less than 10^{-3} for mismatching either the Jaynes-Cummings constants or the cavity energy constants by up to 5% [KWD03]. We will include the effects of gate fidelities as a free parameter in our secret key rate, since they simply contribute a constant overhead at each station.

The error source which required the most consideration is the decoherence of the qubits. This is minimized by utilizing the long-lived nuclear spins, so has little effect on the fidelity for an individual section. The effect becomes pronounced when we consider the full system with $N_S \gg 1$, where N_S is the number of sections that are connected together to make the repeater network.

To see how the effects here may be analyzed, consider first the ideal case where every elementary section connects at the same time since the start of the protocol, $t_{\text{avg}} = p_{\text{c}}^{-1} L/c$, where p_{c} is the probability for us to make a connection between two adjacent stations in one attempt at double-heralding and L is the distance between repeater stations. This is the average time at which a connection between adjacent stations is made. This is given by

$$p_{\text{c}} = 1 - \left(1 - \frac{1}{2} e^{-2L/L_{\text{att}}} \eta^2\right)^q, \tag{5.5}$$

where L_{att} is the attenuation length, q is the number of qubit pairs per station, and η is the efficiency of photon emission and collection. We have set it equal to the product of the detector efficiency, channel transmissivity, and the coupling efficiency between the NV centre and the optical channel (which may be made deterministic [ESR+10]).

The only decoherence effects here will be a factor of $\exp(-N_S L/c\tau_{\text{d}})$ contribution to x as the spins decohere slightly while the signals are being sent from the type R stations to the type S stations. This is independent of p_{c} since the electron spin qubits are reset for each round of double-heralding. Even for $N_S = 100$ stations at $L = 25$ km, $\tau_{\text{d}} = 1$ s this is only a factor of $\sim 1 - 10^{-5}$ contribution to x. Note that we are not considering the contribution of the gate times, since these are mediated by microwave pulses which typically last around 50 ns, compared to the light travel time between stations on the order of tens of microseconds.

A more accurate analysis of the effects of decoherence must take into account the fact that the establishment of Bell pairs across different sections will not all occur at the same time, so the first section to be connected must be kept coherent until the last one has been completed. This is not simply a minor perturbation to the naïve situation described in the previous paragraph, since now the non-unit efficiencies of the detectors play a part.

For the set of N_S sections, let $\{T_k\}$ be the set of *order statistics*. That is to say, T_1 is the time at which the first connection is made, T_2 is the time of the second connection, and so on. For an elementary section between two given stations, let f_t be the probability that the connection is formed at a time t, and F_t be the probability that it is formed at a time less than or equal to t, given by

$$\begin{aligned} f_t &= (1-p_{\mathrm{c}})^{t-1}p_{\mathrm{c}}, \\ F_t &= 1-(1-p_{\mathrm{c}})^{t}. \end{aligned} \tag{5.6}$$

The average value of T_k is then given by

$$\begin{aligned} \langle T_k \rangle = \sum_{t=1}^{\infty} t \sum_{j=0}^{N_S-k} \binom{N_S}{j} \times \\ \left[(1-F_t)^j F_t^{N_S-j} - (1-F_t+f_t)^j (F_t-f_t)^{N_S-j}\right], \end{aligned} \tag{5.7}$$

By taking the worst case scenario that we connect all the odd-numbered sections first (so that we can't make any indirect Bell measurements until as late as possible), we have the following contribution to x from decoherence effects:

$$x_{\mathrm{e}} = \exp\left(-\frac{2L}{c\tau_{\mathrm{d}}}\left[\frac{N_S}{2} + \langle T_{N_S}\rangle + \sum_{k=k_{\mathrm{u}}}^{N_S}\langle T_k\rangle - \sum_{k=1}^{k_{\mathrm{l}}}\langle T_k\rangle\right]\right), \tag{5.8}$$

where $k_{\mathrm{u}} = \lceil (N_S+1)/2+1 \rceil, k_{\mathrm{l}} = \lfloor (N_S+1)/2 \rfloor$, $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ represent the ceiling and floor functions respectively, and τ_{d} is the decoherence timescale (T_2 time) of the nuclear spins. The additional factor N_S comes from the decohering of the nuclear qubits in the time between sending the photons for double heralding and detection

The final ingredient required in finding the overall rate is a decision on when to say that an attempt to make an end-to-end connection has finished, indicating to Alice and Bob that they should then measure their qubits in the Z basis to generate a key bit. This may be accomplished by one of two methods:

Method A: When the final section completes, a message is sent from it to Alice and Bob telling them to make the relevant measurements. This will be favorable when N_S and η are both low.

Method B: Decide on a fixed time, t_{f} (as a function of $\langle T_{N_S}\rangle$), at which Alice and Bob should make their measurements. This will be favourable when η is small (so T_{N_S} has a narrow distribution) or N_S is large. With this method, we will "miss out" on a fraction $\sum_{t=t_{\mathrm{f}}}^{\infty} P(T_{N_S} = t)$ of attempted connections.

We have found that method B, choosing $t_{\mathrm{f}} = \lceil \langle T_{N_S}\rangle + \delta \rceil$ for some buffer value, δ, is better for almost all choices of parameters, with roughly 90% of connection attempts being successful. However, the behaviour of K at a given δ can be highly erratic with varying N_S, so we should optimize our choice of δ individually for each choice of parameters.

Thus we finally arrive at our secret key rate given by

$$K = \underbrace{\max_{\delta} \sum_{t=1}^{\lceil \langle T_{N_S} \rangle + \delta \rceil} P(T_{N_S} = t) \frac{c}{L \lceil \langle T_{N_S} \rangle + \delta \rceil}}_{\text{Raw rate}} \cdot \underbrace{\left[1 - 2h_2 \left(\frac{1}{2} (1 - x_{\mathrm{d}}^{N_S} x_{\mathrm{mm}}^{N_S} x_{\mathrm{ga}}^{N_S - 1} x_{\mathrm{e}}(N_S)) \right) \right]}_{\text{Correction term}}, \tag{5.9}$$

where x_{d}, x_{mm}, x_{ga}, and x_{e} represent the contribution from dark counts, mode mismatching, Bell measurement gates and nuclear spin decoherence respectively. The raw rate is determined by the light travel time between stations, since (being at the millisecond scale) it is orders of magnitude longer than the timescales involved in referencing the NV centres.

5.3.3 Distillation

As we have noted earlier, there are two factors contributing to the key rate, K: the raw key rate, R, and the error rate, ϵ. Of these, ϵ is ostensibly the most important. While R may drop arbitrarily low and still give a $K > 0$, ϵ must stay above 0.11. The effect of the distillation that we consider here is to therefore ensure that the probability of no error occurring, x, is maintained always above some *critical probability*, x_{c}. This works as follows:

First, we calculate a number of sections, N_{long}. This is the largest N_S for which $x(N_S) > x_{\mathrm{c}}$, where $x(N_S)$ is the overall value for x as an explicit function of N_S. Entanglement over this distance is created, and distilled by DEJMPS. Before distillation, we assume that the errors that the state may undergo are entirely random. This means that the state will be a Werner state, given by

$$\rho = \rho_W \left(F = \frac{1 + 3x}{4} \right). \tag{5.10}$$

After distillation, the state is mapped to a non-Werner state. Since calculating the fidelity resulting from repeated application of DEJMPS distillation becomes analytically intractable, we wish to deal only with Werner states. At this point, we might choose to replace the result of the distillation operation with a Werner state of the same fidelity. This gives an upper bound on the error (and hence a lower bound on the rate) since a Werner state is the highest entropy state of a given fidelity. In Eq. 2.57 we show that the fidelity of a Werner state after one round of distillation is given by $F' = (10F^2 - 2F + 1)/(9N_D)$.

However, a better choice is to do one round of distillation, followed by an application of a local Hadamard gate[1] to all qubits, followed by another round of distillation before mapping to a Werner state. In this situation, the first distillation will deal with phase errors, and the second will deal with bit errors. If we only do the first distillation, we will not suppress bit errors. However, if we do two distillations before mapping to a Werner state, then we have a quadratic suppression of errors with respect to the first method. By considering the expressions for the state coefficients given in Eq. 2.54, we may see that the fidelity after two rounds of distillation (without mapping to a Werner state in between) is given by:

$$F'' = \frac{81F^4 + 5(1-F)^4 + 18F^2(1-F)^2}{81N_{D,2}}, \tag{5.11}$$

where $N_{D,2} = \left(F^2 + 3F_e^2\right)^2 + 4\left(FF_e + F_e^2\right)^2$, and $F_e = (1-F)/3$. Note that in all equations here, F is the fidelity of the Werner states with respect to $|\Phi^+\rangle$ before *either* distillation has been performed.

The advantage of distilling twice before mapping to a Werner state may be noticed most clearly when we consider the average of the post-distillation fidelities with the Bell states *other* than $|\Phi^+\rangle$. Before distillation, this average is equal to F_e. When using the method where we distil once, then the average of the relevant fidelities contains a term that is still linear in F_e. However, when we distil twice before mapping to a Werner state, the largest term in the fidelity with the "incorrect" states goes with F_e^3, making the advantage clear.

We then identify another number of sections, N_{short}. To highlight the purpose of N_{short}, it is clearest to indicate what it *would* be, if we were only distilling *once* before assuming we map back to a Werner state. In this case, it would be defined as the largest N_S such that $[4F'/3 - 1/3] \cdot x(N_S) > x_c$, where the factor on the left-hand side is the probability of no error occurring corresponding to a Werner state of fidelity F'. That is to say, we let the state decay to the point where there is a probability of error of $x_c = 4F/3 - 1/3$, we perform a distillation to reach fidelity F', corresponding to a probability of error of $4F'/3 - 1/3$, and then connect over another N_{short} sections, leading to the probability of error being multiplied by $x(N_S)$. It is therefore this number of sections that we can afford to connect to the first N_{long} sections before we need to distil again, such that x stays above x_c.

However, since we are in fact distilling *twice* before reducing our state to a Werner state, we instead say that N_{short} is the solution to

$$\frac{4F'' - 1}{3} \cdot x(2N_S) > x_c. \tag{5.12}$$

This is solved by using Eq. 5.11, where $F = (1 + 3x_c)/4$.

When we distil sets of qubit pairs together, we are at best left with 50% of the number of pairs that we started with. This must be multiplied by the probability that

[1] This maps $|0\rangle \leftrightarrow |+\rangle$ and $|1\rangle \leftrightarrow |-\rangle$. When applied to both parts of the Bell states, it maps $|\Phi^+\rangle \rightarrow |\Phi^+\rangle$, $|\Psi^-\rangle \rightarrow -|\Psi^-\rangle$, $|\Psi^+\rangle \rightarrow |\Phi^-\rangle$, and $|\Phi^-\rangle \rightarrow |\Psi^+\rangle$.

the distillation succeeded. This is given by Eq. 2.56. For a Werner state described by Eq. 5.10 for $x = x_\mathrm{c}$, this is given by

$$N_D = \frac{1 + x_\mathrm{c}^2}{2}. \tag{5.13}$$

Therefore, for $N_S \geq N_\mathrm{long}$ (the regime where we intend to start distillation) we get a secret key rate for our protocol of

$$K_\mathrm{dist}(x_\mathrm{c}) \geq R_\mathrm{raw} \left(\frac{1 + x_\mathrm{c}^2}{4}\right)^{\frac{N_S - N_\mathrm{long}}{N_\mathrm{short}}} \left[1 - 2h_2(\frac{1 - x_\mathrm{c}}{2})\right], \tag{5.14}$$

where we are using $\geq$ instead of $=$ since we fix x at the lower bound of x_c. Here, R_raw is the raw rate term from Eq. 5.9. Unlike Eq. 5.9 this never drops below zero (since we effectively pin x at x_c) but at an exponential cost in the raw rate. To get the best key rate at a given distance, we maximise over choices of $x_\mathrm{c} \in [0.78, 1]$.[2]

We emphasize here that we are considering all noisy pairs to be the same. That is to say, the k^th order statistic, T_k, for any given connection attempt is given by its expectation value. In reality, some connections are going to be established sooner than others and so will have a higher fidelity. There remains the open question of how best to pair up non-identical noisy pairs taken from some distribution, that is addressed in Chapter 6 and Appendix C.

5.4 Performance

In Fig. 5.3 we analyse the achievable secret key rates for a range of distances (using blind distillation). For a fair comparison, we consider two measures of the key rate. Firstly, we consider the key rate *per channel*. This is the key rate normalised by the number of parallel channels, q. This makes for a fair comparison, since K may be made arbitrarily large by increasing q. Note that this is not the same thing as normalising by the total physical resources, which is given by the key rate divided by the total number of qubits, $2qN_S$, which we also show. We note that, unlike q, increasing N_S does not necessarily increase K, since it also increases the probability of a local gate error occurring. We show the achievable results for various values of the local gate error probability, x_ga, which turns out to be of critical importance in determining the key rate. In addition to the realistic scenarios of $x_\mathrm{ga} = 0.95$ and $x_\mathrm{ga} = 0.99$, we show the optimized rate in line with the gate quality that is necessary for fault-tolerant quantum computing, $x_\mathrm{ga} = 0.999$. It is noteworthy that if such gate qualities are reached, then this protocol will enable secure communication on the intercontinental range. The rates shown here are lower bounds, since we are not including the effects of parallelisation. In reality, when one section forms a connection across one of its pairs of qubits, the others will keep attempting to make connections

[2] Choosing x_c to be below 0.78 results in $K = 0$.

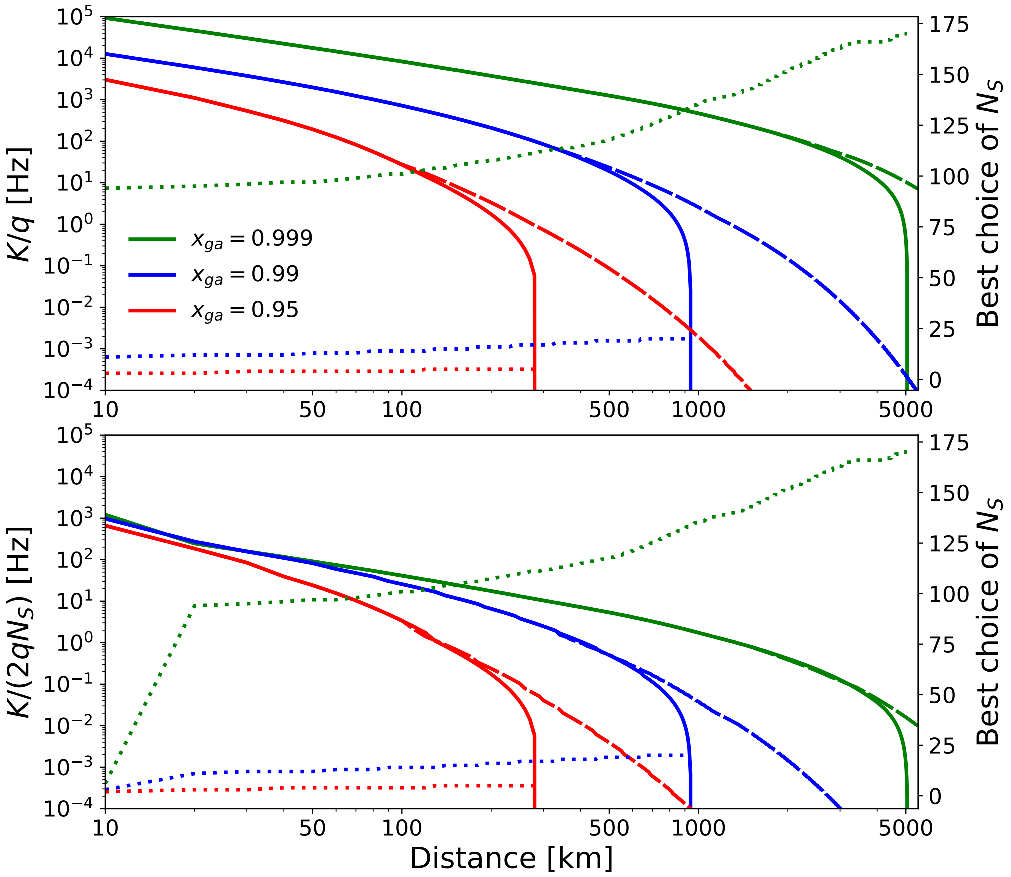


FIGURE 5.3: Attainable secret key rates for different values of gate quality, x_{ga}, optimised over number of repeaters N_S. Top: key rate per channel = K/q, bottom: key rate per qubit = $K/(2qN_S)$. Solid lines are without distillation, long dashed lines are with. Short dashes show the optimal values of N_S, with the scale shown on the right-hand axis.

while waiting for the other sections to connect, meaning the true rate is likely to be far higher.

At each distance we have optimised over the choice of N_S, and shown the optimal choice at each distance. As expected, this increases with distance. We also note that for low values of x_{ga}, the optimal N_S is lower. This is because gate errors become more relevant than decoherence errors in this case. All plots are shown for $\eta = 0.9$, $q = 10$, and $L_{\text{att}} = 25$ km.

We assume here that we attempt to detect all emitted photons. In various physical implementations this may not be the case, and we may wish to post-select on some fraction μ of the photons. An example of this may be in NV centres where we wish to use the zero-photon line (since only these photons are perfectly entangled with the emitting centre). Recent advancements have produced NV centres with Debye-Waller factors of 0.4, requiring a seven-fold increase in the number of qubits needed per station [JDG+15].

5.5 Comparison with other protocols

To get a good idea of the quality of the protocol, it is necessary to have some point of reference. The most naive alternative to our repeater protocol would be a system where we simply have q quantum channels direct from Alice to Bob. Our protocol easily beats this system. At a range of 100 km, our protocol gives normalised rates of $10^2 - 10^4$, whereas the naive protocol, even if there is no loss of fidelity over the long range, would have a key rate of $e^{-4} \approx 10^{-2}$.

We have also compared our setup to various others in the existing literature. For a meaningful comparison, it is of course necessary to match the choices of experimental parameters. Since we are comparing against theoretical proposals, there are only a few parameters which are meaningful across all proposals. One such parameter is the local gate efficiency (x_{ga} in our work), which we have matched with relevant parameters in comparison protocols. Another well-defined value is the resource cost, such as the number of memories, stationary qubits, or subsystems. For a meaningful comparison, we consider here the *normalised* secret key rates, which are the secret key rates divided by the number of qubits used. The third well-defined parameter that we identify is the repetition rate. However, this is included in the overall secret key rate, which we have explained above to be dominated by the light travel time, so does not require additional attention.

One of the closest schemes conceptually is that of Nemoto et. al. [NTD+14]. This also uses NV centres, but transfers entanglement between the nodes by encoding information in the polarisation of a single photon, which requires the nodes to be equipped with single-photon detectors. A value for local-gate errors of 0.3% is used, corresponding to a secret key rate of the top line on our Fig. 5.3. For total distances of 200 km and 500 km, Nemoto's protocol gives normalised secret key rates of approximately 16 and 3 (where the normalisation divides by the number of qubits used), whilst ours gives normalised rates of 159 and 75. This order-of-magnitude improvement is particularly striking when we consider the fact that the two analyses are greatly different. In our protocol we have assumed that the rate of bit errors and phase errors are the same, since the application of indirect Bell measurements to connect two Bell pairs may give any one of the four Bell states as a result, dependent on the outcomes of the measurements. This results in a mixing of phase and bit errors, whereas Nemoto et al. consider phase errors to be dominant, resulting in a key rate correction term of only $1 - h_2(\epsilon)$, rather than the standard $1 - 2h_2(\epsilon)$.

Additionally our protocol beats other realistic linear-optical repeater schemes such as [SDRA+07, DBCZ99, KGD+15, PR15] by some orders of magnitude, however gives lower rates than proposals based on advanced encoding schemes [ATL15, MSD+12]. This is to be expected, since our proposal falls within the category of schemes that are simple to build and do not require large encoded states. In the intermediate regime, there are other protocols. One such is the measurement-based scheme of Ref. [ZDB12], which gives lower normalised rates in the regime of a few thousand kilometers, but has greater reachable distances.

5.6 Summary

We have presented a protocol for a quantum repeater network that allows for greater reachable distances and higher secret key rates than other methods in the literature, yet is implementable using today's technology. Unlike most other proposals for such networks, the fidelity of the elementary links is not affected by photon loss, nor detectors that do not perfectly count photon number. We have demonstrated that this leads to excellent secret key rates over thousands of kilometers, given sufficiently high gate fidelities. This gives a strong indication that we may be able to have absolutely secure communication over intercontinental distances in the near future.

Chapter 6

Statistical analysis of quantum entangled network generation

We have seen already how the ability to construct large-scale quantum networks between two or more parties is a necessary precursor to the general deployment of entanglement-based quantum key distribution as a ubiquitous alternative to classical encryption [Eke91, VV14], as well as the creation of measurement-based quantum computers [BBD+09]. Implementations of such networks would range from Bell states for point-to-point communication over large distances [Bel64, BDCZ98], to highly connected cluster states [KL10] and a complete distributed quantum Internet [Kim08, CCB18], that may itself involve non-trivial routing problems [Cal17]. Many theoretical proposals have been put forward for different schemes to implement these tasks, and in general the construction of these quantum networks requires the use of probabilistic elements. For example, many probabilistic methods for the generation of entanglement between nodes of a network have been proposed [BK05, CB08, BBFM06, CTSL06], as well as many high-level schemes that take advantage of such methods, such as entanglement-based quantum repeaters [PKEG16, BRS10, MBF10, VK17, ATL15, DLCZ01] and entanglement distillation [DEJ+96, DB07]. Probabilistic methods are also used in the implementation of non-linear unitary operations on optical states, such as those used in linear optical quantum computation [KLM01, BR05, KMN+07] and code-based repeaters [MSD+12, RHG05], as well as schemes for making measurements of states in a way that is protected against particle loss [VBR06]. The presence of such probabilistic components means that a complex composite protocol will likely take many attempts before completing its task. When a single element fails, this could result in the entire process, or a subsection of it, needing to be restarted. It may also result in *waiting errors*. This is where one part of the protocol finishes, but accumulates errors while waiting for another part to complete. This was analysed briefly in Chapter 5, particularly in Section 5.3.2, although there is much more to be said on the matter for such an analysis to be treated as a truly accurate representation of the system.

Typically, in many analyses of quantum network systems, the full depth of statistical information that may be gleaned from the full probability distributions over completion times or error distributions is neglected in favor of a simpler analysis, such as analysing the average values. However, this can result in too limited a characterization of the protocol, and one that may miss essential features. For example, a situation that is commonly considered in the context of quantum communication is the time taken to generate a set of entangled states between Alice and Bob, which may be distilled in order to generate a smaller number of higher-fidelity pairs. If we consider that all pairs connect after some average time, t, then the secret key rate will scale linearly with the number of states that we are trying to connect in parallel, and inversely proportional with t. In reality, not all pairs of entangled states will establish at the same time. However, if we intend to use all of them for distillation, then the pairs that establish first will have to be stored on quantum memories, and the fidelities of these states will decay while they wait for the other pairs to complete. It will therefore not necessarily be advantageous to have a greater number of pairs try to establish their entanglement in parallel. A good understanding of the distribution of times taken by a protocol and the corresponding error probabilities is thus essential for any analysis of a protocol.

We begin with some general methods that may be used for the analysis of probabilistic processes using Markov chain analysis. Markov chains have recently been applied to quantum networks by Shchukin, Schmidt and van Loock [SSvL17]. We build upon these techniques in order to include errors in a natural way, as well as introducing new analytic techniques to greatly reduce the computational burden that comes with any deep analysis of Markov chains. In Section 6.1 we explain how one may construct Markov matrices for probabilistic processes, and how the matrices for larger compound processes can be constructed from the matrices of smaller processes. We show also how we can find $\{p_t | t \in \mathbb{N}\}$ from such matrices, where p_t is the probability that the process will complete at time t. In Section 6.2 we show how one may find the probability-generating function (PGF) from the Markov matrix. We then show how one may solve the PGF to find the completion time distribution such that the computational complexity of finding p_t is decreased by a factor of $n_{\mathcal{P}}$ compared to using the matrix alone, where $n_{\mathcal{P}}$ is the dimension of the matrix.

In Section 6.3 we show how to calculate the probability distribution for the number of times that a given event in a process occurs. This rather general method may be used to calculate the distribution of the number of errors that will accumulate in the running of a process, both on average and conditioned on the completion time.

In Section 6.4 we examine a modification of the Innsbruck protocol for distillation-based quantum repeaters [BDCZ98], where the available quantum memories at a repeater station are bunched. By this, we mean we separate the available pairs of quantum memories between each pair of repeater stations into bunches of fixed size which are then distilled once all entanglement connections within the bunch have completed. We apply the techniques developed here to estimate the best values for the sizes of these bunches. This allows for a richer characterization of the secret key rates reachable by a protocol than may be learned from an analysis that does not

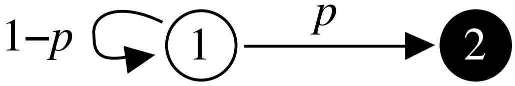


FIGURE 6.1: Graph for simple probabilistic entanglement generation. Each node represents a state that the system may be in at any one time. Transitions between nodes are considered to all take the same length of time, and occur with probabilities indicated by the weight of the edge. Terminating node shown in black.

account for the statistical factors that are captured by the Markov chain formulation. Finally, in Section 6.4.3 we consider a simplification of our statistical analysis of the Innsbruck protocol. By considering bounds on the order statistics of completion times of certain elements within the protocol, we derive bounds on the secret key rates. This allows us to identify minimum experimental parameters that must be reached in order to securely communicate over a repeater network of many sections.

6.1 Markov chains

Let $\mathcal{P}$ be some process that may be decomposed into events taking place across a series of discrete time-steps. This process may be summarised by a directed graph, $G_{\mathcal{P}}$, which is a flowchart showing possible paths of progression. Each node represents a unique state that the process may be in at any one time. The edges leading away from each node are the possible events (with the traversal of an edge being considered to take one time-step), with the weight of each edge representing the probability that that step will be taken. Each graph must include at least one terminating node (with no edges leaving it) representing the termination of the process. For example, if $\mathcal{P}$ is the protocol of establishing entanglement between a single pair of quantum memories by a probabilistic process that succeeds with probability p, then $G_{\mathcal{P}}$ is given by Fig. 6.1 (see [1]). Since $\mathcal{P}$ is probabilistic, the time that it takes to complete is represented by a random variable, T, that takes on value t with probability p_t.

From here, we can form the square Markov matrix for the process, $M_{\mathcal{P}}$, which is the adjacency matrix of $G_{\mathcal{P}}$.[2] That is to say, $[M_{\mathcal{P}}]_{i,j}$ is the weight of the edge of $G_{\mathcal{P}}$ leading from node j to node i for some fixed labeling of $G_{\mathcal{P}}$. This means that the j^{th} column contains the transition probabilities *away* from node j. This immediately gives us an operational method to find $\{p_t\}$. If we let $I_{\mathcal{P}}$ be the set of indices for terminating nodes, then we may say that

$$p_t = \sum_{i \in I_{\mathcal{P}}} \left[M_{\mathcal{P}}^t\right]_{i,1}, \tag{6.1}$$

[1] Here we have used the convention that an absorbing node of the process has no edges leading away from it. Many standard texts on Markov chains use the convention that absorbing nodes should transition to themselves with probability 1.

[2] These Markov matrices are such that the columns sum to either 0 (for terminating nodes) or 1.

where the node representing the start of the process is given the label 1. For example, for the system shown in Fig. 6.1, we have

$$M_{\mathcal{P}} = \begin{bmatrix} 1-p & 0 \\ p & 0 \end{bmatrix}. \tag{6.2}$$

If, for example, we wish to know the probability of a process completing in exactly 2 time steps, we would look at $M_{\mathcal{P}}^2$, which is given by

$$M_{\mathcal{P}}^2 = \begin{bmatrix} (1-p)^2 & 0 \\ p(1-p) & 0 \end{bmatrix}. \tag{6.3}$$

We would then see that $[M_{\mathcal{P}}]_{2,1} = p(1-p)$. This is indeed the probability of completing in 2 time-steps, since event 1 must fail once and succeed once for this to be the case.

From the matrices for simple processes we can build up matrices for more complex processes. Consider two processes $\mathcal{P}_1$ and $\mathcal{P}_2$. We wish to concatenate these to form the process $\mathcal{P}_3$, which consists of $\mathcal{P}_1$ and $\mathcal{P}_2$ being run simultaneously but independently. For each unique pair of states with one chosen from $\mathcal{P}_1$ and one chosen from $\mathcal{P}_2$, we should assign a unique state in $\mathcal{P}_3$. Additionally, for two such pairs of states, $s^1_{\mathcal{P}_1}, s^1_{\mathcal{P}_2}$ and $s^2_{\mathcal{P}_1}, s^2_{\mathcal{P}_2}$, then the independence of $\mathcal{P}_1$ and $\mathcal{P}_1$ implies that the probability to move from the state representing $(s^1_{\mathcal{P}_1}, s^1_{\mathcal{P}_2})$ to $(s^2_{\mathcal{P}_1}, s^2_{\mathcal{P}_2})$ in $\mathcal{P}_3$ should be given by $p(s^1_{\mathcal{P}_1} \to s^2_{\mathcal{P}_1}) \cdot p(s^1_{\mathcal{P}_2} \to s^2_{\mathcal{P}_2})$. Therefore if we consider $\mathcal{P}_3$ to have finished when $\mathcal{P}_1$ *or* $\mathcal{P}_2$ has finished, then our combined matrix is the tensor product of the constituent matrices:

$$M_{\mathcal{P}_3} = M_{\mathcal{P}_1} \otimes M_{\mathcal{P}_2}. \tag{6.4}$$

We may instead wish to wait until both $\mathcal{P}_1$ *and* $\mathcal{P}_2$ have completed before considering $\mathcal{P}_3$ to have completed. In this case we should add an element to the matrix for each subprocess that keeps the system on that terminating node until the other subprocess has completed. The composite matrix is therefore

$$\begin{aligned} M_{\mathcal{P}_3} =& [M_{\mathcal{P}_1} + \mathrm{diag}(\mathbf{I}_{\mathcal{P}_1})] \otimes [M_{\mathcal{P}_2} + \mathrm{diag}(\mathbf{I}_{\mathcal{P}_2})] \\ &- \mathrm{diag}(\mathbf{I}_{\mathcal{P}_1} \otimes \mathbf{I}_{\mathcal{P}_2}). \end{aligned} \tag{6.5}$$

where $\mathbf{I}_{\mathcal{P}}\big|_i = 1$ if $i \in I_{\mathcal{P}}$, and 0 otherwise, and $\mathrm{diag}(\mathbf{I})$ is a matrix with the elements of $\mathbf{I}$ on the diagonal, and with zeros elsewhere.

Suppose that instead we consider $\mathcal{P}_3$ to consist of $\mathcal{P}_1$ followed by $\mathcal{P}_2$. When we reach the terminating nodes of $\mathcal{P}_1$, the next time-step will have us arrive at the first node of $\mathcal{P}_2$. Then

$$[M_{\mathcal{P}_3}]_{i,j} = [M_{\mathcal{P}_1} \oplus M_{\mathcal{P}_2}]_{i,j} + \sum_{k \in I_{\mathcal{P}_1}} \delta_{i,k}\delta_{j,n_{\mathcal{P}_1}+1} \tag{6.6}$$

where $n_{\mathcal{P}_1}$ is the number of nodes in $G_{\mathcal{P}_1}$ or the number of rows or columns in $M_{\mathcal{P}_1}$.

It may also be the case that different parts of a process take different lengths of time, instead of the above construction which assumes that each event takes a single time-step. Suppose that within some process, $\mathcal{P}$, we have some events (edges on $G_{\mathcal{P}}$) that take some time k_1, and some that take k_2, where $k_2 \geq k_1$. We can decompose $M_{\mathcal{P}}$ as $M_{\mathcal{P},k_1} + M_{\mathcal{P},k_2}$, such that all elements in $M_{\mathcal{P},k_1}$ represent events that take k_1, and similar for $M_{\mathcal{P},k_2}$. From this we can create a new process matrix $M'_{\mathcal{P}}$ which properly accounts for the fact that events in subprocess P_1 can be done many times for each time that P_2 can be done. This is given by

$$\begin{aligned}\left[M'_{\mathcal{P}}\right]_{i,j} = \sum_k \Bigg\{ \left[M_{P,k_1}^{\lceil k_2/k_1 \rceil}\right]_{i,k} + \\ \left(1 - \sum_i \left[M_{P,k_1}^{\lceil k_2/k_1 \rceil}\right]_{i,k}\right) \delta_{i,k} \Bigg\} \left[M_{P,k_2}\right]_{k,j}\end{aligned} \tag{6.7}$$

where the term in parentheses ensures that the system does not move off terminating nodes.

As such, p_t calculated from $M'^{t}_{\mathcal{P}}$ will represent the probability that the process completes after t applications of P_2 and tk_2/k_1 applications of P_1. It should be noted that the modification of process matrices to account for timing differences should be done *before* creating composite matrices by tensor products.

Using Markov matrices along with Eq. 6.1 is a simple way to calculate the completion times of a process, although it is not necessarily the most efficient. Multiplying $M_{\mathcal{P}}$ by itself takes $n_{\mathcal{P}}^{k_{\mathrm{MM}}}$ elementary multiplications, where $k_{\mathrm{MM}} = 3$ for the naive approach of matrix multiplication, and the best known asymptotic result is $k_{\mathrm{MM}} \approx 2.373$. Given some algorithm for calculating exponentials that has a number of operations that scales asymptotically as $f_{\exp}(t)$ for the calculation of k^t for some constant k, we find that the calculation of p_t scales asymptotically as

$$\mathcal{O}(p_t \text{ by matrix mult}) = \mathcal{O}\left(n_{\mathcal{P}}^{2.373} f_{\exp}(t)\right). \tag{6.8}$$

In the next section we derive a method by which this may be reduced to being linear in $n_{\mathcal{P}}$.

6.2 Probability generating functions

In this section we show how an approach based on probability generating functions (PGFs) and complex analysis can lead to formulas for p_t that are faster to compute than the matrix multiplications of Eq. 6.1.

The probability generating function of a distribution $\{p_t\}$ corresponding to the completion times for a process $\mathcal{P}$ is defined as the polynomial

$$f_{\mathcal{P}}(z) = p_0 + p_1 z + p_2 z^2 + p_3 z^3 + \cdots , \tag{6.9}$$

where z is a complex variable, and p_t are constants to be determined based on $\mathcal{P}$. Given the PGF associated with some process, the elements p_t may be found by calculating the coefficients of the various terms by finding the derivatives:

$$p_t = \frac{1}{t!} \frac{\mathrm{d}^t f_{\mathcal{P}}(z)}{\mathrm{d}z^t} \bigg|_{z=0} . \tag{6.10}$$

In order to write down the PGF, it may seem like we need to already know all of $\{p_t\}$. However, we can calculate $f_{\mathcal{P}}(z)$ directly from $G_{\mathcal{P}}$. Consider a node in $G_{\mathcal{P}}$, x, with one edge leading to node y with probability 1. Let $f_{\mathcal{P}}^{(x)}(z)$ be the PGF for the system when we start at node x. Since the system will take exactly one timestep longer to complete when we start at x than when we start at y, we can say that $f_{\mathcal{P}}^{(x)}(z) = z f_{\mathcal{P}}^{(y)}(z)$. Now suppose that x has two edges leading away from it to nodes y_1 and y_2 with probabilities $p(y_1)$ and $p(y_2)$ respectively. Then, $f_{\mathcal{P}}^{(x)}(z) = z\, p(y_1) f_{\mathcal{P}}^{(y_1)}(z) + z\, p(y_2) f_{\mathcal{P}}^{(y_2)}(z)$. By extension, we may say that

$$f_{\mathcal{P}}^{(j)}(z) = \begin{cases} \sum_i \left[M_{\mathcal{P}}\right]_{i,j} z f_{\mathcal{P}}^{(i)}(z) & \text{if } j \notin I_{\mathcal{P}}, \\ 1 & \text{if } j \in I_{\mathcal{P}}, \end{cases} \tag{6.11}$$

where the sum runs over all columns in the matrix, which is an eigenvalue equation. The PGF of the process as a whole $[f_{\mathcal{P}}(z)]$ may be identified with the PGF of the initial node $[f_{\mathcal{P}}^{(1)}(z)]$. In particular, $f_{\mathcal{P}}(z)$ is the first element of the eigenvector of $\tilde{M}_{\mathcal{P}}(z)$ with eigenvalue 1, normalised such that the k^{th} element is 1 for any $k \in I_{\mathcal{P}}$, where

$$\tilde{M}_{\mathcal{P}}(z) = z M_{\mathcal{P}}^T + \operatorname{diag}(\mathbf{I}_{\mathcal{P}}). \tag{6.12}$$

However, a problem may arise in the process of finding the set of eigenvectors. We wish to retain z as an open variable in the PGF, which means that many of the fast methods for finding eigenvalues of matrices cannot be used, since they rely on numerical methods. In order to find an eigenvector of a completely general matrix, M, we need to be able to solve the characteristic equation $|M - \lambda \mathbb{1}| = 0$. This involves exactly solving a polynomial of order $n_{\mathcal{P}}$, which cannot in general be done for $n_{\mathcal{P}} \geq 5$. Instead, we use the fact that the eigenvalue is 1, so that $\tilde{M} \boldsymbol{f}_{\mathcal{P}} = \boldsymbol{f}_{\mathcal{P}}$, where $\boldsymbol{f}_{\mathcal{P}}$ is the vector with i^{th} element equal to $f_{\mathcal{P}}^{(i)}$, and say that

$$\begin{aligned} f_{\mathcal{P}}(z) &= [\boldsymbol{f}_{\mathcal{P}}]_1 \,/\, [\boldsymbol{f}_{\mathcal{P}}]_k \\ \boldsymbol{f}_{\mathcal{P}} &= \operatorname{Null}\left[\tilde{M}_{\mathcal{P}}(z) - \mathbb{1}\right], \end{aligned} \tag{6.13}$$

for any $k \in I_{\mathcal{P}}$. Note that we have used a slight abuse of notation and specified that $\boldsymbol{f}_{\mathcal{P}}$ is equal to the null space itself and not a particular vector in the null space. This is because the null space has a dimension of 1. We can see this by the fact that, if $M_{\mathcal{P}}$ is a Markov matrix, then $\tilde{M}^T_{\mathcal{P}}(z)$ must also be Markovian at $z = 1$. Moreover, the sum of all values in each column of $\tilde{M}^T_{\mathcal{P}}(1)$ will equal 1, which means that $\tilde{M}^T_{\mathcal{P}}(1)$ fits the usual definition of a stochastic matrix found in standard Markov chain textbooks. All such stochastic matrices have exactly one eigenvalue at 1 [Pri13], and so the other eigenvalues of $\tilde{M}_{\mathcal{P}}(z)$ must either be never equal to 1 or be z-dependent.

Having found the PGF, we want to use it with Eq. 6.10 to find $\{p_t\}$. Manually calculating the first few derivatives of the PGF may be possible. However the task soon becomes difficult for higher-order terms. By using Cauchy's differential formula [MH12], we find not only an easy way to compute higher derivatives, but a closed-form expression for an *arbitrary* derivative that can easily be calculated without needing to calculate all lower derivatives. The formula states that for some arbitrary point $a \in \mathfrak{S}$,

$$\frac{1}{t!}\frac{\mathrm{d}^t f_{\mathcal{P}}(z)}{\mathrm{d}z^t}\bigg|_{z=a} = \frac{1}{2\pi i}\oint_{\partial\mathfrak{S}} \frac{f_{\mathcal{P}}(z)}{(z-a)^{t+1}}\mathrm{d}z, \tag{6.14}$$

where $\partial\mathfrak{S}$ is the boundary of $\mathfrak{S}$; a compact subset of $\mathbb{C}$ on which $f_{\mathcal{P}}(z)$ is analytic.

Let $a = 0$ and $g_t(z) = f_{\mathcal{P}}(z)/z^{t+1}$. We will evaluate the integral of $g_t(z)$ on a circle centered on $z = 0$. If the contour encloses no poles except the one at 0 due to the $z^{-(t+1)}$ term, then this is equivalent to finding the residue of the pole of $g_t(z)$ at 0. Suppose that $f_{\mathcal{P}}(z)$ scales as $\mathcal{O}(z^{t_0})$ as $|z| \to \infty$. Then the integral of $g_t(z)$ on a circular path of radius R will tend to 0 as $R \to \infty$ for all $t > t_0$ (since $\mathrm{d}z = |z|\,\mathrm{d}\theta$). However, by Cauchy's residue theorem, this integral must *also* be equal to $2\pi i$ times the sum of all residues of $g_t(z)$ in $\mathfrak{S}$. This includes the pole at 0, which we get from the $z^{-(t+1)}$ term, and the poles elsewhere in the complex plane, which are the poles of $f_{\mathcal{P}}(z)$. Therefore the sum of the residues of all poles must be equal to 0 for $t > t_0$. The residue at $z = 0$ cannot be easily directly calculated since it is a non-simple pole, but we can calculate it indirectly since we know it must be equal to the negative of the sum of the residues of the other poles, which are in general simple. We therefore arrive at our main original result of this section:

$$p_t = -\sum_i \mathrm{Res}\left[\frac{f_{\mathcal{P}}(z)}{z^{t+1}}, z_i \in \mathbb{P}(f_{\mathcal{P}})\right], \tag{6.15}$$

where $\mathbb{P}(f)$ is the set of singularities of $f_{\mathcal{P}}(z)$.

As a corollary, we may use this method to easily find expectation values for completion times of such processes. Consider that

$$\langle T\rangle = \sum_t t\ p_t. \tag{6.16}$$

If we use the fact that

$$\mathrm{Res}\left[\frac{f_{\mathcal{P}}(z)}{z^{t+1}}, z_i\right] = \frac{\mathrm{Res}\left[f_{\mathcal{P}}(z), z_i\right]}{{z_i}^{t+1}}, \tag{6.17}$$

since $f_{\mathcal{P}}(z)$ has no pole at 0, we can write Eq. 6.16 as

$$\langle T \rangle = -\sum_i \mathrm{Res}\left[f_{\mathcal{P}}(z), z_i\right] \sum_t \frac{t}{z_i^{t+1}}, \tag{6.18}$$

where the z_i sum is implicitly over the poles of $f_{\mathcal{P}}(z)$. We can use the identity $\sum_{n=0}^{\infty} n\,x^{n-1} = (1-x)^{-2}$ to find the sum over the t–dependent terms, giving

$$\langle T \rangle = -\sum_i \frac{\mathrm{Res}\left[f_{\mathcal{P}}(z), z_i\right]}{(1-z_i)^2}, \tag{6.19}$$

Similarly, we can find the probability of the process completing *by* t, and the variance of the completion times:

$$p(T \leq t) = \sum_i \frac{1 - z_i^{-t-1}}{1-z_i} \mathrm{Res}\left[f_{\mathcal{P}}(z), z_i\right], \tag{6.20}$$

$$\mathrm{Var}(T) = \sum_i \frac{1+z_i}{(1-z_i)^3} \mathrm{Res}\left[f_{\mathcal{P}}(z), z_i\right] - \langle T \rangle^2, \tag{6.21}$$

We may note now that, if $f_{\mathcal{P}}(z)$ is built up constructively, as in Eq. 6.11, each non-terminating node contributes a single factor of z to the PGF. This means that $f_{\mathcal{P}}(z)^{-1}$ must be of order $n_{\mathcal{P}} - |\mathbf{I}_{\mathcal{P}}|^2$ at most, and so have no more than $n_{\mathcal{P}} - |\mathbf{I}_{\mathcal{P}}|^2$ poles. When calculating p_t by Eq. 6.15, the residues of $f_{\mathcal{P}}(z)/z$ are not t-dependent. Therefore, when we vary t, we simply need to calculate z_i^t for each $z_i \in \mathbb{P}(f_{\mathcal{P}})$. Given again some algorithm for calculating exponentials k^t in $\mathcal{O}(f_{\exp}(t))$ operations, we have that calculating p_t now scales asymptotically as

$$\mathcal{O}(p_t \text{ by Cauchy}) = \mathcal{O}\Big(\Big[n_{\mathcal{P}} - |\mathbf{I}_{\mathcal{P}}|^2\Big] f_{\exp}(t)\Big), \tag{6.22}$$

which represents an improvement of a factor of $n_{\mathcal{P}}$ over the matrix multiplication method. This factor may be very significant if we are using Eq. 6.4 or Eq. 6.5 to construct Markov descriptions of large processes. Additionally, when using the PGF method we may retain all transition probabilities as open variables, without running out of memory issues (if calculating results with a computer) or paper (if doing so by hand).

6.2.1 Double-heralding completion times

Here we will give an explicit, simple example of this method of finding p_t with generating functions, to highlight its utility. We will then use this example to show how the method may be used to find analytic expressions for the Fisher information. We will analyse the distribution of completion times of the double-heralding protocol for entanglement generation, as described in Section 3.1.4. This involves two rounds of photon transfer, the failure of either of which will cause the process to be restarted. We may therefore draw a simple Markov graph for this process as shown in Fig. 6.2,

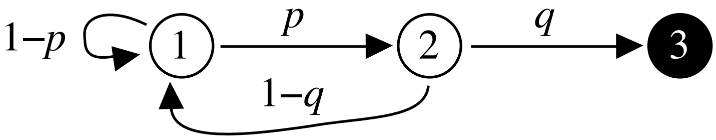


FIGURE 6.2: Markov graph for the two stages of double-heralding.

where p is the probability of passing the first round, and q the probability of passing the second, conditional on passing the first. If the transmissivity of the channel is η_T and the efficiency of Bob's detector is η_D, these may be written in terms of $\eta = \eta_T \eta_D$ as

$$\begin{aligned} p &= \frac{\eta}{2} + \frac{1-(1-\eta)^2}{4}, \\ q &= \frac{2\eta^2}{4\eta - \eta^2}. \end{aligned} \tag{6.23}$$

where the second term in p comes from the chance that both emitters emitted a photon.

From this graph, we may start to build up the PGF in an iterative, constructive way. We start by writing down the PGF for the first node according to the rules of Eq. 6.11:

$$\begin{aligned} f_{\mathcal{P}}^{(1)}(z) &= zpf_{\mathcal{P}}^{(2)}(z) + z(1-p)f_{\mathcal{P}}^{(1)}(z), \\ \to f_{\mathcal{P}}^{(1)}(z) &= \frac{zpf_{\mathcal{P}}^{(2)}(z)}{1-z(1-p)}. \end{aligned} \tag{6.24}$$

In the construction of $f_{\mathcal{P}}^{(1)}(z)$, we can see that we will only get non-trivial behaviour when the system includes closed loops, which will give the PGF a z-dependent denominator, and so at least one non-analytic singularity.[3] We then substitute in $f_{\mathcal{P}}^{(2)}(z) = zq + z(1-q)f_{\mathcal{P}}^{(1)}(z)$, and simplify to get the PGF for the whole system (implicitly starting from node 1), $f_{\mathcal{P}}(z) := f_{\mathcal{P}}^{(1)}(z)$, given by

$$f_{\mathcal{P}}(z) = \frac{z^2pq}{1-z(1-p)-z^2p(1-q)}. \tag{6.25}$$

Solving $f_{\mathcal{P}}(z)^{-1} = 0$ finds the poles:

$$z_{\pm} = \frac{p-1 \pm \sqrt{(1-p)^2 + 4p(1-q)}}{2p(1-q)}. \tag{6.26}$$

The residue of $f_{\mathcal{P}}(z)$ at a pole $z = z_{\pm}$ is given by $\lim_{z\to z_{\pm}} [(z-z_{\pm})f_{\mathcal{P}}(z)]$. Using Eq. 6.15, we therefore get

[3] Note that the method still technically works when there are no loops, but in that case we will have $t_0 = \infty$, so it reduces to calculating all derivatives "by hand."

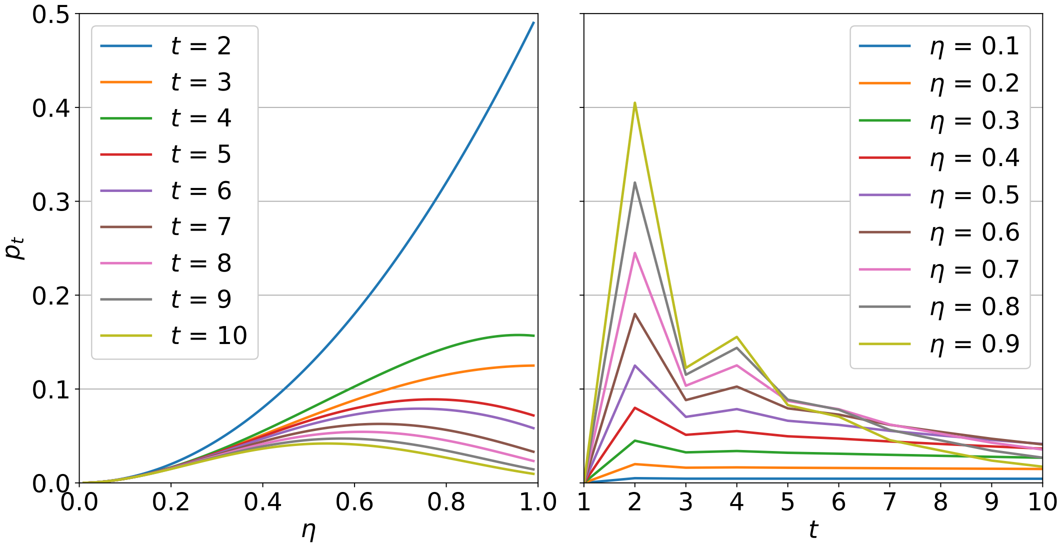


FIGURE 6.3: Completion time probabilities for double-heralding.

$$p_t = \frac{pq}{\sqrt{(1-p)^2 + 4p(1-q)}} \left(z_-^{1-t} - z_+^{1-t} \right). \tag{6.27}$$

This is illustrated in Fig. 6.3.

We have already covered the fact that the method of PGFs introduces a $n_{\mathcal{P}}$-fold speedup in calculating p_t. Here we discuss another benefit, that becomes most apparent when calculating derivatives. Suppose that we wish to calculate $\partial_\eta p_t$. To do this, it is necessary that we have an expression for p_t that retains η as an open variable, which is the case for Eq. 6.27. Note also that this equation is a general one that applies for all t. If, on the other hand, we wanted to calculate the $\partial_\eta p_t$ from the matrix multiplication method, then we would have much more difficulty. This is because there are two ways we can approach the calculation of $\partial_\eta p_t$ if we are using Markov matrices directly, which are performing the differentiation before or after the matrix exponentiation. i.e., assuming $I_{\mathcal{P}} = \{n_{\mathcal{P}}\}$ we can calculate $\partial_\eta \left[M_{\mathcal{P}}^t\right]_{1,n_{\mathcal{P}}}$ or $\left[\partial_\eta M_{\mathcal{P}}^t\right]_{1,n_{\mathcal{P}}}$. In the first of these two forms, we will have to calculate the derivative separately for each t. Not only that, but p_t $\left(= \left[M_{\mathcal{P}}^t\right]_{1,n_{\mathcal{P}}}\right)$ will include a super-exponential $2^{2^{t-1}-1}$ terms, with each term being a polynomial in η up to order $t-1$. The latter form is no easier to analyse, particularly since $\partial_\eta M_{\mathcal{P}}$ does not generally commute with $M_{\mathcal{P}}$. It is therefore clear that the PGF method provides a great advantage when we wish to calculate derivatives of distributions with respect to parameters.

One may very well ask now why we should wish to do such a thing. Consider the following situation. Alice and Bob are trying to distribute entangled pairs by double-heralding. Eve wants to attack this system by a photon-number splitting attack.[4] Suppose the Alice's source has a probability p_{multi} to emit a multi-photon state. Eve could take only the states that contain more than one photon, keep the excess photon to herself, and then send the remaining photon through a channel of transmissivity $\eta_T / p_{\mathrm{multi}}$, such that Bob will not notice any change in his detection rate. However, this requires that Eve can accurately estimate η_T. Assuming that η_D is a known parameter set by the device manufacturers and that Eve can access $\{T_i \mid i \in \mathcal{S}, |\mathcal{S}| = n_{\mathrm{est}}\}$, for some set of qubits $\mathcal{S}$ indexed by i, then she would want to know the extent to which her estimate of η is likely to differ from its true value. This equates to calculating the variance of $\hat{\eta}$, which is her estimator for η.

This variance, when estimated from these n_{est} independent measurements of T, is given by the Cramér-Rao bound:

$$\mathrm{Var}(\hat{\eta}) = \frac{1}{n_{\mathrm{est}} \mathcal{I}_\eta}, \tag{6.28}$$

where $\mathcal{I}_\eta$ is the Fisher information, given by

$$\mathcal{I}_\eta = \sum_t \left(\frac{\partial p_t}{\partial_\eta} \right)^2 \frac{1}{p_t}. \tag{6.29}$$

The ability to analytically calculate this derivative will therefore be useful to Eve in estimating whether she can launch such an attack without being detected.

We may also note here that from Eq. 6.27 we may find an analytic expression for the derivatives with respect to time, $\partial_t p_t$. We may see, for example, that $(\partial_t p_t)/p_t = -\left[\log_e(z_-)\, z_-^{1-t} - \log_e(z_+)\, z_+^{1-t}\right] / \left[z_-^{1-t} - z_+^{1-t}\right]$, which would be significantly harder to find, if at all possible, if we were finding p_t for each new t by a matrix multiplication.

6.3 Error distributions

In any process, there will be events that have some probability to cause an error. For example, if an event represents a state being stored on a quantum memory, then in each time-step there is some non-zero probability that the memory fails and the information stored on it is lost. When carrying out the process, we wish to know $p(k|t)$; the probability that we will pass such an edge k times, given an overall process completion time of t. This implicitly assumes that such a process is *heralded*. That is, we always know what stage of $\mathcal{P}$ we are at, and so can count the number of occurrences of an error-carrying event. A non-heralded process would be one where we have a description of $G_{\mathcal{P}}$, but we do not know how close we are to completion at any time, but instead are simply informed when the process completes. If each

[4] We mentioned in Section 3.1.5 that decoy states defend against such attacks. However, we can imagine situations where such a defence would fail. For example, such states may be implemented such that an additional side-channel exists that indicates which states are decoys and which are not, perhaps by some frequency or timing information.

occurrence of an error-carrying event has a probability ϵ to cause an error, then the overall probability that an error will have occurred is given by

$$\begin{aligned} p(\text{error heralded}) &= 1-(1-\epsilon)^k, \\ p(\text{error non-heralded}) &= \sum_{k=0}^{\infty} p(k|t) \left[1-(1-\epsilon)^k\right]. \end{aligned} \tag{6.30}$$

In order to include this in our analysis, we must first identify which events (edges in $G_{\mathcal{P}}$) may cause an error. Then, for each event in question between edges j and i, we multiply $[M_{\mathcal{P}}]_{i,j}$ by an open complex variable, w, which we will call the counting variable.

Now note that the value for p_t, calculated either by Eq. 6.1 or Eq. 6.15, may be seen as a sum of the probabilities of the different sequences of events by which the process may be completed in time t. When one term in $M_{\mathcal{P}}$ is an open variable, p_t will be expressed as a finite polynomial in w, which we will denote $p_t(w)$. For such a sequence of events that includes k passes of an error-carrying edge and occurs with probability $p(k|t)$, $p_t(w)$ will include a term equal to $p(k|t)w^k$. The full expression for $p_t(w)$ will then be of the form

$$p_t(w) = p_t(1) \sum_{k=0}^{\mathcal{O}(p_t(w))} p(k|t)\, w^k. \tag{6.31}$$

where $\mathcal{O}(p_t(w))$ is the order of $p_t(w)$. Thus by finding the coefficients of this polynomial, we can find the error distributions. This polynomial is finite, with all terms involving w to a non-negative power. Therefore it has no poles, so we cannot use the methods of Section 6.2. Instead we can extract the coefficients by way of a (fast) Fourier transform, which, unlike the complex analysis method, can be done numerically. To do this, we first should identify some number $N_{\mathcal{P}}(t)$ such that $\mathcal{O}(p_t(w)) \leq N_{\mathcal{P}}(t) \leq 2\mathcal{O}(p_t(w))$, where the latter inequality is to avoid aliasing effects.[5] We then evaluate $p_t(w)$ at $N_{\mathcal{P}}(t)$ equally spaced complex points, given by $\{p_t(e^{i2\pi k/N_{\mathcal{P}}(t)})|k = 1, \ldots, N_{\mathcal{P}}(t)\}$. The discrete Fourier transform of these evaluated points reveals the first $N_{\mathcal{P}}(t)$ coefficients of $p_t(w)$ (where all greater coefficients are 0). Applications of this Fourier method for extracting coefficients to more general analytic functions are described in [For81]. This construction may also be used to account for different kinds of errors, by multiplying matrix elements by different complex variables, $w_1, w_2, w_3, \cdots$, and performing a multi-dimensional Fourier transform on $p_t(w_1, w_2, w_3, \cdots)$ to determine $p(k_1, k_2, k_3, \cdots |t)$.

From this we can also read off the *average* error rate for a process completing by time t. i.e. the non-heralded error. Suppose we have only a single type of error. From Eqs. 6.30 and 6.31, we can say that

[5] It may seem like we can always choose $N_{\mathcal{P}}(t) = t$. However, if $M_{\mathcal{P}}$ is constructed from elementary process matrices by Eq. 6.4 or Eq. 6.5 we may need to choose a larger value for $N_{\mathcal{P}}(t)$.

$$p(\text{error non-heralded}) = \frac{1 - p_t(w = 1 - \epsilon)}{p_t(1)} \tag{6.32}$$

The case for multiple types of error follows as a simple extension of this.

6.4 Innsbruck protocol analysis

In this section we will be considering the repeater protocol of the Innsbruck group [BDCZ98]. In the standard implementation of this protocol, there are Q_0 pairs of quantum memories between each adjacent pair of repeater stations. These are all connected in parallel, and then distilled to make $Q_1 \leq Q_2/2$ pairs. By entanglement swapping, these are then connected with adjacent pairs to form entanglement over twice the length, and distilled again to form $Q_2 \leq Q_0/4$ pairs in parallel between each section, and so on. Previous analyses of this protocol have either assumed that the entanglement connection can be done almost deterministically, or considered that entanglements establish after some average time. So if each attempt to establish entanglement between two stations has a probability to succeed of p_c, then a simplified approach to understanding the system and estimating the key rate would be to assume that all pairs establish entanglement after $1/p_c$ attempts.

Considerable progress has been made in understanding and building upon the Innsbruck protocol, since it is one of the most promising routes to constructing long-distance quantum communication. Much of this work has focused on aspects such as the relation between the key rate and experimental imperfections [ABB+13], the specifics of how to implement the system with atomic ensembles [DLCZ01], understanding and improving the robustness against channel noise [SSDRG11, VK17] or side-channel attacks [LWW+10, LLK07, NFSM+09, VK18]. However, one important aspect is often overlooked, namely the statistical factor of waiting times arising from probabilistic completion times of different elements. We now show that this has severe implications for the performance of the protocol.

In our analysis, we let the Q_0 parallel pairs of memories between each pair of repeater stations be divided into bunches of q_0 pairs. When all pairs within such a bunch have completed, then they are distilled to $q_1 \leq q_0/2$ pairs. This is an inequality, since distillation (as described more fully in Appendix C) is a probabilistic process. This is shown in Fig. 3.4. Here, there is a trade-off inherent in the size of q_0. When q_0 is small, the bunch will complete quickly on average. This means that the first entangled pair to complete will not have to wait long before the last one completes, and so is less likely to accrue memory errors. However, one then has fewer options for distilling a high fidelity state. Given a large set of states, we can instead find a better optimal strategy for combining states under a distillation protocol to result in a higher final secret key rate, at the expense of longer waiting times.

6.4.1 Constructing the matrix

Here we consider at first a repeater consisting of two sections, separated by a distance L. Alice tried to establish entanglement between herself and Richard (a repeater station), and Richard between himself and Bob. The Markov graph for the establishment of a single Bell pair is shown in Fig. 6.1. Let its Markov matrix be M_{Bell}. We shall consider one time-step in this process to be $2L/c$, where the factor of 2 arises since the receiving party needs to send a classical signal back to the sending party to confirm whether the previous photon was received or not.

We now use Eq. 6.4 to construct the matrix for q_0 pairs connecting in parallel between a pair of repeater stations. We want to include a complex counting variable, w_0 that counts how many time-steps a given quantum memory has to wait before the others finish. However, we should note that we include this only on a single factor of the matrix for the section, M_{sect}, to avoid multi-counting errors. The counter w_0 therefore counts how many errors accumulate on a particular entanglement link. By symmetry, we can say that this error distribution is equal across all such entanglement links. Therefore,

$$\begin{aligned} M_{\text{sect}} = &\left[M_{\text{Bell}} + w_0 \operatorname{diag}(\mathbf{I}_{\text{Bell}})\right]\otimes \\ &\left[M_{\text{Bell}} + \operatorname{diag}(\mathbf{I}_{\text{Bell}})\right]^{\otimes q_0 - 1} \\ &- w_0 \operatorname{diag}(\mathbf{I}_{\text{Bell}}^{\otimes q_0}), \end{aligned} \tag{6.33}$$

where

$$M_{\text{Bell}} = \begin{bmatrix} 1 - p_{\text{c}} & 0 \\ p_{\text{c}} & 0 \end{bmatrix}, \quad \mathbf{I}_{\text{Bell}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{6.34}$$

where p_{c} is the probability of entanglement being established in any particular attempt. In Appendix D we describe how this may be reduced in dimension by symmetry arguments.

Once all pairs between two stations have established their entanglement, we want to perform a distillation on these. These are matched up into $\lfloor q_0/2 \rfloor$ pairs, which are then distilled using the DEJMPS protocol [DEJ+96]. A DEJMPS distillation between two noisy Bell pairs has some non-unity chance of success, which depends on the fidelities of the states involved. However if the success probabilities were fidelity-dependent, that would mean including terms in the matrix which depend on the time taken for the process to reach that event, a modification which would move us outside the realm of Markovian dynamics. Therefore we will choose some minimum distillation success probability, λ, corresponding to the success probability two states of fidelity F_{min} being distilled with each other. We will later exclude any runs of the process that would have used states of fidelity less than F_{min}, as explained in Appendix C. In this way, any choice of λ will give us a lower bound on the secret key rate, and we may freely maximise over choices of λ. We therefore add a row and column to M_{sect} to form M_{dist}. The new column has an element representing distillation success, with a probability of $1-(1-\lambda)^{\lfloor q_0/2 \rfloor}$. The "failure" event (of all distillations failing) resets the process of creating entanglement on that section. Therefore, if M_{sect} in Eq. 6.33 takes the form

$$M_{\text{sect}} = \begin{bmatrix} M'_{\text{sect}} & \mathbf{0} \\ \mathbf{a}^T & 0 \end{bmatrix}, \tag{6.35}$$

for some matrix M'_{sect} and vector $\mathbf{a}$, then M_{dist} is given by

$$M_{\text{dist}} = \begin{bmatrix} M'_{\text{sect}} & \mathbf{b} & \mathbf{0} \\ \mathbf{a}^T & 0 & 0 \\ \mathbf{0}^T & s & 0 \end{bmatrix}, \tag{6.36}$$

where the vector $\mathbf{b} = \left[(1-\lambda)^{\lfloor q_0/2 \rfloor}, 0, 0, \cdots, 0\right]^T$, and $s = 1-(1-\lambda)^{\lfloor q_0/2 \rfloor}$.

Finally, we construct the matrix for the entire system, M_{full}, in a similar way to Eq. 6.33 by considering two copies M_{dist}. We again include a complex counting variable to account for all memories on one section needing to wait until the other side has been connected and distilled. For this we make sure to use a different counting variable, w_1, so we can keep track of the distribution of errors that occur before and after the first round of distillation.

6.4.2 Analysis of results

We can now analyse $\{p(k_0, k_1|t)\}$, where k_0 and k_1 are the number of passes of edges weighted by w_0 and w_1 respectively. By doing this for a fixed q_0, we can find a distillation strategy that gives the maximum possible achievable secret key rate for a given completion time, t, averaged over the error distribution (explained in detail in Appendix C), which we shall call $K(t|q_0, p_c, \epsilon_W)$. Since the key rate of a protocol goes inversely with the time taken to establish a raw bit, and linearly with the number of parallel "bunches" of states that are used, Q_0/q_0, we will use the normalised average key rate as a function of q_0 as our main figure of merit for the system analysis:

$$K(q_0, p, \epsilon_W) = \frac{1}{q_0} \sum_{t=1}^{\infty} \frac{p_t}{t} K(t|q_0, p_c, \epsilon_W) \tag{6.37}$$

For the error probability per time period, ϵ_W, we assume the probability for a quantum memory to not undergo an error decays exponentially with time as $\epsilon_W = 1 - \exp(4L/c\tau)$, where the extra factor of 2 is due to the fact that each entangled pair involves 2 memories. Here, τ is the memory lifetime. In Fig. 6.4 we show the a few examples of calculated normalised key rates for different values of q_0. In order to compare the insight gained from this method to the estimations that might result from a less nuanced analysis, we have also shown the *simplified secret key rate*. Here, we consider only the average connection time of an entanglement link. That is to say, we assume that all links wait for a time $1/p$, and then connect

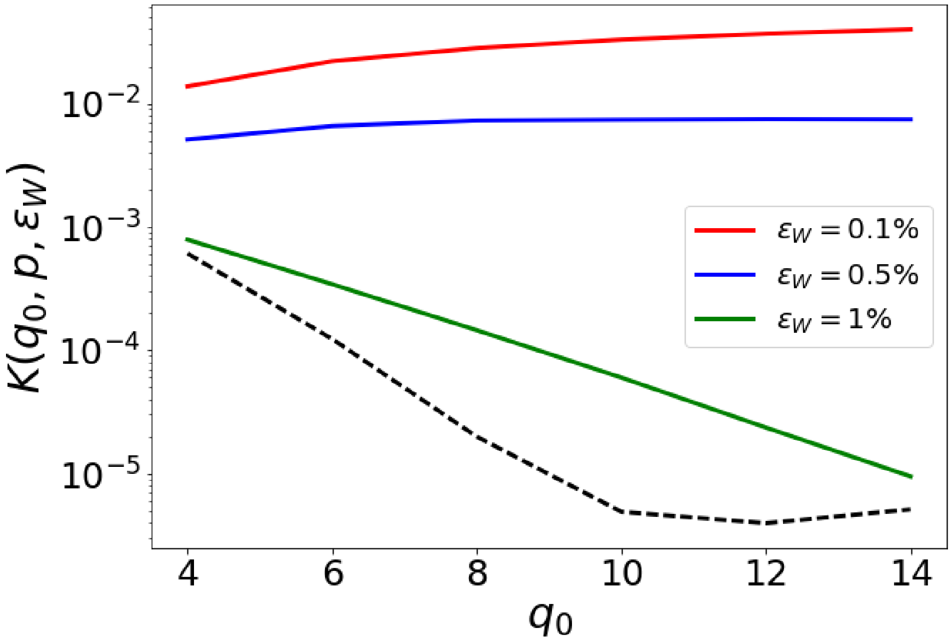


FIGURE 6.4: The normalised secret key rate, in the case where $p = 0.1$. Plots are in the same order as in the legend. Solid lines show the key rate including statistical effects (K) and the dashed line shows the simplified key rate (K_{simp}), which is identical for all ϵ_W.

deterministically. For a fair comparison, we have retained the same maximization over distillation strategies that is outlined in Appendix C. The simplified secret key rate may therefore be written as

$$K_{\text{simp}}(q_0, p_c, \epsilon_W) = \frac{1}{q_0} \sum_{t=1}^{\infty} \frac{p_t}{t + 1/p_c} K(t|q_0, 1, \epsilon_W) \tag{6.38}$$

When we consider Fig. 6.4, we can see that an estimation of the secret key rate that considers only the average completion time severely underestimates the performance of the protocol. This is particularly striking when we note that no errors accumulate in the simplified analysis, due to the fact that no elements are left waiting while others complete. In particular, the difference between the statistical and simple key rates in the $\epsilon_W = 0.1\%$ case reaches 3.8 orders of magnitude, which could mean the difference between communicating in kilobits and megabits per second.

We may also note that, for low values of ϵ_W, the key rate rises with increased bunch size. This means that increasing the pool of states available to be distilled has a greater-than-linear effect on the key rate, highlighting the power and importance of distillation to quantum technologies.

6.4.3 Analytic key rate bounds

The techniques presented above allow for a thorough investigation of the contribution of statistical factors arising from non-deterministic protocol elements towards the secret key rate of a general quantum communication protocol. While this has been presented in-depth for a two-section repeater, practical systems will often demand the application of a series of many repeaters. The current limit for repeater-unassisted quantum communication is on the order of a hundred kilometers. If we therefore wish to securely communicate on an intercontinental scale, we require a method of analysis that can scale up to dozens of repeater sections. This presents a limitation in our protocol: while the dependence of $n_{\mathcal{P}}$ on q_0 can be made linear (Appendix D), the dependence on the number of sections, N_S, remains exponential.

In the original analysis of the Innsbruck protocol, the fidelity of the final shared state was not considered to be fundamentally dependent on the number of sections. This is because the l^{th} level of the protocol, which consists of taking entangled pairs over some distance $2^l L$, distilling them, and connecting with adjacent sections to form pairs over a distance $2^{l+1}L$, would produce pairs of a fidelity that did not depend on l.

In addition to waiting times increasing with q_0, there is also the issue that the classical communication time grows with the distance over which pairs are entangled. However, if statistical factors are ignored then this could be dealt with using a "blind" protocol, where distillation and attempts are assumed to have succeeded at every stage, and communication after the termination of the protocol allows for a post-selection on the attempts that succeeded. This allows for the final fidelity to be kept above the minimum level required for secure communication at the expense of a hit to the raw key bit generation rate.

Once we include the statistical waiting times in our analysis, it is no longer true that a non-zero secret key rate can be guaranteed for all length scales. In this section we examine the behaviour of a repeater network with a minimum requirement of physical qubits, which is $q_0 = 2^{N_S}$. Instead of finding best key rates, we look here for the parameters for which the secret key rate is lower bounded above zero. By finding the threshold parameters that need to be reached for the protocol to operate, we can identify concrete values for component designers to aim for, and give benchmarks by which we can compare performances.

We will consider that after every quantum operation, a Werner twirling procedure [BBP+96] is applied to all states. This involves unitarily mapping the state to the singlet state, $|\Psi^-\rangle$, then applying a randomly chosen local unitary Pauli operation identically each part of the entangled pair. This maps all states to Werner states of the same fidelity (see Appendix C). By doing this, we can simply consider the effect of the repeater network as a recursive function on a single real variable - that of the average fidelity. Note that this operation is not actually carried out, it is simply used to repeatedly map states to the analytically simple Werner states. This may be done since applying local operations cannot increase the strength of entanglement by any measure, and so cannot increase the secret key rate. Note that this equivalent to the argument used in Section 5.3.3 where we stated that the Werner state was the highest entropy state in the Bell basis of fixed fidelity.

The analysis then proceeds as follows. Level 0 of the protocol consists of all pairs within one section connecting at initial fidelity F_{init}. Instead of considering the full probability distribution of waiting times, we consider that all pairs wait for a number of time-steps equal to the estimated time for the last pair to connect, k_L, which upper bounds the waiting time for each pair. This time is equal to the expectation value for the largest order statistic from a sample of q_0 chosen from the distribution with cumulative distribution function $1-(1-p_c)^t$. By choosing $p_c \ll 1$, such that we may allow the distributions to be approximated by continuous functions, we may use results from [AG+79] to bound this by

$$k_L \leq \left(\frac{q-1}{\sqrt{2q-1}}+1\right)\frac{1}{|\log(1-p_c)|}+1. \tag{6.39}$$

These states are then distilled to produce states of fidelity

$$F_0 = J[D(F_{\text{init}}, \epsilon_W, k_L)], \tag{6.40}$$

where the functions describing the effect of the decay of quantum memories over time k_L on the average fidelity and DEJMPS distillation are given respectively by

$$\begin{aligned} D(F,\epsilon_W,k_L) &= (1-\epsilon_W)^{k_L}F+\frac{1-(1-\epsilon_W)^{k_L}}{4}, \\ J(F) &= \frac{10F^2-2F+1}{8F^2-4F+5}, \end{aligned} \tag{6.41}$$

respectively. The l^{th} level of the protocol consists of the following when $l \geq 1$. Within each pair of two sections, one section will complete first, and wait for a time no longer than $k_{A,l}$ for the latter to complete. We show in Appendix E that this is bounded by

$$k_{A,l} \leq 2^l \left[\frac{H(2^{N_S-l+1})}{|\log(1-p_c)|}+1\right], \tag{6.42}$$

where $H(n) = \sum_{m=1}^{n} 1/m$ is the n^{th} harmonic number. The average fidelity after level l can then be defined recursively as

$$F_l = J(C(F_{l-1}, \tilde{F}_{l-1}, \epsilon_L)), \tag{6.43}$$

where C gives the average fidelity after connecting two adjacent sections and twirling, where we have allowed here for local gate errors. This is given by

$$C(F_a, F_b, \epsilon_L) = D\left(\frac{1}{3}(1-F_a)(1-F_b)+F_aF_b, \epsilon_L, 1\right), \tag{6.44}$$

where ϵ_L is the probability that an error occurs when performing the local operations involved in entanglement swapping, equal to $1-x_{\text{ga}}$, and $\tilde{F} = D(F_{l-1}, \epsilon_W, k_{A,l-1})$.

In Fig. 6.5 we show the minimum quantum memory lifetimes required for a non-zero key rate as a function of p_c. From this it can be seen that the probability for an entanglement attempt to succeed is the biggest factor in affecting the ability to securely communicate. For comparison we also include the requirements for the case that does not include statistical effects, where $k_{A,l} = 2^l$ for all l. It can be seen that the minimum memory requirements are slightly higher in the case where statistical effects are included, but this effect decreases with the number of sections over which we connect. For an even comparison, we have not used blind distillation in the non-statistical case. We see that there is a constant-factor increase in the required lifetime of the memories. In some cases this reaches as high as a factor 2 increase in the required lifetime of the quantum memories. The resultant bounds are just reachable by the lifetimes of atomic ensembles, which can have lifetimes up to 40ms [KMJ$^+$11]. However, all bounds are well within the lifetimes of the nuclear spin states of NV centers [LMY$^+$05]. This implies that the main challenge towards implementation of DLCZ-type protocols [DLCZ01] is the construction of optical elements with high transmission and detection efficiencies, whereas NV-center-based protocols [NTD$^+$14, NTD$^+$16, VK17] may be more suitable when these efficiencies are low.

6.5 Summary

Many of the practical quantum technologies that are being proposed are inherently probabilistic in nature, which leads to uncertain completion times and error distributions. We have developed techniques that allow for thorough characterizations of such statistical distributions in both the computational and analytic directions. In terms of computational techniques, we have used Markov chains to analyse quantum entanglement generation. We have shown how to form composite systems from smaller elements in a constructive manner. We have then shown how to use the Markov matrices for such composite systems to calculate the probability distribution over the number of errors that occur in the running of a general protocol. This allows for a complete characterization of the fidelities of the states that are produced by a quantum protocol. As an example, we have analysed the Innsbruck quantum repeater protocol with a memory-error model. A thorough understanding of the set of resultant errors has been shown to lead to a tighter bound on the secret key rate than an analysis based only on an averaged approach. In some cases this resulted in tightening the bounds on key rates by over 3 orders of magnitude – a clear indication that a consideration of statistical effects does not simply provide a minor correction to performance, but instead is fundamental to understanding the quantitative behaviour of a system.

In terms of analytic techniques, we have shown how elements of the eigenvectors of a transformed form of the Markov matrix correspond to the probability generating function of the process. This has been solved for an arbitrary term in the probability distribution over completion times by using results from complex residue analysis. This was done in a way such that the number of computational operations required

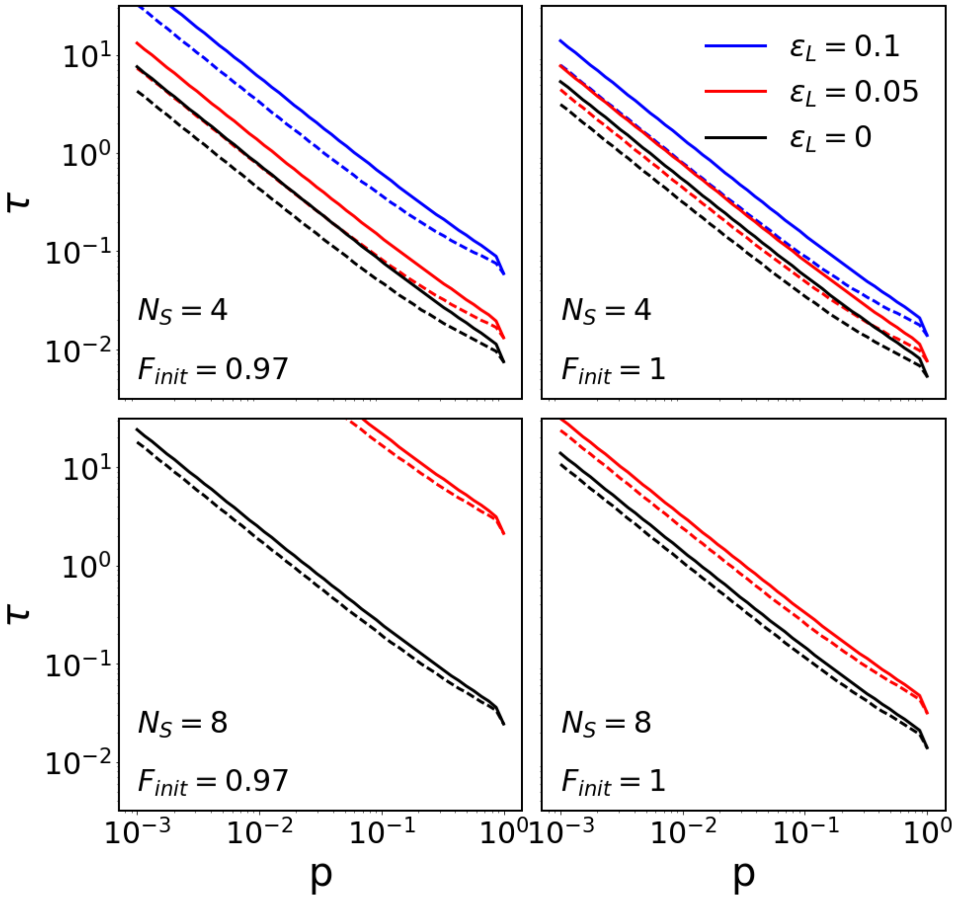


FIGURE 6.5: Minimum quantum memory lifetime, $\tau = -4L/\log(1-\epsilon_W)c$, required for a minimum-resources quantum repeater to securely communicate over a network of 8 sections of length $L = 25$km. F_{init} is the initial fidelity of entangled pairs before distillation or connection, and N_S is the number of sections. Solid lines indicate the cases that include statistical factors, and dashed lines do not. Both solid and dashed lines are ordered the same in the plots and the legend. Blind distillation is not used in either case.

scales only linearly with the number of states in the Markov process, compared to the quadratic scaling of a more direct approach. This has been highlighted for deriving an analytic expression for probability for double-heralding to complete in a given time, and is shown to be instrumental in calculating the Fisher information of this and similar processes. Finally we have shown how the theory of order statistics can put bounds on the statistical effects on the secret key rate, and used this to bound the minimum quantum memory lifetimes needed to run the Innsbruck repeater protocol.

Chapter 7

Conclusion and future developments

It is here that we look back at the main results of this thesis, and what has been achieved over the course of this doctoral study, as well as looking forward at what the future path of development for this field might be.

7.1 What has been done

In this thesis we have given an overview of the current field of quantum key distribution, and presented protocols, analyses, methods and results which strengthen the security of such systems and allow for a more thorough understanding of their statistical dynamics.

In Chapter 4 we examined the susceptibility of the BB84 QKD protocol to the Trojan-Horse Attack, where an eavesdropper, Eve sends an optical state into a system of Alice's that encodes her qubits, and measures the reflected state to try to learn the value of the qubit. We initially used a framework where Eve was permitted to use any Gaussian state. Additionally, we allowed Alice to include an amount of thermal noise with her outgoing state, which could muddy Eve's estimation of the qubit. We noted that this inclusion of thermal noise required a different method of combining states than is usually used due to the fact that the strength of the combination should depend only on the average photon number in the noise, and that we could take advantage of the structure of thermal noise and introduce it by applying a two-mode squeezer between the signal and the vacuum, and tracing out a mode.

Within this framework, we showed that Eve does not gain in estimation ability by using a squeezed state. This is because any practical system will only permit a finite number of photons to enter. Squeezing a state increases the average photon number, so one must economically distribute one's photons between squeezing and displacement, and Eve was shown to always benefit by maximising the displacement of her state. From this, we showed that the distinguishability between different qubit states, a parameter that Alice and Bob wish to minimise, was bounded by

$$\Delta \leq \frac{1}{2}\left[1 - \exp\left(-\frac{\mu_D}{1+2\mu_T}\right)\right], \tag{7.1}$$

where μ_D is the average number of photons that Eve receives back from her inputted state, and μ_T is the average photon number of the thermal noise. We calculated the optimal thermal noise strength for Alice to use, while noting that even in situations where using such noise is impractical, the results still stand to prove the counter-intuitive proposition that one should not necessarily try to minimise the noise in one's system.

We then moved to an analysis that generalised Eve's attack to be an arbitrary separable state. This was motivated by the fact that the high attenuation used by Alice would destroy almost all entanglement that Eve used, and was backed-up by the previously-found fact that her optimal Gaussian was separable. By considering the effect of Alice's attenuator as an expansion of Krauss operators, we were able to show that, in the separable case,

$$\Delta \leq \frac{1 - e^{-\mu}\sqrt{1 - 3\mu(2-\mu)/4}}{2}, \tag{7.2}$$

which presented a tighter bound than earlier work. We concluded this chapter with a novel idea for how one may artificially increase the strength of an attenuator with an active shutter system.

Next, we turned our attention from the problem of secure communication between two nearby parties, to that of linking together many QKD channels to communicate over large distances. In Chapter 5 we introduced our design for a novel quantum repeater architecture. In doing so, our hope was to highlight the importance of choosing an efficient method for generating the primary entanglement between repeater stations. Our protocol recognised the utility of using NV centres for the quantum memories, as had been suggested in earlier works, but moved beyond this by proposing to take advantage of the level structure of the centres to perform double-heralded entanglement generation. This has allowed for a method of producing an initial entanglement with a fidelity that depends neither on the efficiency of detectors, nor on the transmissivity of the optical channel. Additionally, the use of NV centres allows for quantum information to be stored on the nuclear spins whilst entanglement is generated on the electronic spins. This provides a natural architecture from which to apply distillation of entanglement, as well as enabling fault-tolerant entanglement-swapping through brokering.

Our analysis included a thorough investigation of dark counts, which proved to have a negligible effect on the secret key rate. We also introduced the notion of the contribution of statistical effects to the secret key rate. This is idea that different sections of a repeater will complete at different times, thus leading to decoherence of quantum memories while the remaining sections catch up. We showed that, even with such statistical effects included, our protocol performed very favourably, allowing for non-zero key rates over thousands of kilometres, and giving high secret key rates, even when normalised by the number of qubits used.

In Chapter 6, we investigated the statistical effects of probabilistic entanglement generation much more deeply. Here we began with developing powerful methods and frameworks to analyse finite Markov processes. The first of these was a method for developing and analysing the probability generating function of a process in order to find a distribution over its completion times. Here, there were two main results. First, we showed that for a process $\mathcal{P}$ described by a Markov matrix $M_{\mathcal{P}}$, related to a classical stochastic matrix, we can derive its generating functions as being as proportional to elements of the vector

$$\boldsymbol{f_{\mathcal{P}}}(z) = \text{Null}\left[zM_{\mathcal{P}}^{T} + \text{diag}(\mathbf{I}_{\mathcal{P}}) - \mathbb{1}\right], \tag{7.3}$$

where $\mathbf{I}_{\mathcal{P}}$ is the vector of states on which $\mathcal{P}$ terminates. Next, we showed that this may be solved, to give the probability that $\mathcal{P}$ completes at time t as

$$p_t = -\sum_i \text{Res}\left[\frac{f_{\mathcal{P}}(z)}{z^{t+1}}, z_i \in \mathbb{P}(f_{\mathcal{P}})\right]. \tag{7.4}$$

This was used to derive an analytic description of the process of double-heralding, and the statistics of its associated completion times.

When applying this to the problem of calculating an accurate estimate for the secret key rate of a repeater network, we noted that is was not sufficient to calculate how long the entire process took to complete, one must also keep track of the accumulated errors. We showed how one may track the number of times that an edge was traversed in a Markov graph, by multiplying the corresponding matrix element by a complex variable. Evaluating the completion time at values of this variable evenly distributed about the unit circle and taking a discrete Fourier transform reveals the probability distribution for the number of times that edge was traversed. This enabled us to perform a thorough analysis of the Innsbruck repeater network on two sections. In order to allow this to be practically extended to multiple sections, we showed how one may use order statistics to get bounds on the statistical effects on the key rate.

Throughout this work, it may be said that we have made multiple studies of a fundamental trade-off, expressed in different forms. This is the notion that, when operating a QKD protocol, one may make an exchange between raw key rate and secrecy. We first encountered this idea in fundamental forms in Chapter 3. Here we saw how the classical cryptographical concept of privacy amplification, by which one may shorten a key to reduce the mutual information between it and an adversary, was reflected in the quantum operation of entanglement distillation. This involves a large number of weakly entangled states being consumed in the production of a small number of more highly entangled states. When these states are used to produce a key by the E91 protocol, then the parallel between this process and privacy amplification is elevated to an equality. In Chapter 4 we saw how privacy could be amplified by the introduction of thermal noise, whilst at the same time, reducing the proportion of key bit generation attempts on which Bob measures a single photon, thus lowering the raw key rate. Similarly, we saw in Chapter 5 how the choice of buffer value, δ, would affect both the raw key rate and the fidelity — we can wait longer to guarantee a good entangled connection, but at the expense of a longer delay between bit generations. In Chapter 6 we saw how our choice of bunching size, q_0, indirectly induced such a trade-off, by influencing the states available to distil.

When we consider these balances, we may find it reasonable to reach a more general conclusion that any theoretical or experimental development must be considered in its proper context (in terms of security, computational power, etc.) to fully understand its merits and demerits.

7.2 What is to be done

Just as the work presented here has been built from the excellent work done by a great many scientists over many decades, we hope that our work will serve as a small stepping-stone to further developments in the field of quantum communication. Here we discuss some potential ideas for further work.

Our protocol for a repeater network has shown that double-heralding on NV centre-based qubits is an excellent method for generating entanglement. Whilst the protocol presented is one way of using this method, it is not unique. The use of double-heralding could provide highly entangled states for other architectures of repeaters, such as ones that enable communication between multiple parties, or take advantage of fault-tolerant encodings. If this method can be combined with a quantum state configuration that is resistant against photon loss, then this would have the potential to communicate over large distances, even *before* the use of repeater stations. The analysis of such architectures may also benefit from using the techniques developed in Chapter 6, since they are rather general in nature.

When repeaters *are* used, it has become clear from Chapters 5 and 6 that the probabilistic nature of the initial entanglement generation is a critical factor in determining the ultimate secret key rate, due to the resulting memory decoherences that result from a statistical distribution of completion times for the sections. Whilst double-heralding produces high-fidelity pairs, it does so on at most 50% of attempts. This is due to the fact that, from the initial emitter state of $|+,+\rangle$, the desired entangled output is $|\Psi^+\rangle$, and $|\langle +,+|\Psi^+\rangle|^2 = 1/2$. In order for us to maximise the secret key rate generated by a repeater network, we want to make the probability of creating an entangled connection in any given attempt as high as possible.

Initial work on such a protocol has been investigated. Here, we might consider that we again begin with two qubits initialised in the state

$$|+,+\rangle = \frac{1}{2}\left[|0,0\rangle + |0,1\rangle + |1,0\rangle + |1,1\rangle\right], \tag{7.5}$$

which are stimulated to cause photon emission as in double-heralding. The outputs of these may be passed through beam-splitters which mix them with a highly-squeezed ancilla state, as shown in Fig. 7.1. In the situation where $\eta = 1$ and there

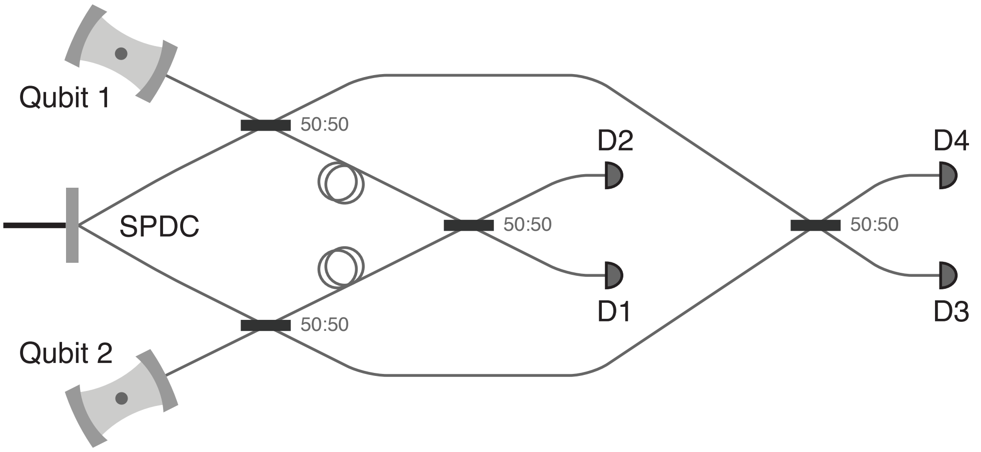


FIGURE 7.1: The set-up for ancilla-assisted double-heralding (drawn by Pieter Kok).

is no photon loss, then a final detection of an even number of photons would project the state onto the $|0,0\rangle, |1,1\rangle$ subspace. As the squeezing parameter ζ is increased, the distribution on the number of photons in the ancilla becomes flat, and the projected state will approach $|\Phi^{\pm}\rangle$. Similarly, a detection of an odd number of photons will project onto a state that approximates $|\Psi^{\pm}\rangle$. However, as we discussed in Section 7.1, this increase in success probability cannot come for free. When $\eta < 1$, as in the case in any practical set-up, then as we increase the ζ, the probability that we will lose a photon also increases. This will decrease the fidelity of our final state. It is therefore important that we carefully calculate the optimal value of ζ if such a protocol is to result in an increase in overall secret key rate. An alternative approach would be to replace the squeezed state with another Bell state, say $|\Phi^{+}\rangle$. This would result in higher-fidelity pairs, but bound the success probability at 75%. In any case, we believe that the statistical analysis of QKD systems is a greatly important field, and one with a great deal left to be discovered.

Appendix A

Proof of the non-equivalence of squeezer and beam-splitter approaches to combined state generation

In Section 4.1, we asserted that the proper way to model the addition of thermal noise into the channel was by modelling the system as using a two-mode squeezer of squeezing parameter ζ, squeezing our signal with a vacuum mode, and then tracing out one mode (situation A). In particular, we claim that this is *not* equivalent to a situation where we mix our signal with a thermal state of average photon number μ_T on a beam-splitter of *complex parameter* $\xi = \phi e^{i\theta}$, and trace out one mode (situation B). When the parameter of the beam-splitter is acts on two modes $\hat{a}_1, \hat{a}_2$ by the unitary

$$U_{\text{BS}} = e^{\xi \hat{a}_1^\dagger \hat{a}_2 - \xi^\star \hat{a}_1 \hat{a}_2^\dagger}, \tag{A.1}$$

and maps

$$\begin{aligned} \hat{a}_1 &\mapsto \cos(\phi)\hat{a}_1 + e^{i\theta}\sin(\phi)\hat{a}_2, \\ \hat{a}_2 &\mapsto \cos(\phi)\hat{a}_2 - e^{-i\theta}\sin(\phi)\hat{a}_1. \end{aligned} \tag{A.2}$$

When $\theta = 0$, this is a beam-splitter of transmissivity $\eta = \cos^2(\phi)$.

Here we will examine the effect of these two situations with the use of symplectic matrices on covariance matrices. Given an opical transformation, T, we can identify a symplectic matrix corresponding to this transformation, S_T. The covariance matrix of the state, C, is then transformed as

$$C \mapsto C' = S_T \, C \, S_T^\dagger. \tag{A.3}$$

Suppose we take as our input signal to both situation A and B to be a single-mode thermal state, with average photon number N. The covariance matrices for situations A and B are, respectively,

$$\begin{aligned} C_A &= \left(\frac{n+1}{2}\,\mathbb{1}_2\right) \oplus \left(\frac{1}{2}\,\mathbb{1}_2\right), \\ C_B &= \left(\frac{n+1}{2}\,\mathbb{1}_2\right) \oplus \left(\frac{\mu_T+1}{2}\,\mathbb{1}_2\right), \end{aligned} \tag{A.4}$$

where $\oplus$ represents a direct sum. The symplectic matrices for a two-mode squeezer (where we assume ζ is real) and a beam-splitter are, respectively,

$$\begin{aligned} S_{\text{2MS}} &= \begin{bmatrix} \cosh(\zeta)\,\mathbb{1}_2 & \sinh(\zeta)\,\mathbb{1}_2 \\ \sinh(\zeta)\,\mathbb{1}_2 & \cosh(\zeta)\,\mathbb{1}_2 \end{bmatrix}, \\ S_{\text{BS}} &= \begin{bmatrix} \cos(\phi)\,\mathbb{1}_2 & \sinh(\phi)\,\mathcal{R}_\theta \\ \sinh(\phi)\,\mathcal{R}_\theta & \cos(\phi)\,\mathbb{1}_2 \end{bmatrix}, \end{aligned} \tag{A.5}$$

where we take the clockwise 2D rotation matrix...

$$\mathcal{R}_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \tag{A.6}$$

Applying the matrices in Eq. A.5 to those in Eq. A.4 by Eq. A.3, we arrive at:

$$\begin{aligned} C_A' &= \left(\frac{n+1}{2}\cosh^2(\zeta)\,\mathbb{1}_2\right) \oplus \left(\frac{1}{2}\cosh^2(\zeta)\,\mathbb{1}_2\right), \\ C_B' &= \left(\frac{n+1}{2}\cos^2(\phi)\,\mathbb{1}_2\right) \oplus \left(\frac{\mu_T+1}{2}\cos^2(\phi)\,\mathbb{1}_2\right), \end{aligned} \tag{A.7}$$

From this we can see that no choice of beam-splitter parameter on a thermal state can reproduce the behaviour of a two-mode squeezer on a vacuum, and situation B is not reducible to situation A.

Appendix B

Proof of Eq. 4.27

We want to show that $\mathrm{Tr}\left[\hat{E}_{\checkmark}\mathcal{E}(\rho)\right] \geq e^{-\mu}$. First note that since $\hat{E}_{\checkmark} = |0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|$, we can assert that

$$\mathrm{Tr}\left[\hat{E}_{\checkmark}\mathcal{E}(\rho)\right] \geq \mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}(\rho)\right]. \tag{B.1}$$

Since $\mathcal{E}$ does not map off-diagonal elements to diagonals and $|0\rangle\langle 0|$ is diagonal, we can consider only the effect of $\mathcal{E}$ on diagonal elements, and say that

$$\begin{aligned}\mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}(\rho)\right] &= \mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}\left(\sum_{k=0}^{\infty} p_{k,k}\,|k\rangle\langle k|\right)\right] \\ &= \sum_{k=0}^{\infty} p_{k,k}\,\mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}(|k\rangle\langle k|)\right],\end{aligned} \tag{B.2}$$

where $p_{k,k}$ is the k-th diagonal element of ρ.

We want our ultimate bounds to be in terms of the average photon number of the states. To relate the above quantity to this, we claim that

$$\sum_{k=0}^{\infty} p_{k,k}\,\mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}(|k\rangle\langle k|)\right] \geq \mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}\left(|\langle\hat{n}\rangle_{\rho}\rangle\langle\langle\hat{n}\rangle_{\rho}|\right)\right]. \tag{B.3}$$

That is to say, the average of the probabilities of losing each of many different photon number states is greater than the probability of losing one state of the average photon number (which we assume without loss of generality to be an integer).

Since a state is mapped to $|0\rangle\langle 0|$ if and only if it loses all of its photons, we can say that

$$\mathrm{Tr}\left[|0\rangle\langle 0|\,\mathcal{E}(|k\rangle\langle k|)\right] = (1-\eta)^{k}. \tag{B.4}$$

We will consider first the case of ρ being a mixture of only two Fock states, with weightings p and $1-p$ and respective photon numbers n and m. By using Eq. B.4, Eq. B.3 then becomes

$$p(1-\eta)^{n} + (1-p)(1-\eta)^{m} \geq (1-\eta)^{pn+(1-p)m}. \tag{B.5}$$

If we let $n = m + \delta$, then this simplifies to

$$-y^p + py - p + 1 \geq 0, \tag{B.6}$$

where we have used $y \equiv (1-\eta)^\delta$. The claim then reduces to proving that this polynomial is satisfied for all $y, p \in [0,1]$.

Let $f(p) = -y^p + py - p + 1$. We have $f(0) = f(1) = 0$. This function has a unique stationary point between 0 and 1, and the curvature $= -[\log(y)]^2 y^p$ is everywhere negative. Therefore $f(p) \geq 0$. This proves the claim for a bimodal initial state. The general claim follows by induction.

We now have that

$$\mathrm{Tr}\left[\hat{E}_{\checkmark}\mathcal{E}(\rho)\right] \geq (1-\eta)^{\langle\hat{n}\rangle_\rho} = \left(1 - \frac{\mu}{\langle\hat{n}\rangle_\rho}\right)^{\langle\hat{n}\rangle_\rho}. \tag{B.7}$$

Since the average *input* photon number is generally of the scale of dozens of orders of magnitude above unity, we may confidently take the limit of $\langle\hat{n}\rangle_\rho \to \infty$, which reduces Eq. B.7 to Eq. 4.27.

Appendix C

Detailed description of the analysis of the Innsbruck protocol

The first step in analysing the modified Innsbruck protocol is to construct the associated Markov matrix, as described in Section 6.4.1 and Appendix D. To do this we fix q_0, which sets the size of the matrix, and p_c, which determines the elements of the matrix. For a fixed t, a joint distribution of k_0 and k_1 is then calculated.

We assume that the states that are initially created after establishment is connected are Werner states of the form

$$\rho = \rho_W(F_{\text{init}}) = \frac{4F_{\text{init}} - 1}{3} \left|\Phi^+\right\rangle \left\langle\Phi^+\right| + \frac{1 - F_{\text{init}}}{3}\mathbb{1}. \tag{C.1}$$

where F_{init} is the fidelity with respect to $|\Phi^+\rangle$. We then choose some k_{max} that gives some largest acceptable error. Then from the marginal distribution on k_0 we then choose $2q_0 - 1$ values for k_0 (all of which are below k_{max}), and we choose a final value for k_0 and a value for k_1 from the full error distribution. These transform the $2q_0$ states (by q_0 on each section of the repeater network) as:

$$\rho \mapsto (1 - \epsilon_{W0})\rho + \epsilon_{W0}\mathbb{1}, \tag{C.2}$$

where ϵ_{W0} are defined as heralded errors as in Eq. (6.30), with $k = k_0$.

We now partition the set of states into a 'left set' and a 'right set', corresponding to the two different section of the network, and randomly apply a distillation to each set. To do this, we pair up the states within a set. If q_0 is odd, one state is randomly chosen to proceed to the next round without being distilled. The remaining pairs are distilled according to the DEJMPS protocol, which succeeds with probability N_D (Eq. 2.56).

As explained in the main text, when translating the distillation success probability to a term in the Markov matrix, we use a constant probability of distillation success, λ. This is related to our choice of k_{max} by

$$k_{\text{max}} = \left\lfloor \log\left(\frac{3\sqrt{2\lambda - 1}}{4F_{\text{init}} - 1}\right) \frac{1}{\log\left(1 - \epsilon_{W0}\right)} \right\rfloor. \tag{C.3}$$

This is because Two Werner states that have waited for $k_{\max}$ will be of fidelity $F_{\min}$. If these are distilled with each other the success probability will be no less than λ.

After the states on each section are distilled, the number of remaining states on each side, q_1^L and q_1^R, are random variables, with $p(q_1^{L,R} = x) = \lambda^x(1-\lambda)^{\frac{1}{2}q_0 - x}$. When we perform entanglement swapping to connect the two sections, the final number of states will be $q_1 = \min(q_1^L, q_1^R)$ with

$$\begin{aligned} p(q_1 = x) = p(q_1^L = x) \cdot \sum_{y=x}^{q_0/2} p(q_1^R = y) + \\ p(q_1^R = x) \cdot \sum_{y=x}^{q_0/2} p(q_1^L = y) - \\ p(q_1^L = x) \cdot (q_1^R = x) \end{aligned} \tag{C.4}$$

One of the two sets only then undergoes waiting errors while waiting for the other side to complete, by evolving according to Eq. (C.2) but with the ϵ_{W1} calculated from k_1.

For a fixed q_1, we then calculate the secret key rate as follows. We choose a random pairing of states on the left with states on the right. They are deterministically connected by applying a CNOT gate to the part of each Bell state stored in the repeater, and then measuring each in the X basis. This maps two states of diagonal coefficients[1] $(a_1, b_1, c_1, d_1), (a_2, b_2, c_2, d_2)$ to one with coefficients

$$\begin{bmatrix} a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 \\ a_1b_2 + a_2b_1 + c_1d_2 + c_2d_1 \\ a_1c_2 + a_2c_1 + b_1d_2 + b_2d_1 \\ a_1d_2 + a_2d_1 + b_1c_2 + b_2c_1 \end{bmatrix}. \tag{C.5}$$

These final states may then be distilled again. We optimize over combinations of distillation pairings to produce $q_2 \leq q_1/2$ final pairs, in order to maximise the secret key rate, given by

$$K(t|q_0, q_1, p_c, \epsilon_W) = \sum_{i=1}^{q_2} 1 - 2h_2(\epsilon_i), \tag{C.6}$$

where ϵ_i is the bit error of the i^{th} entangled pair, averaged between measuring in the Z basis and the X basis.

We must finally multiply p_t by the probability that none of the states involved in completing the process were of a fidelity less than $F_{\min}$. As such, we make the transformation

$$p_t \mapsto p_t \times \left(\frac{1}{p_t} \sum_{k=k_{\max}}^{t} p(k|t) \right)^{q_0}. \tag{C.7}$$

[1] On the basis ordered as $|\Phi^+\rangle, |\Psi^-\rangle, |\Psi^+\rangle, |\Phi^-\rangle$.

This key rate is optimized over distillation strategies (both before and after entanglement-swapping) and entanglement-swapping pairing choices, and averaged over values of q_1 and selections of sets of k_0, k_1 from the distribution to get $K(t|q_0, p_{\mathrm{c}}, \epsilon_W)$, which is used to get $K(q_0, p_{\mathrm{c}}, \epsilon_W)$ by Eq. (6.37).

Appendix D

Simplifying Markov matrices with high symmetry

The Markov matrix for a single section given in Eq. (6.33) is 2^{q_0}-dimensional. While this accurately describes the dynamics of the system, we can take advantage of the fact that the system contains a high degree of symmetry to reduce the size of the matrix. We can use the fact that the probability to move between one state and another is only dependent on *how many* entanglements have been established in the initial and final states, and not on the specifics of *which* entanglements. We can therefore use a technique called "lumping," where we create a partition of the states into sets, as shown in Fig. D.1 (discussed in more detail in [KS83]). From this we can consider a new process, where each set of states is considered as a single state.

When we lump states together, we should ensure that the transition probabilities in the lumped process produce the same system behaviour as in the unlumped process. Let M be the (unlumped) Markov matrix for the process, and $A_1, A_2, \cdots$ be a partitioning of the states, where each A is a set of states disjoint from all other sets. Then in order to be able to lump the process we require that, for each A_m and A_n, $\sum_{i \in A_m} [M]_{i,j}$ should be identical for all $j \in A_n$.

For our system this is true when we simply consider the transition probabilities, but the symmetry is broken when we include the complex counting variable, w_0, since this is only applied to one of the pairs. However, we can re-introduce a symmetry here, since w_0 is designed to capture the error rate on a typical pair, and not a particular pair. We pre-multiply the original Markov matrix by an *in-set maximal mixing matrix*, M_{mix}, which takes us from some state to any other state with the same number of completed entangled pairs with equal probability. This is given by a block diagonal, where each block has all elements equal to $1/n$, where here n is the size of the block. This is shown for the $q_0 = 3$ case below.

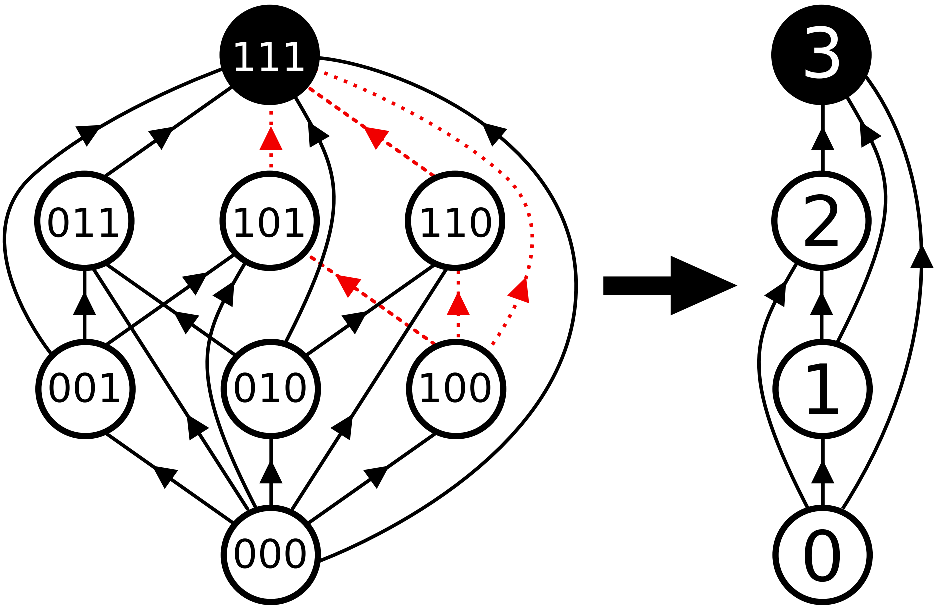


FIGURE D.1: On the left we have the Markov matrix for 3 pairs of entanglement trying to establish in parallel, with transition probabilities not shown for clarity, constructed in a way that tracks the status of each pair. The binary codes on each state show whether the first, second and third pairs are connected (1) or unconnected (0). Shown as red dotted lines are the events involving the first pair waiting after its completion. These edges translate to terms in the matrix that should be multiplied by an error-counting variable, w_0. On the right, we have grouped the states by how many pairs are connected. Terminating nodes shown in black. If the probability for each unentangled pair to establish its entanglement in a given time-step is p_c, then the probability to transition from node j to node i in the lumped process after application of the mixing matrix is given by $\binom{q-j}{q-i} p_\mathrm{c}^{i-j}(1-p_\mathrm{c})^{q-i}[q+(w_0-1)j]/q$.

$$M_\mathrm{mix} = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} \tag{D.1}$$

This effectively distributes the counting variable amongst the states, giving it the symmetry required to lump the states together. This reduces the number of unique states in the process corresponding to one section from 2^q to $q+1$.

Appendix E

Proof of Eq. 6.42

The distance over which communication has to occur at level l scales with 2^l. Given two sections of a repeater, there is some number of time-steps $k_{2\text{sec}}$ between the first completing and the second. After the second section completes, there must be one more round of classical communication to indicate this fact. Therefore $k_{A,l} = 2^l(k_{2\text{sec}} + 1)$.

We wish to calculate $k_{2\text{sec}}$, which is given by $\mathbb{E}[|x-y|]$, where x and y are two times drawn from the distribution, $f(t) = \mathrm{d}_t C(t)$, where the cumulative distribution function is given by $C(t) = [1-(1-p_\mathrm{c})^t]^q$. Approximating these as continuous distributions, we can write this as

$$\begin{aligned}\mathbb{E}[x-y|x>y] + \mathbb{E}[y-x|y>x] = \\ 2\int_0^\infty \int_0^y (y-x)\, f(x)\, f(y)\, \mathrm{d}x\, \mathrm{d}y.\end{aligned} \tag{E.1}$$

Let this inner $(\int \cdots \mathrm{d}x)$ integral be I. Then

$$\begin{aligned}I(y) &= y\,C(y) - \int_0^y x f(x)\, \mathrm{d}x, \\ &= y\,C(y) - \left\{\int_0^y \frac{\mathrm{d}}{\mathrm{d}x}\left[x\,C(x)\right] \mathrm{d}x - \int_0^y C(x)\, \mathrm{d}x\right\}, \\ &= \int_0^y C(x)\, \mathrm{d}x, \\ &\le y\,C(y).\end{aligned} \tag{E.2}$$

Therefore, we have

$$\begin{aligned}\mathbb{E}[|x-y|] &\le 2\int_0^\infty f(y)\, y\, C(y)\, \mathrm{d}y, \\ &= 2\int_0^1 y\, C\, \mathrm{d}C, \\ &= \frac{2}{\log(1-p_\mathrm{c})}\int_0^1 \log\left(1 - C^{1/q}\right) C\, \mathrm{d}C, \\ &= \frac{H(2q)}{|\log(1-p_\mathrm{c})|},\end{aligned} \tag{E.3}$$

where $H(n) = \sum_{m=1}^{n} 1/m$ is the n^{th} harmonic number. Here, q is equal to the total number of elementary pairs that need to connect in each "section" at a given level, which is given by $2^{N_S - l}$, which arrives at Eq. 6.42.

Bibliography

[AAA$^+$13] Junaid Aasi, J Abadie, BP Abbott, Richard Abbott, TD Abbott, MR Abernathy, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, et al. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7(8):613, 2013.

[ABB$^+$13] Silvestre Abruzzo, Sylvia Bratzik, Nadja K Bernardes, Hermann Kampermann, Peter van Loock, and Dagmar Bruß. Quantum repeaters and quantum key distribution: Analysis of secret-key rates. *Physical Review A*, 87(5):052315, 2013.

[AG$^+$79] Barry C Arnold, Richard A Groeneveld, et al. Bounds on expectations of linear systematic statistics based on dependent samples. *The Annals of Statistics*, 7(1):220–223, 1979.

[ATL15] Koji Azuma, Kiyoshi Tamaki, and Hoi-Kwong Lo. All-photonic quantum repeaters. *Nature Communications*, 6:6787, 2015.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India*. IEEE, 1984.

[BBD$^+$09] Hans J Briegel, David E Browne, Wolfgang Dür, Robert Raussendorf, and Maarten Van den Nest. Measurement-based quantum computation. *Nature Physics*, 5(1):19, 2009.

[BBFM06] Simon C. Benjamin, Daniel E. Browne, Joe Fitzsimons, and John J L Morton. Brokered graph-state quantum computation. *New Journal of Physics*, 8, 2006.

[BBP$^+$96] Charles H Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A Smolin, and William K Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76(5):722, 1996.

[BBP15] Leonardo Banchi, Samuel L. Braunstein, and Stefano Pirandola. Quantum Fidelity for Arbitrary Gaussian States. *Physical Review Letters*, 115(26):1–13, 2015.

[BCG$^+$02] Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, and Alain Tapp. Authentication of quantum messages. In *Foundations* *of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 449–458. IEEE, 2002.

[BDCZ98] H-J Briegel, Wolfgang Dür, Juan I Cirac, and Peter Zoller. Quantum repeaters: the role of imperfect local operations in quantum communication. *Physical Review Letters*, 81(26):5932, 1998.

[BDSW96] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996.

[Bel64] John S Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[Ber09] Daniel J Bernstein. Introduction to post-quantum cryptography. In *Post-quantum cryptography*, pages 1–14. Springer, 2009.

[BK98] Samuel L Braunstein and H Jeff Kimble. Teleportation of continuous quantum variables. *Physical Review Letters*, 80(4):869, 1998.

[BK05] Sean D. Barrett and Pieter Kok. Efficient high-fidelity quantum computation using matter qubits and linear optics. *Physical Review A*, 71(6):060310, 2005.

[BKPV99] S. Bose, P. L. Knight, M. B. Plenio, and V Vedral. Proposal for teleportation of an atomic state via cavity decay. *Physical Review Letters*, 83(24):5158, 1999.

[Bla52] Max Black. The identity of indiscernibles. *Mind*, 61(242):153–164, 1952.

[BN01] Fabio Benatti and Heide Narnhofer. Additivity of the entanglement of formation. *Physical Review A*, 63(4):042306, 2001.

[Bol66] Ludwig Boltzmann. *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie:(vorgelegt in der Sitzung am 8. Februar 1866)*. Staatsdruckerei, 1866.

[Bol12] Ludwig Boltzmann. *Lectures on gas theory*. Courier Corporation, 2012.

[BP12] Samuel L. Braunstein and Stefano Pirandola. Side-channel-free quantum key distribution. *Physical Review Letters*, 108(13):1–4, 2012.

[BR05] Daniel E. Browne and Terry Rudolph. Resource-efficient linear optical quantum computation. *Physical Review Letters*, 95(1):6–9, 2005.

[Bra05] Samuel L. Braunstein. Squeezing as an irreducible resource. *Physical Review A - Atomic, Molecular, and Optical Physics*, 71(5):8–11, 2005.

[BRS10] Sean D. Barrett, Peter P. Rohde, and Thomas M. Stace. Scalable quantum computing with atomic ensembles. *New Journal of Physics*, 12:1–8, 2010.

[Cal17] Marcello Caleffi. Optimal routing for quantum networks. *IEEE Access*, 5:22299–22312, 2017.

[CB08] Earl T. Campbell and Simon C. Benjamin. Measurement-based entanglement under conditions of extreme photon loss. *Physical Review Letters*, 101http://(13):16–19, 2008.

[CCB18] Marcello Caleffi, Angela Sara Cacciapuoti, and Giuseppe Bianchi. Quantum internet: from communication to distributed computing! *arXiv preprint arXiv:1805.04360*, 2018.

[CCGFZ99] C. Cabrillo, J. Ignacio Cirac, P. Garcia-Fernandez, and P. Zoller. Creation of entangled states of distant atoms by interference. *Physical Review A*, 59(2):1025, 1999.

[CDT$^+$06] L Childress, MV Gurudev Dutt, JM Taylor, AS Zibrov, F Jelezko, J Wrachtrup, PR Hemmer, and MD Lukin. Coherent dynamics of coupled electron and nuclear spin qubits in diamond. *Science*, 314(5797):281–285, 2006.

[CF06] Yves Couder and Emmanuel Fort. Single-particle diffraction and interference at a macroscopic scale. *Physical Review Letters*, 97(15):154101, 2006.

[CKW00] Valerie Coffman, Joydip Kundu, and William K Wootters. Distributed entanglement. *Physical Review A*, 61(5):052306, 2000.

[CRD03] C. W. Carr, H. B. Radousky, and S. G. Demos. Wavelength Dependence of Laser-Induced Damage: Determining the Damage Initiation Mechanisms. *Physical Review Letters*, 91(12):127402, 2003.

[CS01] Marcos Curty and David J Santos. Quantum authentication of classical messages. *Physical Review A*, 64(6):062309, 2001.

[CTSL06] L. Childress, J. M. Taylor, A. S. Sørensen, and M. D. Lukin. Fault-tolerant quantum communication based on solid-state photon emitters. *Physical Review Letters*, 96(7):96–99, 2006.

[Cur12] Marcos Curty. Device-independent quantum key distribution. In *Quantum Electronics and Laser Science Conference*, pages JTh4K–6. Optical Society of America, 2012.

[CZ95] Juan I Cirac and Peter Zoller. Quantum computations with cold trapped ions. *Physical Review Letters*, 74(20):4091, 1995.

[DB07] W. Dür and H. J. Briegel. Entanglement purification and quantum error correction. *Rep. Prog. Phys.*, 70:1381, 2007.

[DBCZ99] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller. Quantum repeaters based on entanglement purification. *Physical Review A*, 59(1):169–181, 1999.

[DCL91] T J Driscoll, J M Calo, and N M Lawandy. Explaining the optical fuse. *Optics Letters*, 16(13):1046–8, 1991.

[DEJ$^+$96] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Physical Review Letters*, 77(1):2818, 1996.

[Deu85] David Deutsch. Quantum theory, the church–turing principle and the universal quantum computer. *Proc. R. Soc. Lond. A*, 400(1818):97–117, 1985.

[DG15] Bryce Seligman DeWitt and Neill Graham. *The many worlds interpretation of quantum mechanics*. Princeton University Press, 2015.

[DH76] Gordon Davies and MF Hamer. Optical studies of the 1.945 ev vibronic band in diamond. *Proc. R. Soc. Lond. A*, 348(1653):285–298, 1976.

[Die82] DGBJ Dieks. Communication by epr devices. *Physics Letters A*, 92(6):271–272, 1982.

[DK03] L.-M. Duan and H. J. Kimble. Efficient engineering of multiatom entanglement through single-photon detections. *Physical Review Letters*, 90(25):253601, 2003.

[DLCZ01] L M Duan, M D Lukin, J I Cirac, and P Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Nature*, 414(6862):413–418, 2001.

[DLL03] Fu-Guo Deng, Gui Lu Long, and Xiao-Shu Liu. Two-step quantum direct communication protocol using the einstein-podolsky-rosen pair block. *Physical Review A*, 68(4):042317, 2003.

[DLZZ05] Fu-Guo Deng, Xi-Han Li, Hong-Yu Zhou, and Zhan-jun Zhang. Improving the security of multiparty quantum secret sharing against trojan horse attack. *Physical Review A*, 72(4):044302, 2005.

[dOCF14] Thiago R de Oliveira, Marcio F Cornelio, and Felipe F Fanchini. Monogamy of entanglement of formation. *Physical Review A*, 89(3):034303, 2014.

[DS13] Michel H Devoret and Robert J Schoelkopf. Superconducting circuits for quantum information: an outlook. *Science*, 339(6124):1169–1174, 2013.

[DV13] Jeroen Delvaux and Ingrid Verbauwhede. Side channel modeling attacks on 65nm arbiter pufs exploiting cmos device noise. In *Hardware-Oriented Security and Trust (HOST), 2013 IEEE International Symposium on*, pages 137–142. IEEE, 2013.

[EI57] Hugh Everett III. " relative state" formulation of quantum mechanics. *Reviews of modern physics*, 29(3):454, 1957.

[EKB15] Michael Epping, Hermann Kampermann, and Dagmar Bruß. Graph State Quantum Repeater Networks. *arXiv*, 2015.

[EKB16] Michael Epping, Hermann Kampermann, and Dagmar Bruß. Robust entanglement distribution via quantum network coding. *New Journal of Physics*, 18(10):103052, 2016.

[Eke91] Artur K Ekert. Quantum cryptography based on bell's theorem. *Physical Review Letters*, 67(6):661, 1991.

[ESR$^+$10] Dirk Englund, Brendan Shields, Kelley Rivoire, Fariba Hatami, Jelena Vuckovic, Hongkun Park, and Mikhail D Lukin. Deterministic coupling of a single nitrogen vacancy center to a photonic crystal cavity. *Nano Letters*, 10(10):3922–3926, 2010.

[Fey82] Richard P Feynman. Simulating physics with computers. *International journal of theoretical physics*, 21(6-7):467–488, 1982.

[FLO$^+$00] J. Fujita, M. Levy, R. M. Osgood, L. Wilkens, and H. Dötsch. Waveguide optical isolator based on Mach–Zehnder interferometer. *Applied Physics Letters*, 76(16):2158–2160, 2000.

[FMS14] Christopher A Fuchs, N David Mermin, and Rüdiger Schack. An introduction to qbism with an application to the locality of quantum mechanics. *American Journal of Physics*, 82(8):749–754, 2014.

[For81] Bengt Fornberg. Numerical differentiation of analytic functions. *ACM Transactions on Mathematical Software (TOMS)*, 7(4):512–526, 1981.

[FR88] Steven French and Michael Redhead. Quantum physics and the identity of indiscernibles. *The British Journal for the Philosophy of Science*, 39(2):233–246, 1988.

[FRS$^+$03] Mark Friesen, Paul Rugheimer, Donald E Savage, Max G Lagally, Daniel W van der Weide, Robert Joynt, and Mark A Eriksson. Practical design and simulation of silicon-based quantum-dot qubits. *Physical Review B*, 67(12):121301, 2003.

[FS16] Christopher A Fuchs and Blake C Stacey. Qbist quantum mechanics: Quantum theory as a hero's handbook. Technical report, 2016.

[Fuc17] Christopher A Fuchs. On participatory realism. In *Information and Interaction*, pages 113–134. Springer, 2017.

[FVDG99] Christopher A Fuchs and Jeroen Van De Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999.

[GFK$^+$06] Nicolas Gisin, Sylvain Fasel, Barbara Kraus, Hugo Zbinden, and Grégoire Ribordy. Trojan-horse attacks on quantum-key-distribution systems. *Physical Review A*, 73(2):022320, 2006.

[Gib78] J Willard Gibbs. Art. lii.–on the equilibrium of heterogeneous substances. *American Journal of Science and Arts (1820-1879)*, 16(96):441, 1878.

[Gla63] Roy J Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131(6):2766, 1963.

[GLLP04] Daniel Gottesman, H-K Lo, Norbert Lutkenhaus, and John Preskill. Security of quantum key distribution with imperfect devices. In *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, page 136. IEEE, 2004.

[GRTZ02] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74(1):145–195, 2002.

[Haw99] Margaret Hawton. Photon position operator with commuting components. *Physical Review A*, 59(2):954, 1999.

[HBB99] Mark Hillery, Vladimír Bužek, and André Berthiaume. Quantum secret sharing. *Physical Review A*, 59(3):1829, 1999.

[HBD$^+$15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[HEB04] Marc Hein, Jens Eisert, and Hans J Briegel. Multiparty entanglement in graph states. *Physical Review A*, 69(6):062311, 2004.

[Hel69] Carl W Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969.

[HHM84] RT Harley, MJ Henderson, and RM Macfarlane. Persistent spectral hole burning of colour centres in diamond. *Journal of Physics C: Solid State Physics*, 17(8):L233, 1984.

[HK18] David L Hurst and Pieter Kok. Analytic few-photon scattering in waveguide qed. *Physical Review A*, 97(4):043850, 2018.

[Hol95] Peter R Holland. *The quantum theory of motion: an account of the de Broglie-Bohm causal interpretation of quantum mechanics*. Cambridge university press, 1995.

[HPLG07] Yucheng Hu, Xiang Peng, Tiejun Li, and Hong Guo. On the Poisson approximation to photon distribution for faint lasers. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 367(3):173–176, 2007.

[HS13] Michael Hutter and Jörn-Marc Schmidt. The temperature side channel and heating fault attacks. In *International Conference on Smart Card Research and Advanced Applications*, pages 219–235. Springer, 2013.

[Hwa03] Won-Young Hwang. Quantum key distribution with high loss: toward global secure communication. *Physical Review Letters*, 91(5):057901, 2003.

[HWVE14] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic'for quantum computation. *Nature*, 510(7505):351, 2014.

[IAB$^+$99] A Imamog, David D Awschalom, Guido Burkard, David P DiVincenzo, Daniel Loss, M Sherwin, A Small, et al. Quantum information processing using quantum dot spins and cavity qed. *Physical Review Letters*, 83(20):4204, 1999.

[IHBW90] Wayne M Itano, Daniel J Heinzen, JJ Bollinger, and DJ Wineland. Quantum zeno effect. *Physical Review A*, 41(5):2295, 1990.

[JAK$^+$14] Nitin Jain, Elena Anisimova, Imran Khan, Vadim Makarov, Christoph Marquardt, and Gerd Leuchs. Trojan-horse attacks threaten the security of practical quantum cryptography. *New Journal of Physics*, 16(12):123030, 2014.

[JDG$^+$15] Sam Johnson, PR Dolan, Thomas Grange, AAP Trichet, Gaston Hornecker, Yu-Chen Chen, Laiyi Weng, GM Hughes, AAR Watt, A Auffèves, et al. Tunable cavity coupling of the zero phonon line of a nitrogen-vacancy defect in diamond. *New Journal of Physics*, 17(12):122003, 2015.

[JGP$^+$04] Fedor Jelezko, T Gaebel, I Popa, A Gruber, and Jorg Wrachtrup. Observation of coherent oscillations in a single electron spin. *Physical Review Letters*, 92(7):076401, 2004.

[Joz94] Richard Jozsa. Fidelity for mixed quantum states. *Journal of modern optics*, 41(12):2315–2323, 1994.

[JPE$^+$13] Dirk Jalas, Alexander Petrov, Manfred Eich, Wolfgang Freude, Shanhui Fan, Zongfu Yu, Roel Baets, Miloš Popović, Andrea Melloni, John D. Joannopoulos, Mathias Vanwolleghem, Christopher R. Doerr, and Hagen Renner. What is-and what is not-an optical isolator. *Nature Photonics*, 7(8):579–582, 2013.

[JSK$^+$15] Nitin Jain, Birgit Stiller, Imran Khan, Vadim Makarov, Christoph Marquardt, and Gerd Leuchs. Risk analysis of trojan-horse attacks on practical quantum key distribution systems. *IEEE Journal of Selected Topics in Quantum Electronics*, 21(3):168–177, 2015.

[Ken77] David G Kendall. The diffusion of shape. *Advances in applied probability*, 9(3):428–430, 1977.

[KGD$^+$15] Hari Krovi, Saikat Guha, Zachary Dutton, Joshua A. Slater, Christoph Simon, and Wolfgang Tittel. Practical Quantum Repeaters with Parametric Down-Conversion Sources. page 19, 2015.

[Kim08] H Jeff Kimble. The quantum internet. *Nature*, 453(7198):1023, 2008.

[KL10] Pieter Kok and Brendon W. Lovett. *Introduction to Optical Quantum Information Processing*. 2010.

[KLH$^+$15] Boris Korzh, Charles Ci Wen Lim, Raphael Houlmann, Nicolas Gisin, Ming Jun Li, Daniel Nolan, Bruno Sanguinetti, Rob Thew, and Hugo Zbinden. Provably secure and practical quantum key distribution over 307 km of optical fibre. *Nature Photonics*, 9(3):163, 2015.

[KLM01] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. *nature*, 409(6816):46, 2001.

[KMJ$^+$11] Hanna Krauter, Christine A Muschik, Kasper Jensen, Wojciech Wasilewski, Jonas M Petersen, J Ignacio Cirac, and Eugene S Polzik. Entanglement generated by dissipation and steady state entanglement of two macroscopic objects. *Physical Review Letters*, 107(8):080503, 2011.

[KMN$^+$07] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Reviews of Modern Physics*, 79(1):135–174, 2007.

[KMW02] David Kielpinski, Chris Monroe, and David J Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709, 2002.

[Koa09] M. Koashi. Simple security proof of quantum key distribution based on complementarity. *New Journal of Physics*, 11, 2009.

[Koc96] Paul C Kocher. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In *Annual International Cryptology Conference*, pages 104–113. Springer, 1996.

[KPA$^+$18] EO Kiktenko, NO Pozhar, MN Anufriev, AS Trushechkin, RR Yunusov, Yu V Kurochkin, AI Lvovsky, and AK Fedorov. Quantum-secured blockchain. *Quantum Science and Technology*, 3(3):035004, 2018.

[KS83] John G Kemeny and J Laurie Snell. *Finite Markov chains: with a new appendix" Generalization of a fundamental matrix"*. Springer, 1983.

[KWD03] Pieter Kok, Colin P Williams, and Jonathan P Dowling. Construction of a quantum repeater with linear optics. *Physical Review A*, 68(2):022301, 2003.

[Lal01] Franck Laloë. Do we really understand quantum mechanics? strange correlations, paradoxes, and theorems. *American Journal of Physics*, 69(6):655–701, 2001.

[LBSB13] Ying Li, Sean D Barrett, Thomas M Stace, and Simon C Benjamin. Long range failure-tolerant entanglement distribution. *New Journal of Physics*, 15(2):023012, 2013.

[LC99] Hoi-Kwong Lo and Hoi Fung Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283(5410):2050–2056, 1999.

[LCA05] Hoi-Kwong Lo, Hoi Fung Chau, and Mohammed Ardehali. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 18(2):133–165, 2005.

[LCQ12] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Physical Review Letters*, 108(13):130503, 2012.

[LCW$^+$15] Marco Lucamarini, Iris Choi, Martin B Ward, James F Dynes, ZL Yuan, and Andrew J Shields. Practical security bounds against the trojan-horse attack in quantum key distribution. *Physical Review X*, 5(3):031030, 2015.

[LFU17] Mikołaj Lasota, Radim Filip, and Vladyslav C. Usenko. Robustness of quantum key distribution with discrete and continuous variables to channel noise. 062312(June):1–13, 2017.

[LLK07] Antía Lamas-Linares and Christian Kurtsiefer. Breaking a quantum key distribution system through a timing side channel. *Optics express*, 15(15):9388–9393, 2007.

[Llo08] Seth Lloyd. Enhanced Sensitivity of Photodetection via Quantum Illumination. *Science*, 321(5895):1463–1465, 2008.

[LLY06] Hwayean Lee, Jongin Lim, and HyungJin Yang. Quantum direct communication with authentication. *Physical Review A*, 73(4):042305, 2006.

[LMY$^+$05] T. D. Ladd, D. Maryenko, Y. Yamamoto, E. Abe, and K. M. Itoh. Coherence time of decoupled nuclear spins in silicon. *Physical Review B*, 71(1):1–12, 2005.

[LP07] Hoi-Kwong Lo and John Preskill. Security of quantum key distribution using weak coherent states with nonrandom phases. *Quantum Information & Computation*, 7(5):431–458, 2007.

[LWW$^+$10] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature photonics*, 4(10):686–689, 2010.

[LYDS18] M Lucamarini, ZL Yuan, JF Dynes, and AJ Shields. Overcoming the rate–distance limit of quantum key distribution without quantum repeaters. *Nature*, 557(7705):400, 2018.

[LZG$^+$07] Chao-Yang Lu, Xiao-Qi Zhou, Otfried Gühne, Wei-Bo Gao, Jin Zhang, Zhen-Sheng Yuan, Alexander Goebel, Tao Yang, and Jian-Wei Pan. Experimental entanglement of six photons in graph states. *Nature physics*, 3(2):91, 2007.

[Mat92] Richard D Mattuck. *A guide to Feynman diagrams in the many-body problem*. Courier Corporation, 1992.

[MBF10] Yuichiro Matsuzaki, Simon C. Benjamin, and Joseph Fitzsimons. Probabilistic growth of large entangled states with low error accumulation. *Physical Review Letters*, 104(February):1–4, 2010.

[MCG$^+$07] J Majer, JM Chow, JM Gambetta, Jens Koch, BR Johnson, JA Schreier, L Frunzio, DI Schuster, AA Houck, Andreas Wallraff, et al. Coupling superconducting qubits via a cavity bus. *Nature*, 449(7161):443, 2007.

[MH12] John Mathews and Russell Howell. *Complex analysis for mathematics and engineering*. Jones & Bartlett Publishers, 2012.

[MKL$^+$12] Peter Christian Maurer, Georg Kucsko, Christian Latta, Liang Jiang, Norman Ying Yao, Steven D Bennett, Fernando Pastawski, David Hunger, Nicholas Chisholm, Matthew Markham, et al. Room-temperature quantum bit memory exceeding one second. *Science*, 336(6086):1283–1286, 2012.

[MPA11] Lluís Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature communications*, 2:238, 2011.

[MPG05] Stefan Mangard, Thomas Popp, and Berndt M Gammel. Side-channel leakage of masked cmos gates. In *Cryptographers' Track at the RSA Conference*, pages 351–365. Springer, 2005.

[MSD$^+$12] W. J. Munro, a. M. Stephens, S. J. Devitt, K. a. Harrison, and Kae Nemoto. Quantum communication without the necessity of quantum memories. *Nature Photonics*, 6(October):777–781, 2012.

[MW95] Leonard Mandel and Emil Wolf. *Optical coherence and quantum optics*. Cambridge university press, 1995.

[MZH$^+$15] Robert McConnell, Hao Zhang, Jiazhong Hu, Senka Ćuk, and Vladan Vuletić. Entanglement with negative wigner function of almost 3,000 atoms heralded by one photon. *Nature*, 519(7544):439, 2015.

[NC02] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[Neu32] Johann v Neumann. *Mathematische grundlagen der quantenmechanik*. Springer, 1932.

[NFSM$^+$09] Sebastian Nauerth, Martin Fürst, Tobias Schmitt-Manderbach, Henning Weier, and Harald Weinfurter. Information leakage via side channels in freespace bb84 quantum cryptography. *New Journal of Physics*, 11(6):065001, 2009.

[NTD$^+$14] Kae Nemoto, Michael Trupke, Simon J Devitt, Ashley M Stephens, Burkhard Scharfenberger, Kathrin Buczak, Tobias Nöbauer, Mark S Everitt, Jörg Schmiedmayer, and William J Munro. Photonic architecture for scalable quantum information processing in diamond. *Physical Review X*, 4(3):031022, 2014.

[NTD$^+$16] Kae Nemoto, Michael Trupke, Simon J Devitt, Burkhard Scharfenberger, Kathrin Buczak, Jörg Schmiedmayer, and William J Munro. Photonic quantum networks formed from nv- centers. *Scientific Reports*, 6:26284, 2016.

[Ord67] JK Ord. Handbook of the poisson distribution. *Journal of the Operational Research Society*, 18(4):478–479, 1967.

[Orú14] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014.

[Par70] James L Park. The concept of transition in quantum mechanics. *Foundations of Physics*, 1(1):23–33, 1970.

[Par84] Derek Parfit. *Reasons and persons*. OUP Oxford, 1984.

[Pau12] Wolfgang Pauli. *General principles of quantum mechanics*. Springer Science & Business Media, 2012.

[PBC$^+$03] E Pazy, E Biolatti, T Calarco, I D'amico, P Zanardi, F Rossi, and P Zoller. Spin-based optical quantum computation via pauli blocking in semiconductor quantum dots. *EPL (Europhysics Letters)*, 62(2):175, 2003.

[PBG$^+$05] Nicholas A Peters, Julio T Barreiro, Michael E Goggin, Tzu-Chieh Wei, and Paul G Kwiat. Remote state preparation: arbitrary remote control of photon polarization. *Physical Review Letters*, 94(15):150502, 2005.

[Pen91] Roger Penrose. The emperor's new mind. *RSA Journal*, 139(5420):506–514, 1991.

[PHFB18] Giuseppe Pucci, Daniel M Harris, Luiz M Faria, and John WM Bush. Walking droplets interacting with single and double slits. *Journal of Fluid Mechanics*, 835:1136–1156, 2018.

[PKEG16] Mihir Pant, Hari Krovi, Dirk Englund, and Saikat Guha. Rate-distance tradeoff and resource costs for all-optical quantum repeaters. pages 1–15, 2016.

[POS$^+$15] Stefano Pirandola, Carlo Ottaviani, Gaetana Spedalieri, Christian Weedbrook, Samuel L Braunstein, Seth Lloyd, Tobias Gehring, Christian S Jacobsen, and Ulrik L Andersen. High-rate measurement-device-independent quantum cryptography. *Nature Photonics*, 9(6):397, 2015.

[PR15] Nicoló Lo Piparo and Mohsen Razavi. Long-distance trust-free quantum key distribution. *IEEE Journal of Selected Topics in Quantum Electronics*, 21(3):123–130, 2015.

[Pri13] Nicolas Privault. *Understanding Markov chains: examples and applications*. Springer Science & Business Media, 2013.

[PRM17] Nicoló Lo Piparo, Mohsen Razavi, and William J Munro. Memory-assisted quantum key distribution with a single nitrogen-vacancy center. *Physical Review A*, 96(5):052313, 2017.

[RHG05] T. C. Ralph, A. J. F. Hayes, and Alexei Gilchrist. Loss-Tolerant Optical Qubits. *Physical Review Letters*, 95(10):100501, 2005.

[RLY$^+$17] GL Roberts, M Lucamarini, ZL Yuan, JF Dynes, LC Comandar, AW Sharpe, AJ Shields, M Curty, IV Puthoor, and E Andersson. Experimental measurement-device-independent quantum digital signatures. *Nature communications*, 8(1):1098, 2017.

[Rov96] Carlo Rovelli. Relational quantum mechanics. *International Journal of Theoretical Physics*, 35(8):1637–1678, 1996.

[RSA78] Ronald L Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[Sch26] Erwin Schrödinger. Der stetige übergang von der mikro-zur makromechanik. *Naturwissenschaften*, 14(28):664–666, 1926.

[Sch17] Roman Schnabel. Squeezed states of light and their applications in laser interferometers. *Physics Reports*, 684:1–51, 2017.

[SDRA$^+$07] Christoph Simon, Hugues De Riedmatten, Mikael Afzelius, Nicolas Sangouard, Hugo Zbinden, and Nicolas Gisin. Quantum repeaters with photon pair sources and multimode memories. *Physical Review Letters*, 98(19):1–4, 2007.

[SFI$^+$11] Masahide Sasaki, M Fujiwara, H Ishizuka, W Klaus, K Wakui, M Takeoka, S Miki, T Yamashita, Z Wang, A Tanaka, et al. Field test of quantum key distribution in the tokyo qkd network. *Optics express*, 19(11):10387–10409, 2011.

[Sha48] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[Sho94] Peter W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 124–134. Ieee, 1994.

[SK93] Bruce W Shore and Peter L Knight. The jaynes-cummings model. *Journal of Modern Optics*, 40(7):1195–1238, 1993.

[SN16] Sujeevan Sivasundaram and Kristian Hvidtfelt Nielsen. Surveying the attitudes of physicists concerning foundational issues of quantum mechanics. *arXiv preprint arXiv:1612.00676*, 2016.

[SP00] Peter W Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Physical Review Letters*, 85(2):441, 2000.

[SRK$^+$15] Shihan Sajeed, Igor Radchenko, Sarah Kaiser, Jean Philippe Bourgoin, Anna Pappa, Laurent Monat, Matthieu Legré, and Vadim Makarov. Attacks exploiting deviation of mean photon number in quantum key distribution and coin tossing. *Physical Review A - Atomic, Molecular, and Optical Physics*, 91(3):1–13, 2015.

[SSDRG11] Nicolas Sangouard, Christoph Simon, Hugues De Riedmatten, and Nicolas Gisin. Quantum repeaters based on atomic ensembles and linear optics. *Reviews of Modern Physics*, 83(1):33, 2011.

[SSvL17] E Shchukin, F Schmidt, and P van Loock. On the waiting time in quantum repeaters with probabilistic entanglement swapping. *arXiv preprint arXiv:1710.06214*, 2017.

[Sta72] Henry Pierce Stapp. The copenhagen interpretation. *American Journal of Physics*, 40(8):1098–1116, 1972.

[Teg98] Max Tegmark. The interpretation of quantum mechanics: Many worlds or many words? *Fortschritte der Physik: Progress of Physics*, 46(6-8):855–862, 1998.

[Thi16] Harold Thimbleby. Human factors and missed solutions to enigma design weaknesses. *Cryptologia*, 40(2):177–202, 2016.

[TKI03] Kiyoshi Tamaki, Masato Koashi, and Nobuyuki Imoto. Unconditionally Secure Key Distribution Based on Two Nonorthogonal States. *Physical Review Letters*, 90(16):167904, 2003.

[TYC$^+$14] Yan-Lin Tang, Hua-Lei Yin, Si-Jing Chen, Yang Liu, Wei-Jun Zhang, Xiao Jiang, Lu Zhang, Jian Wang, Li-Xing You, Jian-Yu Guan, et al. Measurement-device-independent quantum key distribution over 200 km. *Physical Review Letters*, 113(19):190501, 2014.

[Uhl76] Armin Uhlmann. The "transition probability" in the state space of a*-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[VBR06] Michael Varnava, Daniel E. Browne, and Terry Rudolph. Loss tolerance in one-way quantum computation via counterfactual error correction. *Physical Review Letters*, 97(12):1–5, 2006.

[VK17] Scott E. Vinay and Pieter Kok. Practical repeaters for ultralong-distance quantum communication. *Physical Review A*, 95(5):1–7, 2017.

[VK18] Scott E. Vinay and Pieter Kok. Extended analysis of the Trojan-horse attack in quantum key distribution. *Physical Review A*, 97(April):042335, 2018.

[vLLMN08] Peter van Loock, Norbert Lütkenhaus, WJ Munro, and Kae Nemoto. Quantum repeaters using coherent-state communication. *Physical Review A*, 78(6):062319, 2008.

[VLLS$^+$06] P. Van Loock, T. D. Ladd, K. Sanaka, F. Yamaguchi, Kae Nemoto, W. J. Munro, and Y. Yamamoto. Hybrid quantum repeater using bright coherent light. *Physical Review Letters*, 96(24):1–4, 2006.

[VN27] John Von Neumann. Wahrscheinlichkeitstheoretischer aufbau der quantenmechanik. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1927:245–272, 1927.

[VOMG88] E Van Oort, NB Manson, and M Glasbeek. Optically detected spin coherence of the diamond nv centre in its triplet ground state. *Journal of Physics C: Solid State Physics*, 21(23):4385, 1988.

[VV14] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Physical Review Letters*, 113(14):140501, 2014.

[Wan05] Xiang-Bin Wang. Beating the photon-number-splitting attack in practical quantum cryptography. *Physical Review Letters*, 94(23):230503, 2005.

[WBC11] Christopher J Wood, Jacob D Biamonte, and David G Cory. Tensor networks and graphical calculus for open quantum systems. *arXiv preprint arXiv:1111.6950*, 2011.

[Wie83] Stephen Wiesner. Conjugate coding. *ACM Sigact News*, 15(1):78–88, 1983.

[WWLH09] Yong Wang, Huadeng Wang, Zhaohong Li, and Jinxiang Huang. Man-in-the-middle attack on bb84 protocol and its defence. In *Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on*, pages 438–439. IEEE, 2009.

[WZ82] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802, 1982.

[WZM$^+$16] Julius Wallnöfer, Michael Zwerger, Christine Muschik, Nicolas Sangouard, and Wolfgang Dür. Two-dimensional quantum repeaters. *Physical Review A*, 94(5):052307, 2016.

[ZCY18] Mehdi Mohammad Zadeh, Rahim Charkhi, and Zahra Yari. The mystical symbols in the images of the hell of the mir heydar miraj name. *BAGH-E NAZAR*, 15(60):67–82, 2018.

[ZDB12] M. Zwerger, W. Dür, and H. J. Briegel. Measurement-based quantum repeaters. *Physical Review A*, 85:1–11, 2012.