The Physical Underpinning of Security Proofs for Quantum Key Distribution

by

Jean Christian Boileau

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2007

Jean Christian Boileau

Jean Christian Boileau

Abstract

The dawn of quantum technology unveils a plethora of new possibilities and challenges in the world of information technology, one of which is the quest for secure information transmission. A breakthrough in classical algorithm or the development of a quantum computer could threaten the security of messages encoded using public key cryptosystems based on one-way function such as RSA. Quantum key distribution (QKD) offers an unconditionally secure alternative to such schemes, even in the advent of a quantum computer, as it does not rely on mathematical or technological assumptions, but rather on the universality of the laws of quantum mechanics.

Physical concepts associated with quantum mechanics, like the uncertainty principle or entanglement, paved the way to the first successful security proof for QKD. Ever since, further development in security proofs for QKD has been remarkable. But the connection between entanglement distillation and the uncertainty principle has remained hidden under a pile of mathematical burden. Our main goal is to dig the physics out of the new advances in security proofs for QKD. By introducing an alternative definition of private state, which elaborates the ideas of Mayers and Koashi, we explain how the security of all QKD protocols follows from an entropic uncertainty principle. We show explicitly how privacy amplification protocol can be reduced to a private state distillation protocol constructed from our observations about the uncertainty principle. We also derive a generic security proof for one-way permutation-invariant QKD protocols. Considering collective attack, we achieve the same secret key generation rate as the Devetak-Winter's bound. Generalizing an observation from Kraus, Branciard and Renner, we have provided an improved version of the secret key generation rates by considering a different symmetrization. In certain situations, we argue that Azuma's inequality can simplify the security proof considerably, and we explain the implication, on the security level, of reducing a QKD protocol to an entanglement or a more general private state distillation protocol.

In a different direction, we introduce a QKD protocol with multiple-photon encoding that can be implemented without a shared reference frame. We prove the unconditional security of this protocol, and discuss some features of the efficiency of multiple-photon QKD schemes in general.

Acknowledgements

When I started, five years ago, I was assigned to a NMR spectrometer. I was struggling and passed countless hours testing and experimenting without progress, not to mention the time I quenched one of the magnets... In short, I was far from being in a comfortable situation and it could have ended up pretty badly had it not been for the open-mindedness of my supervisor who gave me the liberty to choose my own project, even if it was not related to his primary interests. I owe Raymond Laflamme for his support and the myriad of opportunities he gave me. When Ray accepted me as his student, he told me I had to teach him new things. I hope this thesis will live up to his expectations.

I wish to thank David Poulin who helped me start my first project in quantum cryptography. He always had the patience to listen to my babbling and always managed to turn it into a fruitful discussion. I also thank Kyoshi Tamaki for everything I learned from him and all the fun we had working together. The main parts of this thesis were elaborated with Joseph Renes. I thank him for all the long discussions that we have had over the past two years, and hope it will continue for a very long time. I thank Daniel Gottesman, Norbert Lütkenhaus and Aswhin Nayak for answering so many of my questions.

I thank Cindy Schnarr, Connie Slaughter, Lorna Schmalz and Wendy Reibel who have been a tremendous help during the last five years. I thank all my colleagues including Donny Cheung, Niel deBeaudrap, Chris Erven, Gus Gutoski, Hauke Häseler, Matthias Heid, Georg Heinrich, Aziz Kara, Oliver Kern, Jonathan Lavoie, Tobias Moroder, Osama Moussa, Geir Ove Myhr, Camille Negrevergne, Harold Ollivier, Carlos Perez-Delgado, Tim Reid, Marcus Silva, Stephanie Simmons, Damien Simon, Devin Smith and Douglas Stebila for making my experience at IQC stimulating and enjoyable.

I thank Lana Sheridan for helping me with a project about covariant map. I thank Martin Laforest and Colm Ryan for being great roommates and colleagues. I thank Colm for the inspired title for this thesis. However, I won't thank Casey Myers for providing me with those vitamins pills that were supposed to keep me awake on my first car trip from Waterloo to Montreal (though, the night before was fun!), nor Veronica for ruthlessly stealing my roommate. Maybe one day I will forgive them. We will see.

On the other hand, I thank Casey's mother for receiving me a month in Brisbane. I would also like to thank Paul Tobin and Michael Bremner for the good time we had when I was there. I thank Professor Tao Yang and his student Qiang Zhang, for welcoming me in their group in Hefei, and Tenyen Chen, for showing me around Suzhou. I thank Jamie Batuwantudawe for being a great co-worker and travel partner. I also thank Stephen Bartlett and Robert Spekkens for allowing me to go to Sydney and start an interesting new project with them.

I would also like to thank my sister Nathalie, and brothers Jean Philippe and Jean-

Francois, my closest friends, Antoine, Hami, Olivier and Simon, and all their partners, Loudevick, Myriam, Noemi, Sandra, Mélissa and Émilie, for giving me unconditional support and for all the wonderful time we had together either in Montreal, Waterloo, Cuba or Paris. Most of all, I thank my love, Namrata. Her encouragements gave a boost to the efforts I put into this work.

I dedicate this thesis to my parents, Thérèse and Robert. After all these years, they are still my best role models. I am proud to be their son and love them with all my heart.

Contents

Chapter 1

Motivations

1.1 Quantum Revolution

Information technologies like radio, television, telephone, computer and internet have shaped modern civilizations. With the exception of some pioneer works, the theory of computation was, until very recently, restricted to information encoded in classical states and where processing could be described by classical mechanics. On the other hand, the laws of quantum mechanics allow a more general class of physical states and operations, and open the doors to a more general theory of information. The basic reason why quantum mechanics was not included at first in information theory (information theory was discovered after quantum mechanics) was that quantum states are very sensitive to thermal noise and are very difficult to manipulate. Most researchers simply thought that controlling quantum states was a crazy idea and were not interested in a more general theory of information. But, recent advances in quantum optics and nanotechnology have brought in a different perspective. The devices used for computation are becoming smaller and smaller every year, and quantum effects in these devices will soon be dominant. A small quantum structure can now be controlled very well, and quantum correlation observed between photons has been proved to be very robust when communicating over many kilometers.

Quantum information is the study of the new possibilities offered by quantum technologies. It has been discovered that some quantum algorithms, like the Grover search algorithm [74], can outperform the best known classical algorithms essential in some crucial day-to-day applications. Quantum mechanics can also be used in cryptography to obtain a higher level of security, which was not possible classically.

A large part of the research is to try and investigate other possible applications, improve our control of larger and more complex quantum structures and obtain a better understanding of quantum mechanics. In the following text, we present some of our contributions to the field of quantum cryptography. For a clearer comprehension of these contributions, we begin our explanation with a brief introduction to one-time pad and the field of contemporary cryptography followed by a history of quantum key distribution.

1.2 Modern Cryptography

1.2.1 One-time Pad and Perfect Security

An ideal cryptographic system sould not allow eavesdroppers to obtain any information whatsoever about the message, whatever resources they may have. Claude Shannon [167] proved that it is necessary and sufficient to use a random and perfectly secure key of the same length as that of the compressed version of the message to be able to make it perfectly information-theoretic secure.

Here is an illustration on how to obtain such perfect information-theoretic security using a *one-time pad*, followed by a discussion on the difficulties in fulfilling the requirements for such a high level of security.

Consider two parties - Alice and Bob. Bob asks Alice if the Holy Grail is hidden either in the sanctuary of Montserrat, the Rosslyn Chapel, the spring at Glastonbury Tor, Jeruzalem, the Louvre, the Vatican, the White House or elsewhere. Before sending her answer, Alice could compress her message in a three bit string (a_1, a_2, a_3) such that "the sanctuary of Montserrat" corresponds to 000, "the Rosslyn Chapel" to 001, etc. The compression algorithm is assumed to be known by Bob, and any third party. To obtain a perfect information-theoric security, Alice uses a random secret key composed of three bit (s_1, s_2, s_3) that she shares with Bob. This secret key is also known as a one-time pad. She sends the following three bit string to Bob:

$$(a_1, a_2, a_3) \bigoplus (s_1, s_2, s_3) = (a_1 \bigoplus s_1, a_2 \bigoplus s_2, a_3 \bigoplus s_3) \tag{1.1}$$

where addition modulo 2 is implicitly understood. If it is intercepted, then the eavesdropper would not be able to gain any information about the original message, and only someone with the one time pad could recover the original message using a simple transformation:

$$(a_1 \bigoplus s_1, a_2 \bigoplus s_2, a_3 \bigoplus s_3) \bigoplus (s_1, s_2, s_3) = (a_1, a_2, a_3) \tag{1.2}$$

followed by the decompression algorithm.

Perfect security in the above protocol is achieved if and only if the one-time pad is perfectly random and secret. Surprisingly, the randomness of the one time pad is not a trivial issue. Programs that generate random numbers on a computer follow algorithms and, in theory, are predictable. Flipping a coin to generate random bits is arguably very unpredictable, but also very inefficient. Nevertheless, those who are fastidious will remark that the flip process respects the laws of classical mechanics, which are also deterministic. In fact, the randomness of a key generated by any classical process is associated with the complexity of the physical system from which the numbers are extracted. Someone who is able to obtain a better and more detailed description of the physical ensemble could predict the outcome of the supposedly random process. For practical purposes, this notion of randomness may be more than enough, but it is interesting to note that the basic laws of quantum mechanics suggest that pure randomness could be achieved experimentally. By the uncertainty principle, the outcomes of a simple experiment, like measuring the circular polarization of a photon that was initially linearly polarized, are perfectly random up to the noise and apparatus imperfection. Assuming that the noise is independent from one outcome to another, it is possible, using error correction, to extract with high probability some perfectly random bits. Without some strict restriction on the dependence of the noise on different outcomes, it is impossible to obtain perfect randomness as shown in [51, 52]. Consequently, while the randomness obtained by a classical protocol lies in the complexity of a mathematical problem or a physical system, in the case of a quantum measurement it is based on the solidity of the laws of quantum mechanics and some assumptions about the noise's model. As we will see, there exists a similar relation between the secrecy of a one-time pad obtained from a classical and a quantum protocol.

To ensure the secrecy of the key, Alice and Bob could have met ulteriorly and agreed upon a random key. This might turn out to be very impractical, especially if Alice and Bob want to secretly communicate a message, which is larger than what they had initially expected. Note that they cannot reuse the same key with a different message without compromising the security of the key since, by adding the two encoded messages together, an eavesdropper would obtain a juxtaposition of the two original messages which is in most cases easy to decode. By the way, this is why this key is called a one-time pad. Alice and Bob would then be in serious trouble since they cannot create or extend an information-theoretic secret key using an insecure classical channel like a phone line, internet, Fedex, etc. See Appendix A.2 for more about that statement.

1.2.2 “Classical” Cryptography

Public-key distribution was invented to overcome the problem of distributing a secret key through an insecure channel. Its security is based on the use of a one-way function, i.e. a function easy to compute, but hard to invert. Following a specific public-key distribution protocol, Alice picks two keys randomly. She keeps one of them private and she broadcasts the other publicly. Assuming that he can authenticate the message sent by Alice (for example by knowing her web address or using a trusted third party), Bob uses the public key to encode his message and sends it to Alice through the insecure channel. With her private key, Alice can easily decrypt Bob's message. The security of this protocol lies on the assumption that any eavesdropper cannot *efficiently* (i.e. within an acceptable computation time) decode the message without the private key. As a classic example, RSA is a public-key distribution scheme for which the decoding is equivalent to factorizing a product of two large primes. If one of the primes is known, the other one can be efficiently calculated. However, no known classical algorithm can efficiently factorize arbitrary large numbers and current technology does not give us enough computational power to break RSA, even in many lifetimes. Other examples of public-key distribution protocol include the elliptic curve cryptography which is broadly used on the internet. Note that its security is based on the complexity of a mathematical problem other than factorization. Nonetheless, quantum computer can break elliptic crypto-systems efficiently using essentially ShorÕs algorithm for the discrete logarithm problem.

Public-key cryptography is also called *asymmetric* cryptography. There also exists *symmetric* cryptography in which Alice and Bob pre-share a secret key used to encrypt and decrypt a message. In fact, the protocol described above, involving the one-time pad, is an example of symmetric cryptography. As mentioned earlier, up to equivalent operations, it is the only symmetric cryptographic protocol that can achieve information-theoretic security. The security of symmetric cryptographic protocols that use a secret key smaller than the compressed message is given assuming that Eve (the eavesdropper) cannot efficiently solve some mathematical protocol, or that she is restricted in the kind of operations she can perform or the resources available to her. For example, there exist symmetric cryptographic protocols that are secure against an adversary that has a bounded storage capacity [48].

1.2.3 Shaky Ground

One thing particular to one-way functions is that no one really knows if they exist or not. The proof of existence of one-way functions would solve one of the most important open questions in computer science: $NP \neq P$ (i.e. not every decision problem for which the solution can be checked in polynomial time can actually be solved in polynomial time, and vice-versa). The difficulties encountered in solving this question have considerably reduced the hope of proving definitively the security of public-key distribution. However, it does not prevent most mathematicians from believing in the existence of a one-way function.

A problem like factoring is considered hard since no one has found an efficient classical algorithm for it even after tremendous efforts put in by the whole mathematics community. On the other hand, we can not rule out the minute possibility that someone could break RSA or elliptic curve cryptography tomorrow using a laptop bought at Future Shop. In a way, the security of cryptographic protocol based on mathematical complexity is a matter of faith more than an objective statement. It is normal to question that faith, especially in light of the new advances in quantum information.

Today, all computers available on the market follow computation rules described by classical mechanics. However, more general computation operations can be performed.

In theory, we could exploit all operations and states allowed by quantum mechanics. A device which such power is called a quantum computer. In 1994, Shor found an algorithm [172] to efficiently factorize large number using a quantum computer, although no one has found an efficient classical algorithm. It has been proven that a scalable quantum computer could perform some operations faster than the optimal strategy performed on a classical computer of similar size, but those speedups were not sufficient to show that quantum computers could solve efficiently (i.e. in polynomial time) some problems that are not efficiently solvable on a classical computer. Nevertheless, Shor's discovery suggests that it is most likely that such problems exist.

One of the main difficulties of building a quantum computer is to make it scalable, meaning that the register of the quantum computer could be extended while keeping the noise level in the devices below some threshold, as determined by the theory of quantum fault-tolerance [4, 98, 102, 146]. It turns out that presently no one knows how to build a scalable quantum computer. In fact, the security of RSA is based on the extra assumption that no one will be able to build a large quantum computer.

An issue arising with the development of quantum information is to find which cryptographic protocols based on computational complexity can stand up to a quantum computer. Finding a function (called quantum one-way function) easy to compute on a classical computer and hard to invert on a quantum computer would imply that there exists a standard one-way function and $NP \neq P$. In the probable case that no one succeeds in proving the existence of quantum one-way function and thus of standard one-way function, the mathematical community could propose some likely quantum one-way functions. The security of the cryptographic protocols will be directly proportional to the community's efforts to invert the quantum one-way functions, to which these protocols are reduced. This can be a good motivation for the study of algorithmic complexity in the context of quantum information.

One problem with cryptographic protocols based either on quantum or classical complexity assumptions is that if algorithmic and/or technological advancements allow the protocol to be broken, then someone with a recording of the past communications could decrypt the old messages. If someone wants to keep a communication secret for more than 20 or 40 years, then he could used a pre-shared one-time pad as described in the previous section or use a protocol based on a one-way function that he believes won't be broken during that whole time. In the next section, a third alternative, that does not depend on algorithmic complexity, has been presented. It is called quantum key distribution (QKD).

1.3 The Appeal of QKD

So far, we have discussed cryptography algorithms that can be described classically. More than ten years before Shor's discovery, Wiesner, Bennett and Brassard proposed protocols that could perform cryptographic tasks that are impossible classically by exploiting features from quantum mechanics. For example, information-theoretic secure key distribution through an authenticated channel is possible if Alice sends quantum states to Bob instead of restricting herself to classical communication. The concept of distributing a key using quantum protocols is called quantum key distribution (QKD). QKD is part of the wider field of quantum cryptography that includes the study of secrecy and privacy in all quantum protocols. It includes secret sharing, zero-knowledge proofs, quantum coin tossing, etc. The focus of our research is QKD.

Assuming that the channels are authenticated, the security of QKD relies on the dependability of the laws of quantum mechanics and not on the unproven complexity of a mathematical problem. QKD has the advantage of being robust against the advent of quantum computing and it could be useful to use it in combination with other standard cryptographic protocols to increase the security in the sense discussed earlier. However, one can argue that by not responding to the threats of side-channels or the man-in-the-middle attack, QKD does not strengthen the weaker security aspects of contemporary cryptography. QKD does not solve the problem of authentication of channels, which prevents an eavesdropper from impersonating one of the parties.

However, it is possible to unconditionally authenticate a channel by using a relatively small pre-shared key [35]. Therefore, QKD is a good method to extend a pre-shared secret key, something that cannot be done classically. There also exist authentication protocols that do not require a pre-shared secret key. However, their security is based on some computational assumptions. Suppose that QKD was combined with such authentication protocols. Then, if Eve is able to break the authentication protocol before the end of the QKD scheme, she can impersonate the receiver and obtain the secret message. However, if the duration of the QKD protocol is kept within a certain length in which Eve cannot break the authentication protocol, then the security of the distributed key is not threatened by the possibility that Eve may break the protocol in the future. This makes a clear contrast with the security of protocols discussed in the previous section, since the message might be decrypted later by using more time and the new technological and algorithmic developments. Another extraordinary feature of QKD is that, in standard conditions, it is possible to recycle the key used for authentication [116, 81] without compromising the information-theoretic security of the protocol —an issue referred to as the *composability* of the authentication scheme.

QKD was received with a lot of skepticism, and Bennett and Brassard had difficulties publishing their seminal paper on the subject. At the time, it was judged very unlikely that QKD could ever be implemented with good results. The experimental progress in the last few years has been encouraging, and the prospect of building good quantum memories and quantum repeaters has kept alive the hope of implementing reliable and fast long distance QKD. Nevertheless, critics point out that extra costs and poorer performance of QKD as compared to the key-distribution methods based on (quantum) one-way function might outbalance the advantages it provides.

Even considering the importance of private communication in our society, many believe that it is unlikely that QKD will replace public-key cryptography on a large scale since the improvement in security it could provide does not seem to be cost effective. Picking the right cryptography method for everyday applications is mostly a question of cost versus saving. Cryptography is used more as a deterrent to pirates and thieves than as an absolute technique for preventing crimes. For example, learning how to crack a DVD and spending time copying it is not worth the effort for most people. It is much simpler to buy it at the store. Thus, it might be worthwhile to upgrade the cryptographic protocol that encrypts the DVD, but only if the costs related to the changes are lower than the benefits engendered. The same is true for QKD. It is a question of whether installing a large scale QKD network would be worth the investment. Critics of QKD argue that it will be more cost effective to choose the key-distribution schemes that depend on more complex (quantum) one-way functions, but the only way to verify this is to do more theoretical and experimental research on QKD. The discovery and experiments on QKD are very recent, and we anticipate that new technologies and discoveries will improve its performance considerably. Even if QKD, or some form of it, would not be suitable for a large network, it could still be used for point-to-point communication like, for example, making a private link between the White House and the Pentagon.

For some, these arguments justify the efforts and time spent on developing QKD. There are other motivations. Looking back, QKD did not become famous because of the hope of cryptographers to improve the security of some useful protocols, but through the physics community because of the work of Ekert who independently discovered QKD and remarked that its security is related to the violation of Bell's inequality[55].

Physicists were attracted to quantum cryptography since it provided practical ways to describe and observe some important differences between quantum and classical mechanics. With the rest of quantum information, it provides a convenient tool to comprehend, teach and investigate the subtleties of quantum mechanics. As it is explained in this thesis, QKD is directly related to the uncertainty principle and to the production of private states, a generalization of the maximally entangled state. Investigating QKD is like analyzing entanglement distillation with constraints. The importance of these studies is underlined by the central role that uncertainty principles and entanglement plays in more complex quantum protocols and in quantum information in general.

1.4 QKD Experiments

In order to give the complete picture, we provide a brief overview of the experiments related to QKD. QKD was first implemented in 1989 by [16]. It exploited the polarization states of photons traveling in free space for a distance no longer than 30 cm. Loud electric sparks could be heard by anyone in the room, revealing full information about the quantum states emitted by the device. From then on, much improvement has been recorded. Although polarization encoding is suitable for free-space transmission of photons, a birefringence effect in optical fiber combined with thermal fluctuation causes random rotation of the polarization states of photon traveling through the fiber. This makes polarization encoding not the most favored encoding for QKD through optical fiber. In 1992, Bennett proposed a QKD scheme which encoded quantum information in the phase between two states of a photon corresponding to two different time bins [15]. The encoding and decoding is performed using a Mach-Zehnder interferometer for each operation. This protocol lead to some very good implementations of QKD, that can be used even over 100 km [63, 70].

The components used for phase-encoding QKD, like phase modulators, are polarization dependent. This implies that if the random polarization transformation in the fiber is not compensated, the performance of the QKD protocol will suffer greatly. An ingenious way to deal with this was proposed by [129] and implemented for QKD by [133, 178]. It consists of using a Faraday mirror such that photons can travel back and forth in the fiber so that the random polarization rotation can be effectively cancelled. An extra advantage of this scheme is that it is robust against phase instability of the interferometer (see [69] for details).

In the above set-ups, the most popular method to produce an approximation of single photons is to use faint laser pulses. The intensity of the pulses is so low that, on an average, they possess less than one photon each. However, a significant fraction of the pulses will contain two or more photons. Each of the photons contains a copy of the quantum states that Alice would have liked to send to Bob. Eve could simply take one of those copies to gain, about Alice's state, as much information as Bob. Therefore, a multi-photon pulse considerably threatens the security of QKD. To deal with this problem, some used entangled pairs of photons generated by parametric down-conversion [112], and used a measurement on one of the photons to determine the presence of a sister photon sent through the channel. There are also other methods to produce better single-photon pulses. But all those methods, including the one involving entangled pairs of photons generated by parametric down-conversion, have so far provided significantly less non-empty pulses than if a faint laser was used.

A famous technique to restrict the possible eavesdropping attacks on multiple-photon pulses is by using *decoy states* [88]. Some extra — and relatively simple — operations allow

Alice and Bob to make upper bound on the number of multiple-photon pulses that Eve has exploited, and generally considerably improve the key generation of the QKD protocol. Many long distance and fast QKD experiments have been performed using this method [190, 157, 141].

Other QKD schemes have been proposed for the simplicity of their set-up and their high raw key generation rates. For example, continuous-variable QKD using homodyne or heterodyne detection has been particularly popular [125]. New protocols that are not invariant under permutation, exploit the phase between photon pulses and are more robust against the photon splitting attack than conventional QKD protocols [179, 177]. For a more complete review of the different experiments related to QKD, we suggest [69, 126, 123].

1.5 Security Proofs

1.5.1 Types of Security Proofs for QKD

There are two main reasons for doubting QKD. First, because of the difficulties and related costs of implementing it as explained above, and second, because of the lack of trustworthy security proofs. An answer to the first concern was given in the previous section; the second needs more explanation. If QKD is to replace conventional cryptography, we need to make sure that it really does provide better security. Here is an overview of some of the huge progress of the last few years to obtain decent security proof for QKD.

One of the most cited principles to explain the difference between quantum and standard cryptography is the non-cloning theorem. Contrary to the fact that classical communication can be copied faithfully, it is impossible to obtain a perfect copy of an arbitrary quantum state. Generally, trying to extract information from an arbitrary state will perturb it. If Eve tries to eavesdrop on judiciously chosen quantum states sent between Alice and Bob, then she will necessarily add noise to them so that she can be detected. This argument is quite loose and does not constitute a proof, but it explains intuitively how QKD could be secure and that the amount of information that Eve could have obtained must be an increasing function of the noise she induced. To be on the safe side, the worst case scenario is usually assumed when analyzing the security of QKD. Therefore, it is considered that all or most of the experimental noises are caused by Eve since she could have replaced imperfect devices by better ones. For example, it is generally assumed that Eve can replace a standard optic fiber by a loss and birefringence-free channel.

A great advancement in the development of security proof for QKD was the discovery of privacy amplification [19], a classical protocol which allows Alice and Bob to get rid of Eve's information on a string of bits they share, by sacrificing a number of bits that approximately equal the amount of information that Eve had on the string (see appendix B.1 for some more details about privacy amplification). It is therefore sufficient to prove security to be able to find (as a function of the noise in the channel) a bound on Eve's information on the raw key Alice and Bob obtained from the QKD protocol.

Without any assumption of restriction on Eve's access to Alice and Bob's laboratories, it is impossible to make a security proof. In general, it is supposed that Alice and Bob's labs are isolated, meaning that nothing else leaves the labs besides the state that Alice intended to send to Bob. However, this assumption is very strong, and in reality, Eve could make a *trojan-horse* attack by using some light or beam to scan Alice and Bob's lab. In theory, any standard cryptographic protocol could also be threatened by such an attack, but QKD is especially sensitive to it. It is believed that the amount of information that Eve can extract from such an attack can be limited by some common sense arguments [67], but more theoretical and experimental work will be needed on this important and sensitive subject. Another common assumption is that Alice and Bob's devices work in an ideal way. Considering imperfect devices is one of the greatest complications of security proof and many (including [73, 105, 2]) have put forth proposals that consider various general cases of imperfect devices.

Instead of considering an all powerful Eve, the problem of finding a security proof can be simplified by restricting operations that Eve could perform. We present different types of attacks considered on a QKD protocol where Alice sent to Bob a series of independent and identical mixed states[1], but their unconditional security proof are still open). For a *collective* attack, Eve uses one ancilla of some any dimension for each of the independent mixed state sent by Alice. On each of the sent states, Eve will independently apply the same operation on each state with his associated ancilla. However, at the end of the protocol, Eve can do any quantum measurement simultaneously on all her ancillas. The *individual* attack (also called *incoherent* attack) is similar to the collective case except that at the end of the protocol, Eve must do similar and independent measurement on the ancillas. Another interesting restriction is to suppose that the ancillas used by Eve are classical, meaning that she can't store quantum information and wait until the end of the protocol. There are good practical motivations to study weaker attacks, since it is believed that any real eavesdropper should be limited in the kind of quantum operations that he can performed or in the size of the quantum storage he has access to.

The *joint* attack is the strongest possible attack that Eve can perform, assuming she can't access or control Alice and Bob's devices. Eve can perform any quantum operation simultaneously on all the sent quantum states and on some extra ancillas. At the end of the QKD protocol, after Alice and Bob had declared everything they had to, she can make any quantum measurement on her ancillas. In spite of the controversy around the use of the following word, security proof against joint attack are often qualified as *unconditional* — a word borrowed from the computer science community. Most of this thesis will be focusing on unconditional security proofs assuming Alice and Bob's devices function perfectly and are isolated from Eve, but occasionally, there will be a need to refer to weaker kinds of attacks.

To claim that a proof is unconditionally secure, no matter what, is absurd. There is always some assumption behind an allegation of security. QKD assumes that the laws of quantum mechanics are universal i.e. Eve cannot do operations or possess states that are not allowed by quantum mechanics. Such an assumption might be surprising, especially after what happened to classical mechanics less than a century ago. We should expect the existence of some unknown physical theory that could threaten the security of QKD.

[1] Some proposed QKD protocols do not have this property (see [89] for an example

Interestingly, some work has been done to prove the security of the protocol despite the failure of the laws of quantum mechanics. These security proofs are based on some weaker conditions such as non-signaling or restriction on non-local correlations. Therefore, even if quantum mechanics is not the most general physical law, their might still be ways to do information-theoretic secure key distribution. The charm of quantum mechanics lies in its robustness. Many efforts have been made without success in demonstrating non-quantum effects. One important advantage of assuming quantum mechanics rather than some weaker condition is that, as far as we know, it provides higher secret key rates.

Resuming the two previous sections, the level of security of QKD protocols depends on how confident Alice and Bob are about Eve's limits. It also depends on how much Alice and Bob can trust their own devices not to reveal extra information to Eve by some side-channel that they were not aware of. Further improvement in the control and understanding of the inner mechanics of the devices used will increase trust in the QKD schemes. Analogously to the security of cryptography schemes based on algorithmic complexity, the security of QKD depends on the quantity of unsuccessful efforts made in breaking it. It is necessary to ensure that Eve cannot gain more information, than we thought she could, by trying to control or influence Alice and Bob's devices, or by using some special attack like the trojan-horse attack. Two important differences with contemporary cryptography is that QKD depends only on physical assumptions, and that its security is not threaten by future technological developments.[2]

1.5.2 Overview of Proofs of Security Against Joint Attack

When Bennett and Brassard proposed the first QKD protocol [17], called BB84, they did not have a security proof for their protocol. However, they believed that the security would follow from the uncertainty principle which implies that the value of a classical bit that is encoded in either of the two conjugated basis $\{|0\rangle, |1\rangle\}$ or $\{|+\rangle, |-\rangle\}$ where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ cannot be guessed with certainty. This is the same uncertainty principle that prevents the angular momentum of a spin$-\frac{1}{2}$ to be measured simultaneously in two orthogonal directions. More than ten years later, using a different uncertainty principle for entropy and mutual information, Mayers [130] gave the first rigorous unconditional security proof for BB84 (assuming Alice and Bob's devices work properly and are isolated from Eve). He observed that there was a tradeoff between the information that Bob can obtain from states that are encoded in a basis $\{|0\rangle, |1\rangle\}$ ($\{|+\rangle, |-\rangle\}$), and the information that Eve obtains from the states encoded in the conjugated basis $\{|+\rangle, |-\rangle\}$ ($\{|0\rangle, |1\rangle\}$). Soon after, a different proof [22] was published using a similar idea. The proofs were however very mathematically oriented. Remark that

[2] An interesting discussion about this subject can be found on Michael Nielson's blog.

Mayers' proof was greatly simplified in [49, 107, 104] by using arguments from security proof based on entanglement distillation.

Reducing BB84 to some entanglement distillation protocol (EDP) [20] for proving its security was proposed by [120, 173]. EDP is a protocol in which Alice and Bob, represented by the system A and B respectively, use local operations and classical communication (LOCC) to produce several approximate copies of the maximally entangled state

$$|\Phi_d\rangle_{AB} := \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}|k\rangle_A|k\rangle_B = \frac{1}{\sqrt{d}}\sum_{x=0}^{d-1}|\widetilde{x}\rangle_A|\widetilde{d}-\widetilde{x}\rangle_B \tag{1.3}$$

for dimensions d and where $\{|\widetilde{x}\rangle \mid 1 \leqslant x \leqslant d-1\}$ is a conjugated basis to $\{|z\rangle \mid 1 \leqslant z \leqslant d-1\}$ as defined in Appendix A.6. To understand intuitively why this reduction from QKD to EPD is sufficient to obtain security, let us assume that Alice and Bob share $|\Phi_d\rangle_{AB}$. If they measure respectively their register in the computational basis, they will obtain a uniformly random correlated dit. However, we can invoke an uncertainty principle [127] to show that Eve's cannot have any information about their outcome.

Consider any state ρ. Let $H(\Gamma)$ be the Shannon entropy of the outcomes of the measurement of the observable Γ on ρ. Then

$$H(Z) + H(X) \geqslant \log_2 d \tag{1.4}$$

where Z and X are non-degenerate observable with eigenvectors in the computational and a conjugated basis, respectively. Suppose that Bob, instead of measuring in the computational basis, measures in the conjugated basis, then he would be able to predict with certainty Alice's measurement outcome in the conjugated basis since $|\Phi_d\rangle_{AB}$. In that case, by the uncertainty principle, the outcome of the measurement of the observable Z on Alice's state will be uniformly-distributed. Since Eve cannot differentiate the case where Bob measured in the conjugated or in the computational basis, then she cannot have any information about the outcome of the measurement of the observable Z on Alice's state even if Bob measures in the computational basis.

After Shor and Preskill proved the security of BB84 using this reduction to a EDP, the method was generalized to other QKD protocols [118, 180, 27, 181, 150] and to include two-way communication [72].

Meanwhile, there are several other proposed methods for proving security of QKD against joint attack. A general proof for permutation invariant protocols was first given in [41]. In [153, 110], it was shown that pre-processing (a classical operation on the raw key that is performed before error correction) can improve the secret key-generation rate of QKD protocols. For a specific pre-processing procedure and one-way classical communication, Devetak-Winter found the "optimal" asymptotic secret key-generation rate against collective attack [50]. In his Ph.D. thesis [152], Renner proved the exponential quantum de Finetti theorem — which can be used to reduce joint attack to collective attack for permutation invariant protocol — and gave a security proof including finite size effect. It implies that the secret key generation rates obtained by Devetak-Winter for collective attack are also valid against joint attack.[3] The success of those security proofs went beyond the ones based on either the uncertainty principle or entanglement distillation. Principally, higher key generation rates were obtained, but the proofs were usually much more mathematical and were less intuitive. Some believed at the time that proofs based on those physical principles would not be able to capture all the subtleties required to optimize the secret key generation rate. As we will see in the next section, although not all QKD protocols can be interpreted as entanglement distillation, every QKD protocol must correspond to a *private states* distillation protocol.

1.5.3 Private States

By definition, a uniformly-distributed key of dimension d shared by Alice and Bob is independent of Eve if and only if it can be written as

$$\left(\frac{1}{d}\sum_{k=0}^{d-1}|k\rangle_A\langle k|\otimes|k\rangle_B\langle k|\right)\otimes\rho_E \tag{1.5}$$

for any ρ_E, where the index E means that the state is in Eve's possession. A *private state* is defined as any pure state of the form

$$|\psi\rangle_{ABSE}=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}|k\rangle_A|k\rangle_B V_S^{(k)}|\xi\rangle_{SE}=U_{ABS}|\Phi\rangle_{AB}|\xi\rangle_{SE}, \tag{1.6}$$

for arbitrary unitaries $V_S^{(k)}$ restricted to the system S and some state $|\xi\rangle_{SE}$. The new system S is out of reach of Eve and could be under the control (or not) of Alice and Bob. It is called the *shield*, since it protects the state shared by Alice and Bob from being correlated with Eve. $U_{ABS}:=\sum_{k=0}^{d-1}|k,k\rangle_{AB}\langle k,k|\otimes V^{(k)}$ is called the *twisting* operator, where $|k,k\rangle_{AB}:=|k\rangle_A|k\rangle_B$. In [86], it was shown that any uniformly-distributed secret key can be obtained from a private state, and any private state can give a uniformly-distributed secret key. We review this proof in Section 2.1. A private state is a generalization of the maximally entangled state given by Equation 1.3. Since the goal of QKD is to produce a uniformly-distributed secret key, we conclude that every QKD protocol must correspond to a private state distillation protocol where the private states have the form given by Equation 1.6.

[3] Remark that it was previously shown that collective attack was as strong as joint attack in the case of BB84 [22].

1.5.4 Definition of Security

Consider the interpretation of QKD as private state distillation. The resulting ideal state would be a perfect private state, but QKD protocol does not give such a perfect state. Instead, an approximation of a private state is obtained. The question is how good is that approximation. There are several ways of measuring the distance between quantum states, and the definition and strength of security will depend on the one that is chosen.

A good definition of security would allow the key generated by the QKD protocol to be used with any other protocol, such that the security of the combination of the protocols is guaranteed. In that case, the security is said to be a *universal composable* [21]. This turns out to be a very important issue, since many previous proofs for QKD were using a definition of security that was not a *universal composable*. Security was often claimed by showing that Eve's mutual information was exponentially small with respect to the size of the quantum communication. In [108], it was shown that even if Eve's mutual information was exponentially small, she could still get one bit of information in certain circumstances.

A good definition of security would be that Alice and Bob's quantum state would be indistinguishable from a private state with a probability that is exponentially close to $\frac{1}{2}$ in relation to some security parameter s. The parameter s should be easy to increase and could be, for instance, the size in qubits of quantum communication between Alice and Bob. A polynomial number of protocols respecting such a definition of security could then be combined without worrying about the security of the whole protocol.

According to a famous result called the Helstrom's Theorem [82], the probability of distinguishing between the two quantum states ρ_0 and ρ_1 is bounded by $\frac{1}{2} + \frac{1}{4}\mathrm{Tr}[|\rho_0 - \rho_1]$. It motivated the following definition of security[154, 21, 108]:

Definition: A random variable V on $\mathcal{V}$ is called an ϵ-secure key with respect to E if

$$\mathrm{Tr}[\rho_{VE} - \rho'_W \otimes \rho'_E] \leqslant 2\epsilon, \tag{1.7}$$

for some ρ'_E and where $\rho'_W := \sum_{v\in\mathcal{V}} \frac{1}{|\mathcal{V}|}|v\rangle\langle v|$ is the completely mixed state.

If Alice, Bob, the shield and Eve share a state ρ_{ABSE} such that

$$\mathrm{Tr}[|\rho_{ABSE} - \psi_{ABSE}|] \leqslant 2\epsilon \tag{1.8}$$

for some private state ψ_{ABES} given by Equation 1.6, then the key obtained by Alice and Bob by measuring their register in the computational basis will be ϵ-secure. It follows from the fact that the trace distance cannot increase under the action of a completely positive trace preserving (CPTP) map, like the partial trace.

Suppose n is the size in qubits of the quantum communication between Alice and Bob. We say that a secret key generation rate τ is achievable by a QKD protocol, if there exist a series of private states $|\psi^{(n)}\rangle_{ABSE}$ such that $\mathrm{Tr}[|\rho^{(n)}_{ABES} - |\psi^{(n)}\rangle_{ABSE}\langle\psi^{(n)}||] \leqslant 2\epsilon_n$, where $\epsilon_n \to 0$ and $0 \to \infty$. The rate of convergence is irrelevant in the asymptotical case, but is of prime importance when a finite key (i.e. finite n) is considered. As we will see, different methods for proving security of QKD provided different rates of convergence. Therefore, for finite QKD, some methods can be preferable to others. The rate τ of key generation is given by $\frac{1}{n}\log_2 [dim[\mathrm{Tr}_{BSE}[|\psi^{(n)}\rangle_{ABSE}\langle\psi^{n}|]] \to \tau$.

1.6 Contributions

Alternative ways of doing security proofs can be useful. For example, inspired by entanglement distillation, [169] showed that degenerate quantum error correction code could be used to improve the key generation rate. In [105], entropic uncertainty principle was used to prove the security of BB84 implemented with threshold detectors (i.e. detectors that do not differentiate the number of received photons). It is one of the only security proofs that take into account this feature. It is important since there are no efficient photon-counting detectors.

It is known that any secure QKD protocol can be understood as a private state distillation protocol - a straight generalization of the entanglement distillation protocol. For example, [151] described pre-processing as a private-state distillation protocol. We also saw that many proofs were based on an uncertainty principle. A legitimate question is whether it is restrictive to use the uncertainty principle to prove the security of QKD protocols in general. In Chapter 2, we provide an answer by proving that all private states can be interpreted as an instance of the uncertainty principle. Inspired by the work of [104], we give a new definition, related to the uncertainty principle, for a private state. It implies that some entropic uncertainty principle is behind any security proof for QKD. We use this observation to build from [173, 104] more intuitive proofs of various famous theorems regarding privacy amplification. Our method is based on a modified version of the HSW theorem [80, 83, 164] based on two-universal families of hashing functions.

In Chapter 3, we give a generic security proof for permutation-invariant QKD protocols by reducing it to a private state distillation protocol. Our proof of security is based on [150], but treats several new aspects. In [109], it was shown that a different *symmetrization*[4] allowed higher key generation rate in the case of a specific protocol (SARG04 [160]). We prove a similar result in a more general setting and explain how and why such improvement is observed.

Another improvement compared to other generic proofs [41, 153, 110, 152] is the use of Azuma's inequality [7] to simplify the analysis of estimation of parameters in certain situations [27]. In our proof, we emphasize the difference between reducing the QKD protocol to an entanglement distillation protocol or considering a more general private state distillation protocol, and give the different secret key generations rates that correspond to each of those scenarios. Sometimes, in order to achieve some secret key generation rates, it is enough to use the Azuma's inequality for parameter estimation. As far as we know, some higher rates can only be achieved by using the quantum de Finetti theorem. We explain some differences between those two cases.

[4] By symmetrization, we do not refer to extra operations that Alice and Bob must perform, but a trick to simplify the security analysis by exploiting some extra symmetries of the protocol.

Most of the proposed QKD protocols require that Alice and Bob share some reference frame. Without it, the error rates in Alice and Bob's shared raw keys will exceed the threshold under which the protocol is known to be secure. In Chapter 4, we explore the possibility of performing quantum key distribution without a reference frame by exploiting noiseless subsystems. We analyze in details the performance and the security of a four-photon QKD protocol [24]. For this task, the framework of the security proof given in Chapter 3 is not sufficient since Bob's measurement cannot be restricted to the encoded space of the state sent by Alice, as it is assumed there. Our results apply to other rotational invariant QKD protocols – specifically to the three-photon protocol proposed in [24] and the two-photon protocol proposed in [25].

Chapter 2

Private States and the Uncertainty Principle

2.1 Backgroung: Private States and Secret Key

We first review the statement that any uniform secret key can be obtained from a private state [86]. Suppose we have a secret ket (i.e. suppose that ρ_{ABE} is of the form of Equation 1.5). Consider a purification

$$|\psi\rangle_{ABES} = \frac{1}{\sqrt{d}} \sum_k |kk\rangle_{AB} |\xi^k\rangle_{ES}. \tag{2.1}$$

for some orthonormal states $|\xi^k\rangle_{ES}$. Using Theorem 11 and the fact that $\mathrm{Tr}_S[|\xi^k\rangle_{ES}\langle\xi^k|] = Tr_S[|\xi^{k'}\rangle_{ES}\langle\xi^{k'}|]$ for all k, k', then we obtain that there exist $|\xi\rangle_{ES}$ and unitaries $V_S^{(k)}$ such that

$$|\xi^k\rangle_{ES} = V_B^{(k)} |\xi\rangle_{ES} \tag{2.2}$$

for all k. Therefore, $|\psi\rangle_{ABES}$ is a private state as defined by Equation 1.6. Conversely, it is easy to see that if Alice and Bob measure a state of the form of Equation 1.6 in the computational basis, then after tracing the system S, they obtain a state ρ_{ABE} of the form of Equation 1.5. The same argument can be used to prove the following:

Theorem 1 *The pure state ψ_{ABES} is a private state iff there exist orthonormal bases $\{|k\rangle_A \mid 0 \leqslant k \leqslant d-1\}$ and $\{|k\rangle_B \mid 0 \leqslant k \leqslant d-1\}$ such that*

$$\begin{aligned} {}_{AB}\langle k, k'|\mathrm{Tr}_{ES}[\psi_{ABES}]|k, k'\rangle_{AB} &= \frac{1}{d}\delta_{k,k'} && \textit{(uniform)} \\ \mathrm{Tr}_{ABS}[\psi_{ABES}] &=_{AB} \langle k, k|\mathrm{Tr}_S[\psi_{ABES}]|k, k\rangle_{AB} && \textit{(independence)} \end{aligned} \tag{2.3}$$

for all $1 \leqslant k, k' \leqslant d$, where $|k, k'\rangle_{AB} = |k\rangle_A |k'\rangle_B$.

2.2 Dual Interpretation of Private State

Inspired by [104], we propose in [148] an alternative definition for a private state. Keeping in mind that a private state is a state of the form given by Equation 1.6, we define

$$|\widetilde{x}\rangle = \frac{Z^x}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle \tag{2.4}$$

where Z is the generalized Pauli operator in d-dimension (i.e. $Z|k\rangle = e^{\frac{2k\pi i}{d}}|k\rangle$) [71]. A different characterization of private states is obtained from considering a hypothetical measurement by Alice in the $\widetilde{x}$-basis. Then one has

Theorem 2 *A pure state $|\psi\rangle_{ABSE}$ is a private state if and only if*

$$_{AB}\langle k, k'|\mathrm{Tr}_{ES}[|\psi\rangle_{ABSE}\langle\psi|]|k, k'\rangle_{AB} = \frac{1}{d}\,\delta_{k,k'}, \tag{2.5}$$

for all k and k', and

$$\sigma_{BS}^{x'}\sigma_{BS}^{x} = 0 \tag{2.6}$$

for all $x' \neq x$, where $\sigma_{BS}^{x} = d\,{}_A\langle\widetilde{x}|\mathrm{Tr}_E[|\psi\rangle_{ABSE}\langle\psi|]|\widetilde{x}\rangle_A$.

Proof.

Suppose that $|\psi\rangle_{ABSE}$ is a private state. Condition 2.5 follows immediately from the definition of a private state (Equation 1.6). We now try to derive Condition 2.6. Consider

$$\begin{aligned}
{}_A\langle\widetilde{x}|\psi\rangle_{ABES} &= \frac{1}{d}\sum_{k'=0}^{d-1} {}_A\langle k'|Z_A^{\dagger x}\sum_{k}|k,k\rangle_{AB}V_S^{(k)}|\xi\rangle_{SE} &(2.7)\\
&= \frac{1}{d}\sum_{k'=0}^{d-1} {}_A\langle k'|Z_B^{\dagger x}\sum_{k=0}^{d-1}|k,k\rangle_{AB}V_S^{(k)}|\xi\rangle_{SE} &(2.8)\\
&= \frac{1}{d}Z_B^{\dagger x}\sum_{k=0}^{d-1}|k\rangle_B V_S^{(k)}|\xi\rangle_{SE}. &(2.9)
\end{aligned}$$

So

$$\begin{aligned}
{}_A\langle\widetilde{x}|\psi\rangle_{ABES} &= \frac{1}{d}Z_B^{\dagger x}U_{ABS}\sum_{k=0}^{d-1}|k\rangle_B|\xi\rangle_{SE} &(2.10)\\
&= \frac{1}{d}U_{ABS}Z_B^{\dagger x}\sum_{k=0}^{d-1}|k\rangle_B|\xi\rangle_{SE} &(2.11)\\
&= \frac{1}{\sqrt{d}}U_{ABS}|\widetilde{d-x}\rangle_B|\xi\rangle_{SE} &(2.12)
\end{aligned}$$

where $U_{ABS} = \sum_{k=0}^{d-1} |k,k\rangle_{AB}\langle k,k| \otimes V_S^{(k)}$. The states $|\widetilde{x}\rangle_B$ are understood to be

$$|\widetilde{x} \bmod |B|\rangle_B.$$

In this case, $|B| = d$. Let $\xi_S := \mathrm{Tr}_E[|\xi\rangle_{SE}\langle\xi|]$, so

$$\begin{aligned} \sigma^x_{BS} &= d\,{}_A\langle\widetilde{x}|\mathrm{Tr}_E[\psi_{ABES}]|\widetilde{x}\rangle_A & (2.13)\\ &= U_{ABS}(|\widetilde{d}-\widetilde{x}\rangle_B\langle\widetilde{d}-\widetilde{x}| \otimes \xi_S)U^\dagger_{ABS} & (2.14) \end{aligned}$$

Therefore $\sigma^{x'}_{BS}\sigma^x_{BS} = 0$ for all $x' \neq x$, which is Condition 2.6. Now, we prove the converse by invoking the following uncertainty principle [127]. Consider any state ρ. Let $H(\Gamma)$ be the Shannon entropy of the outcomes of the measurement of the observable Γ on ρ. Then

$$H(Z) + H(X) \geqslant \log_2 d \tag{2.15}$$

where X is an non-degenerate observable with eigenvectors given by the $|\widetilde{x}\rangle$ for all x. Suppose that Alice, Bob and the shield share some state ρ_{ABS} for which $|A| = d$, and that there exists a measurement $\mathcal{M}^X_{BS}$ on ρ_{BS} that determines with certainty the outcome of the measurement of the observable X on the state ρ_A. From the above uncertainty relation and conditional on performing $\mathcal{M}^X_{BS}$, the outcome of a measurement of Z on ρ_A must be uniformly-distributed (i.e. $H(Z) = d$) from Eve's perspective. Therefore, $\mathcal{M}^X_{BS}$ had be performed, then Eve cannot obtain any information about the outcome of Alice's Z-measurement. Since Eve cannot distinguish between the cases where $\mathcal{M}^X_{BS}$ has been performed or not, Eve will still be unable to gain any information about the outcome of Alice's Z-measurement even if Bob makes the measurement on ρ_B that gives him full information about Alice's outcome in the Z basis. Consequently, Eve's state cannot dependent on $|k,k\rangle_{AB}$ and we obtain Condition 2.4. ■

Theorem 2 is a generalization of a similar theorem found in [148]. The difference is that in [148], only the $d = 2$ case was considered. Theorem 2 implies that a pure state is a private state if and only if Alice and Bob's measurements in the computational basis are perfectly correlated (and otherwise uniformly-distributed), and if Bob, with the use of the shield (S), is able to guess without errors the outcome of Alice's hypothetical measurement in the $\widetilde{x}$-basis. A physical interpretation of the previous statement is that if Bob has the capacity to obtain full information about the outcome of Alice's measurement in a complementary basis (i.e. the $\widetilde{x}$-basis), then Eve does not have any information about Alice's measurement in the computational basis. As we saw in the proof of Theorem 2, this trade-off between Bob and Eve's information is a manifestation of the uncertainty principle given in [127].[1].

[1] We could also have shown that it follows from an uncertainty principle relating mutual information [77].

2.3 Private State Distillation and Privacy Amplification

2.3.1 With Symmetrization

Let σ_{ABE} be a state shared by Alice, Bob and Eve. What is the maximal amount of ϵ-secure key that can be generated from σ_{ABE} or equivalently, what is the largest state that can be created from σ_{ABE} using $LOCC$ such that it is 2ϵ close in trace distance to a private state? The answer is not known in general, but there are some interesting related results that can be very useful in analyzing the security of standard QKD protocols.

Suppose that Alice and Bob's systems are decomposable in $n+m$ equivalent subsystems $A_1, ..., A_{n+m}$ and $B_1, ..., B_{n+m}$ of dimension d, and that σ_{ABE} is invariant under any joint permutation of those subsystems. By the quantum de Finetti theorem ([152]), the state of the n first subsystems is close to a form $(\sigma_{A_1B_1E_1})^{\otimes n}$, for some state $\sigma_{A_1B_1E_1}$ and where m can be very small compared to n. Here, we are neglecting the extra complications encountered by using the quantum de Finetti theorem, and leave them for a future publication [149].

Suppose that Alice's state is classical and uniformly-distributed:

$$\sigma_{A_1B_1E_1} = \frac{1}{d}\sum_{k_1=d}^{d-1} |k_1\rangle_{A_1}\langle k_1| \otimes |\phi^{(k_1)}\rangle_{B_1E_1}\langle\phi^{(k)}|$$

for some states $|\phi^{k_1}\rangle_{B_1E_1}$. Devetak and Winter [50] proved the following theorem:

Theorem 3 ([50]) *Suppose that Alice, Bob and Eve share the state*

$$(\frac{1}{d}\sum_{k_1=d}^{d-1} |k_1\rangle_{A_1}\langle k_1| \otimes |\phi^{(k_1)}\rangle_{B_1E_1}\langle\phi^{(k_1)}|)^{\otimes n}, \tag{2.16}$$

then there exists a one-way classical communication scheme to extract

$$n[I(K_1{:}B_1) - I(K_1{:}E_1)] \tag{2.17}$$

bits that are ϵ_n-secure where $\epsilon_n \rightarrow 0$ exponentially for $n \rightarrow \infty$ and K_1 is the classical variable associated to Alice's key.

In the above theorem, the function I is the quantum mutual information defined in appendix A.4. $nI(K_1{:}B_1)$ represents the maximal rate of error-free key obtained using one-way error correction. However, to achieve this rate, sometimes a complicated measurement like the one described in the proof of the HSW theorem (see appendix B.2) must be performed. For simplicity, suppose that Bob performs a measurement in a fixed basis. In that case, the rate of achievable error-free key is $(d - H(e_{X^{j_1}})) \leqslant I(K_1 : B_1)$ where

$e_{X^{j_1}}$ is the error rates in Bob's measurement. After information reconciliation is done, the eavesdropper might have information about the key shared by Alice and Bob. This information is bounded by $nI(K_1{:}E_1)$. Using privacy amplification (see appendix B.1), Alice and Bob can reduce this information to zero by sacrificing a number of bits proportional to the amount of information that Eve has on the key.

Instead of applying the quantum de Finetti theorem before information reconciliation, suppose that we apply it after. Alice and Bob will share a state that is close (how close depends on the efficiency of the error correction) in trace distance to a state of the form

$$(\frac{1}{d}\sum_{k_1=0}^{d-1}|k_1,k_1\rangle_{A_1B_1}\langle k_1,k_1|\otimes\sigma_{E_1}^{(k_1)})^{\otimes n}. \tag{2.18}$$

It follows from [154, 50] that

Theorem 4 *There exists a privacy amplification scheme to extract* $n[\log_2 d - I(K_1{:}E_1)]$ *secret bits from*

$$(\frac{1}{d}\sum_{k_1=0}^{d-1}|k_1,k_1\rangle_{A_1B_1}\langle k_1,k_1|\otimes\sigma_{E_1}^{(k_1)})^{\otimes n}, \tag{2.19}$$

for $n\rightarrow\infty$*. Moreover, this is the maximum possible rate.*

In Appendix B.3, we prove an alternative version of this theorem:

Theorem 5 *There exists a privacy amplification scheme to extract* $n[I(X_1{:}B_1S_1)]$ *secret bits from*

$$(\frac{1}{d}\sum_{k_1=0}^{d-1}|k_1,k_1\rangle_{A_1B_1}\langle k_1,k_1|\otimes\sigma_{E_1}^{(k_1)})^{\otimes n}, \tag{2.20}$$

for $n\rightarrow\infty$, $X := \sum_{x_1=0}^{d-1} e^{\frac{-2\pi i x_1}{d}}|\widetilde{x_1}\rangle\langle\widetilde{x_1}|$*, and* $|\widetilde{x_1}\rangle$ *is defined by Equation 2.4. The shield* S *is the extra system required to purify the above state. Moreover, this is the maximum possible rate.*

Our proof in Appendix B.3 is based on a modified version of the HSW-theorem given in Appendix B.2. Even if it does not explicitly refer to the uncertainty principle as in Section 2.2, it is intimately related to it. From our dual definition of a private state, we found that a perfect private state is a state for which there exist a measurement on BS that can predict the measurement on A in the conjugated and in the computational bases. From this observation, it is obvious to conjuncture that an approximate private state is a state for which there exists a measurement on BS that can predict with high probability the measurement on A in the conjugated and in computation bases. We prove this statement in Appendix B.3 and use it to prove Theorem 5.

Not very surprisingly, Theorem 4 and 5 are equivalent to each other. To see this, consider any state of the form $\frac{1}{d}\sum_{k_1=0}^{d-1}|k_1,k_1\rangle_{A_1B_1}\langle k_1,k_1|\otimes\sigma_{E_1}^{(k_1)}$. Take one of its purifications:

$$|\psi\rangle_{A_1B_1S_1E_1} = \frac{1}{\sqrt{d}}\sum_{k_1=0}^{d-1}|k_1,k_1\rangle_{A_1B_1}|\phi^{(k_1)}\rangle_{S_1E_1} \tag{2.21}$$

for some pure states $|\phi^{(k)}\rangle_{S_1E_1}$. We can write

$$|\psi\rangle_{A_1B_1S_1E_1} = \frac{1}{d}\sum_{x_1=0}^{d-1}|\widetilde{x_1}\rangle_{A_1}Z_{B_1}^{-x_1}\sum_{k_1=0}^{d-1}|k_1\rangle_{B_1}|\phi^{(k_1)}\rangle_{S_1E_1}. \tag{2.22}$$

Define

$$\rho_{A_1B_1S_1E_1} := |\psi\rangle_{A_1B_1S_1E_1}\langle\psi|, \tag{2.23}$$

$$\bar{\sigma}_{S_1B_1} := \mathrm{Tr}_{E_1}[\frac{1}{d}\sum_{k_1,k_2=0}^{d-1}|k_1\rangle_{B_1}|\phi^{(k_1)}\rangle_{S_1E_1}\langle\phi^{(k_2)}|\ _{B_1}\langle k_2|], \tag{2.24}$$

$$\sigma_{S_1B_1}^{(x_1)} := Z_{B_1}^{-x_1}\bar{\sigma}_{B_1S_1}Z_{B_1}^{x_1}, \tag{2.25}$$

and

$$\rho_{E_1}^{(k_1)} := \mathrm{Tr}_{S_1}[|\phi^{(k_1)}\rangle_{S_1E_1}\langle\phi^{(k_1)}|]. \tag{2.26}$$

Since all the $\sigma_{B_1S_1}^{(x_1)}$ are related by unitaries to $\bar{\sigma}_{S_1B_1}$, they share the same eigenvalues which implies that $S(\sigma_{B_1S_1}^{(x_1)}) = S(\bar{\sigma}_{S_1B_1})$ for all x_1. One of the consequences of the Schmidt Decomposition is that ρ_{E_1} and $\bar{\sigma}_{S_1B_1}$ have the same eigenvalues. So $S(\rho_{B_1S_1}) = S(\rho_{A_1E_1})$. Therefore,

$$\begin{aligned}
I(X_1 : B_1S_1) &= S(\rho_{B_1S_1}) - \frac{1}{d}\sum_{x_1=0}^{d-1}S(\sigma_{B_1S_1}^{(x_1)}) && (2.27)\\
&= S(\rho_{A_1E_1}) - S(\bar{\sigma}_{S_1B_1}) && (2.28)\\
&= S(\frac{1}{d}\sum_{k_1=0}^{d-1}|k_1\rangle_A\langle k_1|\otimes\rho_{E_1}^{(k_1)}) - S(\rho_{E_1}) && (2.29)\\
&= \log_2 d + \frac{1}{d}\sum_{k_1=0}^{d-1}S(\rho_{E_1}^{(k_1)}) - S(\rho_{E_1}) && (2.30)\\
&= \log_2 d - I(K_1 : E_1). && (2.31)
\end{aligned}$$

Consequently, Theorem 4 and 5 are equivalent.

2.3.2 Without Symmetrization

In the case where Alice, Bob and Eve don't share a state of the symmetric form given by Equation 2.18, we have the following theorem:

Theorem 6 *There exists a privacy amplification scheme to extract a ($2^{-\frac{\eta}{2}+\frac{1}{2}}$)-secure key of size*

$$n\log_2 d - \log_2(||\frac{1}{d^n}\sum_{x=1}^{d^n}\sigma_{BS}^{(x)}||_\infty) - \log_2(\sum_{x=1}^{d^n} Rank[\sigma_{BS}^{(x)}]) - \eta \tag{2.32}$$

secret bits from

$$\frac{1}{d^n}\sum_{k=1}^{d^n}|k,k\rangle_{AB}\langle k,k|\otimes\rho_E^{(k)}, \tag{2.33}$$

where $\frac{\eta}{\log_2 d}$ *is an integer,* $\sigma_{BS}^{(x)} = d\,{}_A\langle\widetilde{x}|\mathrm{Tr}_E[|\psi\rangle_{ABSE}\langle\psi|]|\widetilde{x}\rangle_A$*,* $|\psi\rangle_{ABSE}$ *is a purification of the above state and* $||\cdot||_\infty$ *is the infinite norm.*

This theorem can be showed by applying the comment at the end of Appendix B.2 about the $n=1$ case to the proof given in appendix B.3 (using Equation B.43).

We now prove an well-known version of Theorem 6. Rewrite

$$|\psi\rangle_{ABSE} = \frac{1}{\sqrt{d^n}}\sum_{k=1}^{d^n}|k,k\rangle_{AB}|\phi^{(k)}\rangle_{SE} \tag{2.34}$$

for some pure states $|\phi^{(k)}\rangle_{SE}$. So

$$|\psi\rangle_{ABSE} = \frac{1}{d^n}\sum_{x=1}^{d^n}|\widetilde{x}\rangle_A Z_B^{-x}\sum_{k=1}^{d^n}|k\rangle_B|\phi^{(k)}\rangle_{SE}. \tag{2.35}$$

Define

$$\rho_{ABSE} := |\psi\rangle_{ABSE}\langle\psi|, \tag{2.36}$$

$$\bar{\sigma}_{SB} := \mathrm{Tr}_E[\frac{1}{d^n}\sum_{k,k'=1}^{d^n}|k\rangle_B|\phi^{(k)}\rangle_{SE}\langle\phi^{(k')}|\ {}_B\langle k'|], \tag{2.37}$$

and

$$\rho_E^{(k)} := \mathrm{Tr}_S[|\phi^{(k)}\rangle_{SE}\langle\phi^{(k)}|]. \tag{2.38}$$

So

$$
\begin{aligned}
\mathrm{Rank}[\sigma_{BS}^{(x)}] &= \mathrm{Rank}[Z_{B_1}^{-x_1}\bar{\sigma}_{BS}Z_B^x] && (2.39)\\
&= \mathrm{Rank}[\bar{\sigma}_{BS}] && (2.40)\\
&= \mathrm{Rank}[\rho_E] && (2.41)\\
& && (2.42)
\end{aligned}
$$

where in the last step, we implicitly used the Schmidt decomposition. The Schmidt decomposition also imply that

$$
\begin{aligned}
||\frac{1}{d^n}\sum_{x=1}^{d^n}\sigma_{BS}^{(x)}||_\infty &= ||\frac{1}{d^n}\sum_{k=1}^{d^n}|k\rangle_A\langle k|\otimes\rho_E^{(k)}||_\infty && (2.43)\\
& && (2.44)
\end{aligned}
$$

Therefore, we have provided an alternative proof of:

Theorem 7 *: There exists a privacy amplification scheme to extract a ($2^{-\frac{\eta}{2}+\frac{1}{2}}$)-secure key of size*

$$
S_\infty(\frac{1}{d^n}\sum_{k=1}^{d^n}|k\rangle_A\langle k|\otimes\rho_E^{(k)}) - S_0(\rho_E) - \eta \tag{2.45}
$$

secret bits from

$$
\frac{1}{d^n}\sum_{k=1}^{d^n}|k,k\rangle_{AB}\langle k,k|\otimes\rho_E^{(k)}, \tag{2.46}
$$

where $\frac{\eta}{\log_2 d}$ is an integer.

where $S_0(\rho) := \log_2(\mathrm{Rank}[\rho])$ and $S_\infty(\rho) := -\log_2||\rho||_\infty$. A better version of Theorem 7 was proved in [154] (i.e. the key is ($2^{-\frac{\eta}{2}-2}$)-secure instead of ($2^{-\frac{\eta}{2}+\frac{1}{2}}$)-secure and S_2 replace S_∞ in the key generation rate). The reason why someone might be interested by such results is that it might be useful to analyze the security of QKD protocols that do not have permutation symmetry. In [152] a more optimal formula in term of smooth min-entropy for privacy amplification was derived.

2.4 Summary of the Chapter

2.4.1 Summary

In this Chapter, we have given a new interpretation of private states highlighting a fundamental relation between any security proofs for QKD and the entropic uncertainty principle. We use the framework of the security proof by Koashi [104] — which combines and generalizes the proofs of security of Mayers and Shor-Preskill [130, 173] based on the uncertainty principle and entanglement distillation, respectively — to prove some famous theorems about privacy amplification (Theorem 4 and 7). One of our important contributions was to include the concept of Shield in the framework of [104]. Another was to prove and exploit a generalized version of the HSW theorem involving two-universal family of hashing function. Contrarily to [104], our analysis in Chapter 2 is not restricted to a specific QKD protocol and the dimension of the state considered can be arbitrary — i.e. not only a power of 2.

Neglecting information-reconciliation, we gave a proof of Devetak-Winter secret key generation rate for n copies of a classical-classical-quantum states [50] (we assume that Alice and Bob start with an identical and uniformly distributed key). The main difference between the two proofs is that we used a random linear code instead of random codewords, which simplifies certain steps. As a direct consequence of using our method, we obtain a weaker version of the theorem about privacy amplification given by Renner and Koenig [154]. Remark that we gave the secret key generation rates in function of the state of the Shield.

2.4.2 Connections to Other Results

The main results of this Chapter were initially given in [148]. Soon after, Koashi submitted a paper about the role of complementarity in QKD [106]. By introducing an extra channel, he showed that if there exist two measurements on the system B that could predict the outcome of Alice's measurement in the computational and conjugated bases with some small error probability, then the state shared by Alice and Bob would be very close to a secret key (if the channel is quantum) or a maximally entangled state (if the channel is classical). Noting that the extra channel is directly related to the concept of Shield, Koashi's theorem could help us to considerably simplify our proofs of Theorem 4 and 7. Furthermore, instead of only considering privacy amplification, we can use his theorem to include information-reconciliation and prove the complete theorem of Devetak-Winter about secret key generation rate using a symmetrized state. These results will be explained in detail in a future publication [149].

Chapter 3

Generic Proof of Security for Permutation Invariant QKD Protocols

In Chapter 2, we showed how to extract a secret key from a state ρ_{ABE} shared by Alice, Bob and Eve. To obtain a secret key, we assume implicitly that Alice and Bob knew the exact form of ρ_{ABE}. We are now interested in knowing how Alice and Bob can learn the form of ρ_{ABE}. We consider permutation-invariant QKD protocols that have the nice feature of being easy to implement experimentally. Our final result is a complete proof against joint attack for a large number of QKD protocols that are invariant under permutations.

We first review some important QKD protocols in Section 3.1. In Section 3.2, we present a general prepared-and-measured QKD protocol that we want to prove as secure. This protocol includes the examples given in Section 3.1. In Section 3.3, we introduce an entanglement-based QKD protocol, which our general prepared-and-measured QKD scheme can we be reduced to. We also show how symmetry can be exploited to simplify the form of the state shared by Alice, Bob and Eve. In Section 3.4, we cover information reconciliation. In particular, we explain how dit error correction can be used to build a Shield that can improve the secure key generation rate of the QKD protocol. In Section 3.5, we describes two methods for parameter's estimation and we explain the difference in the rates obtained. One of the methods is based on the Azuma's inequality, which achieves a higher security level for the same key length, and the other on the quantum de Finetti theorem, which achieve higher rate.[1]

[1] In this thesis, we do not consider all the details related to reducing joint attack to collective attack by using the quantum de Finetti theorem. Instead of applying the quantum de Finetti theorem, we simply assume that the state is in the symmetrized form (i.e. $\rho_{ABE}^{\otimes n}$). When this happens, we effectively give a

3.1 Examples of Prepared-and-Measured QKD Protocols

3.1.1 The BB84 protocol and the Six-State Protocol

The first QKD protocol was constructed by Bennett and Brassard [17] (BB84), and used four different single-qubit quantum states. Alice picks at random a n-bits string. For each bit, she randomly selects a basis among the two conjugates bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, encodes her bit accordingly, and sends it to Bob. We use the notation $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The qubits could be encoded in photon polarization, $|0\rangle$ and $|1\rangle$ could correspond to horizontal and vertical polarization states, and $|\pm\rangle$ to polarization at ±45 degree angle. For each qubit received, Bob randomly chooses one of the two conjugate bases, in which he measures the qubit. Then, Alice announces the encodings publicly, and Alice and Bob discard the qubits which were prepared and measured in different bases. The leftover bits are used to build the secret key. If the eavesdropper (Eve) tries to get any information about the quantum states sent by Alice, she will inevitably perturb the states with some significant probability and induce errors in the key. Alice and Bob can use standard classical error correction to get rid of the errors in the key. This step is called information reconciliation. Then, Alice and Bob use privacy amplification, as described in Appendix B.1 to erase Eve's information about the key.

The Six-State protocol [33] is identical to BB84, except that Alice and Bob choose randomly between three mutually unbiased bases instead of two, and keep the result only if they correspond to the same basis.

3.1.2 Bennett 1992

The B92 QKD protocol [15] requires only two quantum states. Alice encodes her bits using any two non-orthogonal states $|\psi_1\rangle$ and $|\psi_2\rangle$. Bob makes a positive operator-valued measurement (POVM) of the qubit described by

$$\{\alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1|, \alpha|\overline{\psi}_2\rangle\langle\overline{\psi}_2|, \mathbb{1} - \alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1| - \alpha|\overline{\psi}_2\rangle\langle\overline{\psi}_2|\} \tag{3.1}$$

where $|\overline{\psi}_1\rangle$ and $|\overline{\psi}_2\rangle$ are orthogonal to $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively, and $\alpha = \frac{1}{1+|\langle\psi_1|\psi_2\rangle|}$. The outcome $|\overline{\psi}_1\rangle$ indicates to Bob that the state sent by Alice was $|\psi_2\rangle$, the outcome $|\overline{\psi}_2\rangle$ indicates $|\psi_1\rangle$, and the other outcome is inconclusive so the associated key bits are discarded. The value of α is chosen to minimize the probability of getting an inconclusive

security proof against collective attack instead of joint attack. The additional details required to consider the most general case are quite straightforward, but cumbersome. We will add them in a future publication including all the results of this Chapter. Remark that when the Azuma's inequality is used instead of the quantum de Finetti theorem, the security analysis given here is against joint attack.

result. As in BB84, any attempt by Eve to gain information about the encoded key bit will perturb the quantum state and induce errors with some finite probability [18].

The unconditional security of B92 was demonstrated in [180]. The main inconveniece of B92, as compared to BB84, is its low tolerance to noise. To establish a secure key, Alice and Bob must evaluate the rate of bit and phase error induced by Eve. This turns out to be easier to perform with the highly symmetric setting of BB84 than with the asymmetric B92 protocol. Another reason for B92's low tolerance to error is the possibility for Eve to perform unambiguous discrimination of the encoded key bit [37], which is impossible in the BB84 setting since Alice's encoding is chosen at random. Thus, Eve can replace the communication channel by a perfect noiseless channel, perform the measurement Equation 3.1 on a random subset of Alice's encoded bits, discard the inconclusive results and retransmit the conclusive qubits to Bob. The overall effect of this intervention would simulate a lossy channel, so by tuning her measurement rate, Eve could be confounded with natural noise. This attack considerably lowers the security threshold of B92 [182] when there are photon losses.

3.1.3 Phoenix, Barnett, and Chefles 2000

A modification of the B92 protocol was proposed by Phoenix, Barnett, and Chefles (PBC00) to overcome the low qubit loss tolerance and enhance the phase error rate estimation [145]. The solution was to reduce the inherent asymmetries of B92 by adding a third state $|\psi_3\rangle$. In this new three state QKD protocol, Alice first randomly chooses two out of the three states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ and encodes her key bit on those states as in B92. When Bob receives a qubit, he performs the POVM

$$\{\alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1|, \beta|\overline{\psi}_2\rangle\langle\overline{\psi}_2|, \gamma|\overline{\psi}_3\rangle\langle\overline{\psi}_3|, \Gamma\} \tag{3.2}$$

where α, β and γ are real positive numbers chosen to optimize the POVM performance and $\Gamma = \mathbb{1} - \alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1| - \beta|\overline{\psi}_2\rangle\langle\overline{\psi}_2| - \gamma|\overline{\psi}_3\rangle\langle\overline{\psi}_3|$. Alice then announces the encodings publicly. Given this information, some of Bob's measurement outcomes will be conclusive, yielding a key bit, and some will be inconclusive and get discarded.

An interesting version of this protocol occurs when the three states $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\psi_3\rangle$ form an equilateral triangle in the X-Z plane of the Bloch sphere. This is the symmetric trine spherical code. In this case, Bob's POVM reduces to

$$\mathcal{M}_{\text{PBC00}} = \{\frac{2}{3}|\overline{\psi}_1\rangle\langle\overline{\psi}_1|, \frac{2}{3}|\overline{\psi}_2\rangle\langle\overline{\psi}_2|, \frac{2}{3}|\overline{\psi}_3\rangle\langle\overline{\psi}_3|\} \tag{3.3}$$

and $\Gamma = 0$. Henceforth, "PBC00" will refer to this specific version of the three state protocol. In [147], Renes demonstrated how the rate of conclusive results can be used to estimate the error rates so that no test bits are required ($e_{\text{bit}} = \frac{1-2I}{1-I}$, $e_{\text{phase}} = \frac{5}{4}e_{\text{bit}}$ where

I is the rate of inconclusive results among the received qubits). This technique could also apply to our robust four-photon protocol, but for simplicity, we will assume that a random sample of test bits is used to estimate the error rates.

In [27], we give a security proof of PBC00 independently of the channel's qubit loss rate. To compare the security level of BB84, B92, and BPC00 in the absence of qubit loss, we can model Eve's intervention by a depolarizing channel that transforms ρ into $(1-p)\rho + \frac{p}{3}\sum_{a=x,y,z} \sigma_a \rho \sigma_a$, the Pauli operators are defined as

$$\sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Excluding pre-processing and considering one-way entanglement distillation, the protocols B92, PBC00, B84 and the six-state protocol are secure up to $p \approx 3.4\%$ [180], $p \approx 11.6\%$ [27], $p \approx 16.5\%$ [173] and $p \approx 19.1\%$ [118], respectively.

3.2 A General Prepared-and-Measured QKD protocol

We will prove the security of the following prepared-and-measured[2] protocol:

1. Consider some Hilbert space $\mathcal{H}_d$ of dimension d. Alice and Bob pre-agree on d linearly independent states $|\psi_i\rangle$ for $0 \leqslant i \leqslant d-1$ and a series of unitary transformations, $U^{(i)} : \mathcal{H}_d \to \mathcal{H}_d$ for $1 \leqslant i \leqslant \ell$. We assume that the states are related by $|\psi_i\rangle = (X^i)^\dagger|\psi_0\rangle$, where X is the generalized Pauli-operator defined in Appendix A.6, and X^i is the i^{th} power of X. Each unitary transformation is associated to a change of basis.

2. Alice randomly picks one of the d states $|\psi_i\rangle$ and one of the transformations $U^{(j_a)}$ and sends $U^{(j_a)}|\psi_i\rangle$ to Bob. She repeats those operations n' times.

3. Upon reception of each of the n' quantum states, Bob picks at random one of the transformation $U^{(j_b)}$ and measures the state with the POVM

   $$\{\alpha U^{(j_b)}|\psi_1^\perp\rangle\langle\psi_1^\perp|U^{(j_b)\dagger}, ..., \alpha U^{(j_b)}|\psi_d^\perp\rangle\langle\psi_d^\perp|U^{(j_b)\dagger}, \mathbb{1} - \sum_{i=0}^{d-1} \alpha U^{(j_b)}|\psi_i^\perp\rangle\langle\psi_i^\perp|U^{(j_b)\dagger}\}$$

   where α is the maximal positive number such that

   $$\sum_{i=0}^{d-1} \alpha U^{(j_b)}|\psi_i^\perp\rangle\langle\psi_j^\perp|U^{(j_b)\dagger} \leqslant \mathbb{1}. \tag{3.4}$$

   The state $|\psi_i^\perp\rangle$ is defined as the unique state in $\mathcal{H}_d$ that is orthogonal to $|\psi_j\rangle$ for all $j \neq i$.

4. Alice and Bob publicly declare which transformation they used, and discard the results associated to cases where they used different bases. They also discard results that are inconclusive (i.e when Bob measures $\mathbb{1} - \sum_{i=0}^{d-1} \alpha U^{(j_b)}|\psi_i^\perp\rangle\langle\psi_i^\perp|U^{(j_b)\dagger}$). We define n as the number of not discarded results.

This protocol includes BB84, B92, the six-state protocol, PBC00 and a many more protocols. In BB84, Alice chose between two states $|0\rangle$ and $|1\rangle$, and between two transformations $\mathbb{1}$ and H. In PBC00, Alice chose between two states $\frac{1}{2}|+\rangle + \frac{\sqrt{3}}{2}|-\rangle$ and $\frac{1}{2}|+\rangle - \frac{\sqrt{3}}{2}|-\rangle$, and between three rotations that correspond to 0, $\frac{2}{3}\pi$ and $\frac{4}{3}\pi$ rotations around the Y-axis (see [27] for details).

[2] By definition, in a prepared-and-measured QKD protocol, Alice sends to Bob n' quantum states, where the i^{th} quantum state $|\psi_i\rangle$ is randomly chosen from ℓ different pure states that follow a probability distribution p_j (i.e. $\sum_{j=1}^{\ell} p_j = 1$). Upon reception of the quantum states, Bob measures them immediately.

We could have generalized the above protocol by supposing that Alice could have different sets of states — the states being related by a power of X within each set — and that for each of those sets, Alice and Bob could associate a different set of random transformations which do not have to be uniformly distributed. Alice and Bob could then have the choice of extracting the key from only one (or some) set of states and use the others to probe the channel. The protocol studied in [64] is an example where there are two different sets of states $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle\}$, and that one of the sets is used to probe the channel. Our proof method could be extended to include these cases.

3.3 State Preparation

3.3.1 Entanglement based QKD protocol

Our first step is to give an entanglement based QKD protocol to which the protocol given in subsection 3.2 can be reduced.[3] Actually, we will consider a more general entanglement based QKD protocol that could reduce to other interesting prepared-and-measured schemes. Define the *filtering* operator

$$F = \beta \sum_{i=0}^{d-1} |i\rangle\langle\psi_i| \tag{3.5}$$

such that β is the maximal positive number such that $F^\dagger F \leqslant \mathbb{1}$ and the states $|\psi_i\rangle$ are defined in Subsection 3.2. It is easy to see that $F^\dagger F$ is diagonal in the basis given by the eigenvalues of X. In fact,

$$[F, X] = 0. \tag{3.6}$$

We also suppose that Alice and Bob pre-agree on a set of unitary operators $U_A^{(i)}$ and $V_B^{(j)}$ for $1 \leqslant i \leqslant \ell$ and$1 \leqslant j \leqslant \kappa$, and on a specific subset W of $\{(i,j)|1 \leq i \leq k,\, 1 \leq i \leq l)\}$.

The protocol goes as follows, Alice creates $n+m$ pairs of maximally entangled states $|\Phi\rangle = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}|k,k\rangle$. She sends the second qudits of each pair to Bob. On each pair, Alice and Bob perform at random one of the operators $U_A^{(i)}$ and $V_B^{(j)}$ followed by the filtering operations F and $F^\dagger$, respectively. Since F is not unitary, some of the pairs will be lost. Alice and Bob publicly declare which of the operators they have chosen for each pair. They discard the pairs for which $(i,j) \notin W$. Then, the normalized state of the n remaining pairs is:

$$\frac{(F_{A_1} \otimes F_{B_1}^T)^{\otimes n}}{N} \sum_{x,y\;:\;(x_i,y_j)\in W} U_A^{(x)} \otimes V_B^{(y)} (\mathbb{1} \otimes \widehat{E})|\Phi\rangle^{\otimes n}_{A_1B_1}|x,y\rangle_R \tag{3.7}$$

where $x, y \in \{0,...,d-1\}^{\otimes n}$, Eve's super-operator is represented by $\widehat{E}$, N is a normalization factor, $U_A^{(x)} := U_{A_1}^{(x_1)} \otimes ... \otimes U_{A_n}^{(x_n)}$ and $V_B^{(y)} := V_{B_1}^{(y_1)} \otimes ... \otimes V_{B_n}^{(y_n)}$. The system R is public. However, the super-operator $\widehat{E}$ is independent of the system R.

Afterwards, Alice and Bob perform quantum error correction using two commuting linear codes — also known as CSS-code [34, 175]— to correct dit and phase errors. Finally, Alice and Bob can extract the secret key by measuring the observable Z.

[3] In this thesis, we have not explained in detail how the reduction can be done since it follows straightforwardly from observations in [173] — Eve cannot differentiate the cases whether Alice is entangled (or not) with the state she sends to Bob and whether Bob delays (or not) his measurement —combined with arguments from Appendix B.3 about random linear hashing functions.

3.3.2 Exploiting A Symmetry

In theory, there exists a larger space such that Eve's operation is unitary over that space. In the worst situation, Eve possesses that whole larger space and we can assume that Alice, Bob and Eve share the pure state

$$|\psi\rangle_{ABRE} = \frac{(F_{A_1} \otimes F_{B_1}^T)^{\otimes n}}{N} \sum_{x,y\ :\ (x_i,y_j)\in W} U_A^{(x)} \otimes U_B^{(y)} (\mathbb{1} \otimes E_{BE}) |\Phi\rangle_{A_1B_1}^{\otimes n} |0\rangle_E |x,y\rangle_R. \quad (3.8)$$

We can rewrite the above state in a different basis:

$$|\psi\rangle_{ABRE} = \sum_{\mu,\nu\in\{0,\ldots,d-1\}^{\otimes n}} \sqrt{p_{\mu\nu}} \mathbb{1} \otimes X_A^{\mu} Z_B^{\nu} |\Phi\rangle_{AB} |\phi_{\mu\nu}\rangle_{ER} \quad (3.9)$$

where

$$|\phi_{\mu\nu}\rangle_{ER} = \frac{1}{N_{\mu,\nu}}\ {}_{AB}\langle\Phi|(\mathbb{1}_A \otimes Z_B^{\dagger\nu} X_B^{\dagger\mu})|\Psi\rangle_{ABER} \quad (3.10)$$

and

$$p_{\mu\nu} = \mathrm{Tr}[{}_{AB}\langle\Phi|\mathbb{1}_A \otimes Z_B^{\dagger\nu} X_B^{\dagger\mu} |\psi\rangle_{ABRE}\langle\psi| \mathbb{1} \otimes X_A^{\mu} Z_B^{\nu} |\Phi\rangle_{AB}] \quad (3.11)$$

for some normalization factors $\frac{1}{N_{\mu,\nu}}$, $Z^{\nu} = Z^{\nu_1} \otimes \ldots \otimes Z^{\nu_n}$ and $X^{\mu} = X^{\mu_1} \otimes \ldots \otimes X^{\mu_n}$. Recall that X and Z are the generalized Pauli operators defined in Appendix A.6. We now use an idea proposed by [153] about the symmetry of the protocol to prove the following Lemma:

Lemma 1 *Without lost of generality,*

$${}_{ER}\langle\phi_{\mu'\nu'}|\phi_{\mu\nu}\rangle_{ER} \propto \delta_{\nu'\nu}. \quad (3.12)$$

Proof. Consider an extended version of our prepared-and-measured protocol presented in Subsection 3.3. Suppose that before information reconciliation and privacy amplification, Alice and Bob agree publicly on some random transformation $X_A^{\omega} \otimes X_B^{-\omega}$ and flip accordingly the dits of their raw key

$$|c,c\rangle_{AB} \rightarrow X_A^{\omega} \otimes X_B^{-\omega} |c,c\rangle_{AB} \quad (3.13)$$

Note that the extra operation is classical and could easily be performed by Alice and Bob. On a first look, Alice and Bob will not share an identical key after that operation, but a simple relabeling of Bob's state will correct that situation.

The corresponding entanglement-based protocol is modified as follows. On each pair, $|\phi\rangle_{AB}$, Alice and Bob apply randomly one of the transformation $X^{\omega} F U^{(i)}$ and $X^{-\omega} F^T V^{(j)}$, respectively, and discard the pair if $(i,j) \notin W$. Since $[F,X] = 0$, then we obtain $X^{\omega} F U^{(i)} = F X^{\omega} U^{(i)}$ and $X^{-\omega} F^T V^{(j)} = F^T X^{-\omega} V^{(j)}$. Equation 3.8 becomes

$$|\psi\rangle_{ABRE} = \mathcal{F} \sum_{\omega,x,y\ :\ (x_i,y_j)\in W} X_A^{\omega} U_A^{(x)} \otimes X_B^{-\omega} U_B^{(y)} (\mathbb{1} \otimes E_{BE}) |\Phi\rangle_{A_1B_1}^{\otimes n} |0\rangle_E |x,y,\omega\rangle_R$$

where $\mathcal{F} := \frac{(F_{A_1} \otimes F_{B_1}^T)^{\otimes n}}{N}$.

Therefore,

$$X_A^{\omega} \otimes X_B^{-\omega} |\psi\rangle_{ABRE} \quad = \quad T_R(\omega)|\psi\rangle_{ABRE} \tag{3.14}$$

where T_R is the translation operator such that $T_R(\omega')|x, y, \omega\rangle_R = |x, y, \omega - \omega'\rangle_R$.

We can calculate that

$$\begin{aligned}
&{}_{ER}\langle\phi_{\mu'\nu'}|\phi_{\mu\nu}\rangle_{ER} = \langle\phi_{\mu'\nu'}|T_R(-\omega)T_R(\omega)|\phi_{\mu\nu}\rangle_{ER} && (3.15)\\
&= \frac{1}{N_{\mu\nu\mu'\nu'}}\ {}_{ABER}\langle\Psi|T_R(-\omega)(\mathbb{1}_A \otimes X_B^{\mu'} Z_B^{\nu'})(|\Phi\rangle_{A_1B_1}\langle\Phi|)^{\otimes n}(\mathbb{1}_A \otimes Z_B^{\dagger\nu} X_B^{\dagger\mu})T_R(\omega)|\Psi\rangle_{ABER} &&\\
&= \frac{1}{N_{\mu\nu\mu'\nu'}}\ {}_{ABER}\langle\Psi|X_A^{-\omega} \otimes X_B^{\omega+\mu'} Z_B^{\nu'})(|\Phi\rangle_{A_1B_1}\langle\Phi|)^{\otimes n} X_A^{\omega} \otimes Z_B^{\dagger\nu} X_B^{\dagger\mu+\omega}|\Psi\rangle_{ABER} &&\\
&= (e^{\frac{2\pi i}{d}})^{\omega\cdot(\nu-\nu')}\ {}_{ER}\langle\phi_{\mu'\nu'}|\phi_{\mu\nu}\rangle_{ER} && (3.16)
\end{aligned}$$

where $N_{\mu\nu\mu'\nu'} := N_{\mu\nu}N_{\mu'\nu'}$. In the last step, we used the relation $ZX = e^{\frac{2\pi i}{d}}XZ$. Since the above equation is true for all ω, then we conclude

$${}_{ER}\langle\phi_{\mu'\nu'}|\phi_{\mu\nu}\rangle_{ER} \propto \delta_{\nu'\nu}. \tag{3.17}$$

■

3.4 Information Reconciliation: Forging a Shield

We now consider dit error correction. For this purpose, Alice and Bob choose a family of classical error correction codes. As explained in Appendix A.7, each one-way deterministic classical error correction code is associated[4] to a projector

$$P^C := \sum_{c\in C} |c\rangle\langle c|. \tag{3.18}$$

where C is a set of codewords. We restrict our attention to a family of error correction codes which respects $\sum_C p_C P^C = \mathbb{1}$, where p_C is the probability distribution of the codes. Alice performs on her system the measurement corresponding to the POVM $\{p_C P^{(C)}\}$. She declares publicly her outcome C_a. Then, using extra ancillas in a new system B', Bob can apply the unitary recovery operation $R^{(C_a)}_{BB'}$ as described in Appendix A.7.

The recovery operation depends on the model of noise that we want to correct. If the model of noise is unknown, we cannot determine which recovery operation should be applied. One of the methods to determine the noise model is to take at random a subset of the pairs of qudits and use them as *test dits*. Suppose that Alice and Bob measure m random pairs of qudits in the Z basis and declare to each other their results, they just made a measurement of the dit error rates $e^{test}_{\mu=i}$ for $0 \leqslant i \leqslant d-1$ in the test dits.[5] Since test dits were chosen at random, then $e^{test}_{\mu=i}$ should be an exponential reliable estimation of the dit error rate $e_{\mu=i}$ in the remaining dits.[6] So it would be sufficient to find a code that is able to correct all error patterns that are compatible with our measurement.

It is often sufficient to guess what the error rate will be and to apply accordingly an error correction code that is easy to implement and afterwards, verify that the errors have really been corrected. Here is how to efficiently check if the error has really been eradicated. Suppose that Alice and Bob each possess a string of n dits s_a and s_b. They want to verify with an exponential reliability whether $s_a = s_b$ or not. They pick a random ℓ by n matrix M with entries $\{0, 1, ..., d-1\}$. They both compute Ms_a and Ms_b and declare publicly

[4] Every one-way deterministic classical error correction code is defined by a set C of codewords and a recovery function $r^{(C)}$. There are many different error correction codes with the same set C of codewords each of those being associated to a different recovery function. To simplify our formula, we will suppose that there is only one recovery function associated to each set C (i.e. Alice and Bob always use the same recovery function for a specific set of codewords.). However, the generalization is straightforward and do not modify any of our conclusions.

[5] The dit error rates $e_{\mu=i}$ for $0 \leqslant i \leqslant d-1$ are a direct generalization of the bit error rate. The error rate associated with $e^{test}_{\mu=i}$ is a rate of event where Alice and Bob's paired dits are $|k\rangle$ and $|k+i\rangle$, respectively, for any $0 \leqslant k \leqslant d-1$.

[6] Suppose that $e^{tot}_{\mu=i}$ are the dit error rates in n pairs, then $(n-m)e_{\mu=i} + me^{test}_{\mu=i} = ne^{tot}_{\mu=i}$ and $e_{\mu=i} - e^{test}_{\mu=i} = \frac{n}{n-m}(e^{tot}_{\mu=i} - e^{test}_{\mu=i})$. As a consequence of the large deviation theorem, $\Pr[|e^{test}_{\mu=i} - e^{tot}_{\mu=i}| > \delta]) < e^{-\frac{m\delta^2}{2}}$. Therefore, $\Pr[|e^{test}_{\mu=i} - e_{\mu=i}| > \delta]) < e^{-\frac{mn^2\delta^2}{2(n-m)^2}}$.

their result. If $s_a \neq s_b$, then $\Pr[Ms_a = Ms_b] = \frac{1}{d^\ell}$. The number of dits spent for this test is ℓ (Alice and Bob simply discard a different pair for each row. Each pair must be associated to a non-zero entry in the corresponding row.).

After the bit error correction, the resulting state is

$$\sqrt{\frac{d^n}{|C_a|}} P_A^{(C_a)} \otimes R_{BB'}^{(C_a)} |\psi\rangle_{ABRE} |0\rangle_{B'} \tag{3.19}$$

which is — or can be verified to be — close in trace distance to a state of the form:

$$\sqrt{\frac{d^n}{|C_a|}} \sum_{\nu \in \{0,\ldots,d-1\}^{\otimes n}} \sqrt{p_\nu} P_A^{(C_a)} \otimes Z_B^\nu |\Phi\rangle_{A_1 B_1}^{\otimes n} \sum_{\mu \in \{0,\ldots,d-1\}^{\otimes n}} p_{\mu|\nu} |\mu\rangle_{B'} |\phi_{\mu\nu}\rangle_{ER} \tag{3.20}$$

where $p_\nu = \sum_\mu p_{\mu\nu}$ and $p_{\mu|\nu}$ is the probability of bit error pattern μ conditional to the phase error pattern ν. We use the fact that the recovery operation stores with high probability the bit error pattern in the system B' (see Appendix A.7). Using Lemma 1 and taking the partial trace over the systems E and R, we obtain

$$\rho_{ABB'} \quad = \quad \frac{d^n}{|C_a|} \sum_\nu p_\nu Z_B^\nu P_A^{(C_a)} (|\Phi\rangle_{A_1 B_1} \langle \Phi|)^{\otimes n} P_A^{(C_a)} Z_B^{-\nu} \otimes \sigma_{B'}^{(\nu)} \tag{3.21}$$

where $\sigma_{B'}^{(\nu)} := \sum_{\mu,\mu'} \sqrt{p_{\mu|\nu} p_{\mu'|\nu}} \; {}_{ER}\langle \phi_{\mu',\nu} | \phi_{\mu,\nu} \rangle_{ER} |\mu\rangle_{B'} \langle \mu|$. The system B' can be used as a Shield (see Section 1.5.3). This system B' might help Bob to improve the rate of phase error correction which corresponds, in our case, to privacy amplification. Note that in Chapter 2, we did not have this equivalence between phase error correction and privacy amplification since the orthogonality of the eavesdropper given by Lemma 1 was not necessarily satisfied.

3.5 Privacy Amplification: Different Rates

The efficiency of phase error correction depends on the method used for evaluating the model of the phase noise. This model is described by some parameters that we need to estimate. We will investigate two methods for parameter's estimation. The first involve Azuma's inequality and the second, the quantum de Finetti theorem. As we will see, higher key rates are achievable using the quantum de Finetti theorem, but using Azuma's inequality simplify considerably the proof and improve — as far as we know — the security level of the protocol. The key ingredient to achieve higher rate will be to exploited the Shield that we constructed from the dit error correction in Section 3.4.

3.5.1 Parameter's Estimation Using Azuma's Inequaltiy

Until now, we have always written the state of all the pairs together. Recalling Equation C.2, let us consider the state of any one pair shared by Alice and Bob:

$$\rho_{A_iB_i} = \frac{(F_{A_i} \otimes F_{B_i}^T)}{N_0} \sum_{(x_1,y_1)\in W} U_{A_i}^{(x_1)} \otimes V_{B_i}^{(y_1)} \widehat{E}_{B_i}^{(i)}[|\Phi\rangle_{A_iB_i}\langle\Phi|] U_{A_i}^{(x_1)\dagger} \otimes V_{B_i}^{(y_1)\dagger} \tag{3.22}$$

for some normalization factor N_0. The super-operator $\widehat{E}_{B_1}^{(1)}$ depends on which pair we are looking at. It also depends on any measurement's outcome obtained from other pairs.

Their are several types of errors. Each of them corresponds to a probability:

$$p_{\mu\nu}^{(i)} = {}_{A_iB_i}\langle\Phi|\mathbb{1}_{A_i} \otimes X_{B_i}^{-\mu} Z_{B_i}^{-\nu} \rho_{A_iB_i} \mathbb{1}_{A_i} \otimes Z_{B_i}^{\nu} X_{B_i}^{\mu} |\Phi\rangle_{A_iB_i} \tag{3.23}$$

where $\mu, \nu \in \{0, ..., d-1\}$. Note that those probabilities are not independent (for different i). Define

$$p_{\mu}^{(i)} := \sum_{\nu} p_{\mu\nu}^{(i)} \tag{3.24}$$

and

$$p_{\nu}^{(i)} := \sum_{\mu} p_{\mu\nu}^{(i)}. \tag{3.25}$$

Consider the Kraus representation

$$\widehat{E}_{B_i}^{(i)}(\rho) = \sum_j E_j^{(i)} \rho E_j^{(i)\dagger} \tag{3.26}$$

where $\sum_j E_j^{(i)\dagger} E_j^{(i)} \leqslant \mathbb{1}$.

Define

$$\rho_{A_1B_1}^{(j)} = \frac{(F_{A_1} \otimes F_{B_1}^T)}{N_0} \sum_{(x_1,y_1)\in W} U_{A_1}^{(x_1)} \otimes V_{B_1}^{(y_1)} E_j^{(i)}[|\Phi\rangle_{A_1B_i1}\langle\Phi|] E_j^{(i)\dagger} U_{A_1}^{(x_1)\dagger} \otimes V_{B_1}^{(y_1)\dagger} \tag{3.27}$$

so that

$$\rho_{A_1B_1} = \sum_j \rho^{(j)}_{A_1B_1}. \tag{3.28}$$

We can then define

$$p^{(i,j)}_{\mu\nu} = \ _{A_iB_i}\langle\Phi|\mathbb{1}_{A_i} \otimes X^{-\mu}_{B_i} Z^{-\nu}_{B_i} \rho^{(j)}_{A_iB_i} \mathbb{1}_{A_i} \otimes Z^{\nu}_{B_i} X^{\mu}_{B_i}|\Phi\rangle_{A_iB_i}\ , \tag{3.29}$$

$$p^{(i,j)}_{\mu} := \sum_\nu p^{(i,j)}_{\mu\nu} \tag{3.30}$$

and

$$p^{(i,j)}_{\nu} := \sum_\mu p^{(i,j)}_{\mu\nu}. \tag{3.31}$$

By direct calculation, for any value of the matrix $E^{(j)}$, it is often possible to find an upper bound for $p^{(i,j)}_\nu$ which is some linear functions of the $p^{(i,j)}_\mu$'s. For example, in BB84, $p^{(i,j)}_{\nu=1} = p^{(i,j)}_{\mu=1}$. In the six-state-protocol, $p^{(i,j)}_{\mu=1,\nu=0} = p^{(i,j)}_{\mu=1,\nu=1} = p^{(i,j)}_{\mu=0,\nu=1}$. In the PBC00 protocol, $p^{(i,j)}_{\nu=1} = \frac{5}{4}p^{(i,j)}_{\mu=1}$ [27]. When determining such a bound is impossible (like in the case of B92), another method is necessary.

By linearity, if there is a (linear) relation between the $p^{(i,j)}$'s that is independent of the form of the matrix $E^{(i)}$, then this relation must also be true for the $p^{(i)}$'s. Now we need to translate those relations between probabilities to relations between observed error rate.

If the error probabilities in different pairs were independent, then it would follow from the theory of large deviations that the error rates, if measured, would asymptotically follow those relations. Since the error probabilities between different pairs are not independent, we need another mathematical tool which is the Azuma's inequality. Consider the following definition:

Suppose we have a series of events $F_0, F_1, \ldots$. Let $X_0, X_1, \ldots$ be random variables. The sequence is a martingale iff the expectation of X_{i+1} conditional to events $F_i, F_{i-1}, \ldots F_0$ is equal to X_i for all i.

Consider the case of N coin tosses, where the probability of getting heads for each coin may be correlated in some way. Consider a series of events $F_0, F_1, \ldots$. Let h_i be the number of heads from the events $F_i, F_{i-1}, \ldots F_0$. Let X_i be $h_i - \sum_{j=1}^{i} p^{(j)}$ where $p^{(j)}$ is the probability of obtaining a head on the j^{th} coin conditional to events $F_{j-1}, F_{j-2}, \ldots F_0$. The expectation of X_{i+1} conditional to events $F_i, F_{i-1}, \ldots F_0$ is $h_i - \sum_{j=1}^{i} p^{(j)}$ plus the expectation of obtaining a head on the $i+1$ coin, minus $p^{(i+1)}$. Since the expectation of obtaining a head on the $i+1$ coin minus $p^{(i+1)}$ is zero, the sequence $X_0, X_1, \ldots$ is a martingale.

Theorem 8 *Special Case of Azuma's Inequality [7]: Let $X_0, X_1, \ldots$ be a martingale sequence such that for each k, $|X_k - X_{k-1}| \leq 1$. Then, for all $N \geq 0$ and any $\lambda \geq 0$,*

$$Pr[|X_N - X_0| \geq \lambda|] \leq 2e^{-\frac{\lambda^2}{2N}}.$$

Suppose that we measure, in the basis defined by $\mathbb{1}_{A_i} \otimes Z^{\nu}_{B_i} X^{\mu}_{B_i} |\Phi\rangle_{A_iB_i}$ for $0 \leqslant \mu, \nu \leqslant d-1$, exactly $n-m$ pairs one after the other, in a specific order. Let $p_{\mu\nu}^{(i \mid 1,\ldots,i-1)}$ be the probability to obtain the outcome $\mathbb{1}_{A_i} \otimes Z^{\nu}_{B_i} X^{\mu}_{B_i} |\Phi\rangle_{A_iB_i}$ on the i^{th} pair conditional to all our previous measurement. By the Azuma's inequality, we have that

$$Pr[|\frac{1}{n}\sum_{i=1}^{n-m} p_{\mu\nu}^{(i \mid 1,\ldots,i-1)} - e_{\mu\nu}| \geq \lambda| \leq 2e^{-\frac{n\lambda^2}{2}}.$$

where $(n-m)e_{\mu\nu}$ is the number of times the measurement outcome $\mathbb{1}_{A_i} \otimes Z^{\nu}_{B_i} X^{\mu}_{B_i} |\Phi\rangle_{A_iB_i}$ has occurred.

Consequently, suppose that we derived linear relations between the $p^{(i)}_{\mu\nu}$, then it translates to identical exponentially reliable relations between the error rates $e_{\mu\nu}$. At this stage, we should have a exponentially reliable upper bound $e_{\nu\uparrow}$ for the phase error rates in term of the measured dits error rates in the test dits.

We note that although Alice and Bob only have a upper bound for the phase error rates, then could have in theory used the test dits to measure with exponential reliability the exact phase error rates (which would be smaller than $e_{\nu\uparrow}$). Using random linear codes (see [20, 118] or Appendix B.2), it is straightforward to show using the theory of types that Alice and Bob could have corrected the phase error rates with exponentially high probability. This imply that they can achieve a secret key generation rate corresponding to the Hashing bound [20]:

$$log_2 d - H(e_\mu) - H(e_{\nu\uparrow}|\mu). \tag{3.32}$$

The entropy of the phase error rates are conditional to the μ since, as we saw earlier, μ is obtained from the dit error correction and can be used to find a better phase error correction code.

3.6 How to Improve the Key Generation Rate

In the previous section, we gave a method to estimate the phase errors from dit error rates, and how we could achieve a key generation rate corresponding to what is called the Hashing bound. However, we didn't use the Shield that we constructed in Section 3.4, and in light of the results we obtained in Chapter 2, we know that using the Shield might improve our rate.

In this section, instead of using the Azuma's inequality to find a direct relation between dit and phase error rate, we use the quantum de Finetti theroem to transform the noise model to one that is independent and identically-distributed. We will skip the mathematical complication involved with the introduction of the quantum de Finetti theorem, and simply assume that it gives quantum states that are independent and identically-distributed.

Applying the quantum de Finetti theorem to the Equation 3.21 and discarding some complications, we obtain

$$\rho_{ABB'} \quad = \quad d\frac{d^n}{|C_a|}P_A^{(C_a)}(\sum_{\nu=0}^{d-1} p_\nu Z_{B_1}^{\nu}|\Phi\rangle_{A_1B_1}\langle\Phi|Z_{B_1}^{-\nu} \otimes \sigma_{B_1'}^{(\nu)})^{\otimes n} P_A^{(C_a)} \tag{3.33}$$

where $\sigma_{B'}^{(\nu)} := \sum_{\mu,\mu'=0}^{d-1} \sqrt{p_{\mu|\nu}p_{\mu'|\nu}} \; {}_{E_1R_1}\langle\phi_{\mu',\nu}|\phi_{\mu,\nu}\rangle_{E_1R_1}|\mu\rangle_{B'}\langle\mu|$ and the states $|\phi\rangle$ are redefine consequently. By Theorem 5, the rate of secret key that be extracted is

$$\log_2 d - H(e_\mu) - I(\nu : BS). \tag{3.34}$$

As a simple exercise, the reader can verify that this rate is equivalent — in our case — to

$$\log_2 d - H(e_\mu) - H(e_\nu) + I(\nu : \sigma_{B_1'}^{(\nu)}). \tag{3.35}$$

where $e_\mu = p_\mu$, $e_\nu = p_\nu$ and we suppose that $\log_2 |C_a| \rightarrow n \log_2 d - nH(p_\mu)$ as $n \rightarrow \infty$. Since a measurement of $\sigma_{B_1'}^{(\nu)}$ in the computational basis give the dit error pattern, then the key generation rate given by the Formula 3.34 (and consequently also Formula 3.35) must be equal or bigger than the rate given by the Formula 3.32. For example, Formula 3.35 improve the rate of the SAGR04 protocol (this improvement was first found by [109]). Another example is PBC00. Using Formula 3.32, the highest tolerable noise rate in a depolarization channel given in Section 3.1.3 is $p = 11.6\%$ versus $p = 13.4\%$ if we use Formula 3.35.

3.7 Summary of the Chapter

3.7.1 Summary

In this Chapter, one of our contributions was to combine and generalize the results given in [173, 24, 150, 104, 151] in order to obtain a generic security for permutation-invariant QKD protocol against joint attack. Our final result is very similar to [41, 153], except that we didn't consider post-processing — which could have been done by using arguments from [151] — and that we introduced the Azuma's inequality to considerably simplify the proof. Our proof has the conceptual advantage of being directly related to entanglement distillation. The secret key generation rates that we derived using this method are related to the entanglement distillation rate given by the Hashing bound (see Equation 3.32).

A second contribution was to explain how we could improve the known secret key for some important QKD protocols by constructing a Shield from dit error correction (see Equation 3.35). With a different method, this improvement was already shown for SAGR04 [109], but we showed how such improvement could be observed on other QKD protocols.

3.7.2 Connections to Other Results and useful comments

We conclude the Chapter with a few important comments. First, there exists another security proof for QKD that exploits the structure of private states [85]. However, the problem that they solve is very different to ours, and they do not consider the same scenario as ours. For a fixed *twisting* operator of the form $U_{BS}^{\otimes n}$, they show how to estimate the phase errors in the *untwisted* state $(U_{twist}^{\otimes n})^{\dagger}\rho_{ABS}U_{twist}^{\otimes n}$ where ρ_{ABS} is shared between Alice, Bob and the Shield. One of the main differences with their approach is that we don't restrict our *twisting* operator to the form $U_{BS}^{\otimes n}$ and that our *twisting* operator is not fixed, but it does depend on Alice and Bob's state. Contrarily to our work, the phase error estimation that is considered in [85] could necessitate quantum operations that are to hard to implement.

As mentioned earlier, our results can be seen as an extension of [104]. As explained in [105], the method given in [104] leads to a security proof for QKD with uncharacterized detectors. However, the arguments in [105] apply only if Alice performs some measurements in the conjugated basis as it happens — virtually — in BB84. In our generalization of [104], we lose the feature where Alice makes a measurement in the conjugated basis. For this reason, our analysis does not apply, in general, to uncharacterized detectors. It would be very interesting to find a method, which allows uncharacterized detectors, to adapt our framework to.

From the analysis of this chapter, in the case where Azuma's inequality is used for parameter estimation, the security level — the upper bound of the trace distance between Alice and Bob's state with a private state — scales as $e^{-\alpha n}$ for some positive constant α. This agrees with the result reported in [161]. The author has also calculated the scaling of the security level in the case where the quantum de Finetti theorem was used, considering all the extra complications that were ignored in this thesis. From his calculation, the security level scales as $e^{-\beta\sqrt{n}}$ and he found that for practical implementation of QKD, this difference could be considerable. However, the author does not claim that his calculation of the security level (in the case where the quantum de Finetti theorem is used) is optimal, but in fact suggests that it be investigated further.

Chapter 4

QKD Without Reference Frame

4.1 Background

Many QKD protocols has been proposed, and each different implementation of the protocols has a different set of advantages and disadvantages. Some are more sensitive to some type of noise or eavesdropping attacks, others have some important physical limitation. With the technological advances, it is difficult, if not impossible, to foresee which QKD protocol combined with which physical implementation will be the best. The most studied QKD protocols so far are legitimately the ones that involve single photon encoding and measurement, since they are far easier to conceive experimentally. However, multi-photon encodings could have there own set of advantages. There exists channels for which there is no protocol using a single photon encoding that can generate a secret key. On the other hand, by exploiting the decoherence-free subspace of the noise, there are multi-photon schemes that can succeed in generating a secret key. Depending on the situation, encoding could be used to improve considerably the performance of QKD in a channel with a memory.

One of the common problems in implementing QKD protocols in a non-control environment is dealing with thermal and tension fluctuations in the optical fiber. Those fluctuations induce some random rotation of the polarization of the photons traveling through the fiber. For this reason, a phase-encoding is often preferred to polarization-based encoding for implementation of QKD through optical fiber. However, since devices used for phase-encoding QKD are polarization dependent, compensating for the rotation is a vital issue for the performance of the schemes. In some experiments (for examples [128, 63]), they had to calibrate their apparatus every few seconds. A good solution to this problem was provided by reflecting the photon using a Faraday mirror so that they travel back-and-forth in the optical fiber [133, 178]. We provide, here, an alternative solution by using mutli-photon polarization encoding. In a typical present situation, there is no doubt that single photon QKD schemes are easier to implement. However, in the future, it may not necessarily still be so. It is easy to imagine situations where two-way quantum communication could be a huge burden — imagine a QKD protocol between two moving objects, say a satellite and a spaceship. This does not mean that some multi-photon QKD protocol will necessarily be better than single photon QKD protocols in any practical situation, but it is certainly worthwhile to study multi-photon QKD protocols to be able to know to which extent they could be useful. It is always interesting to find some practical application to entanglement, even if it is not obvious that it compensates the difficulties in manipulating it.

In the next sections, we study the properties and the unconditional security of one specific mutli-photon QKD protocol made to be robust against collective noise[24], hoping that our analysis might help to study other more complicated cases. We also consider several new variations of those protocols. Interactions between photons are generally hard to perform, and the use of multiple photon states accentuate the sensitivity of the protocol to channel losses. However, allowing ourself to use not too futuristic technologies (i.e. two-photon source on demand, quantum memories and entanglement-swapping), we demonstrate that these problems could be reduced or avoided by using judicious encodings and technics related to quantum teleportation.

We will first describe the rotationally invariant four-photon protocol introduced in [24]. We show in section Sec. 4.3 how variations of the protocol allow further improvements by making the schemes more efficient against channel losses. In Sec. 4.4, we demonstrate the security of the four-photon protocol. We remark that some of our arguments could also apply to other multi-photon schemes, especially the rotational invariant ones [25].

4.2 Robust Four-Photon QKD Protocol

Consider the problem of performing QKD when Alice and Bob do not share a reference frame. This could happen for example when they communicate through optical fiber, which randomly rotates the polarization of photon states. An other possibility would be that Bob is a satellite in orbit, or on a boat floating on the tumultuous ocean. This challenge can be met by encoding the information in rotationally invariant states, that span the *decoherence free subspace* (DFS) of the collective rotation operation. The singlet state $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ has the property of being invariant under bilateral rotation of the two qubits, and therefore spans a one-dimensional DFS of the collective rotation. For the purpose of encoding an arbitrary qubit state however, one dimension is insufficient. It takes at least four qubits to yield a two dimensional DFS. (Three qubits are enough to generate a full qubit algebra via the noiseless subsystem method.) A possible choice of orthogonal basis for this subspace is [93]

$$|\hat{0}\rangle = \frac{1}{\sqrt{2}}(|a\rangle - |b\rangle) \tag{4.1a}$$

$$|\hat{1}\rangle = \frac{1}{\sqrt{6}}(2|c\rangle - |a\rangle - |b\rangle) \tag{4.1b}$$

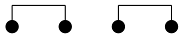
where

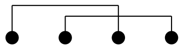
$$\begin{aligned} |a\rangle &= \frac{1}{\sqrt{2}}(|0101\rangle + |1010\rangle) \\ |b\rangle &= \frac{1}{\sqrt{2}}(|0110\rangle + |1001\rangle) \\ |c\rangle &= \frac{1}{\sqrt{2}}(|0011\rangle + |1100\rangle). \end{aligned}$$

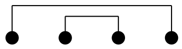
Theoretically, these states could be use to implement any of the aforementioned QKD protocols by preparing and measuring *encoded* qubits restricted to the DFS. Some efforts in this direction have been done using photon-polarization states [29].

Nevertheless, preparing encoded states and performing encoded operations can sometimes be a very difficult task. To implement BB84 for example, at least one of the four encoded states would necessitate four-photon entanglement, so would be difficult to produce efficiently. Similarly, the encoded measurements would require some interaction between all four-photons. Hence, our motivation in [24] was to find a scheme which is adapted to the DFS structure, thus circumventing the difficult challenges of encoded preparation and measurements. It turns out that the encoded version of PBC00 (see Section 3.1.3) fulfills these requirements.

Consider the following three quantum states that are obtained by permuting the qubits of two singlet states [24]:

$|\psi_1\rangle =$  $= \frac{1}{\sqrt{2}}(|a\rangle - |b\rangle)$

$|\psi_2\rangle =$  $= \frac{1}{\sqrt{2}}(|c\rangle - |b\rangle)$

$|\psi_3\rangle =$  $= \frac{1}{\sqrt{2}}(|a\rangle - |c\rangle) \qquad (4.2)$

In the above diagram, linked circles represent two qubits in the singlet state $|\Psi^-\rangle$. The qubits are assumed to be distinguishable, e.g. from their temporal ordering. Such states can be created e.g. using parametric down conversion. The states $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\psi_3\rangle$ are obviously rotationally invariant, and therefore lie inside the DFS. Using the notation of Equation 4.1, they can be re-written as $|\hat{0}\rangle$, $\frac{1}{2}|\hat{0}\rangle + \frac{\sqrt{3}}{2}|\hat{1}\rangle$ and $\frac{1}{2}|\hat{0}\rangle - \frac{\sqrt{3}}{2}|\hat{1}\rangle$, respectively. These are the three states required to implement the PBC00 protocol, so encoded state preparation requires only two-qubit entanglement.

The other ingredient required to implement PBC00 is the POVM $\mathcal{M}_{\text{PBC00}}$ describe by Equation 3.3. This measurement also involves four-photon entangled states, and is therefore technologically very challenging. However, instead of performing this complicated measurement, suppose that Bob measures each of the four qubits in the "computational" basis $\{|0\rangle, |1\rangle\}$, i.e.

$$\mathcal{M}_{\text{comp}} = \{|v\rangle\langle v| \text{ for } v = 0000, 0001, 0010, ..., 1111\}. \qquad (4.3)$$

Under the assumption that the state of the qubits received by Bob lies in the DFS, then $\mathcal{M}_{\text{comp}}$ surprisingly reduces to $\mathcal{M}_{\text{PBC00}}$. Indeed,

$$P_{\text{dfs}}|v\rangle = \begin{cases} \frac{1}{\sqrt{3}}|\overline{\psi}_1\rangle & \text{if } v = 0011 \text{ or } 1100 \\ \frac{1}{\sqrt{3}}|\overline{\psi}_2\rangle & \text{if } v = 0101 \text{ or } 1010 \\ \frac{1}{\sqrt{3}}|\overline{\psi}_3\rangle & \text{if } v = 0110 \text{ or } 1001 \\ 0 & \text{else} \end{cases} \qquad (4.4)$$

where $P_{\text{dfs}} = |\hat{0}\rangle\langle\hat{0}| + |\hat{1}\rangle\langle\hat{1}|$ is the projection operator onto the DFS of the collective noise. Notice that when the state is indeed in the DFS, the measurement outcome should always indicate a balanced number of 0's and 1's.

Thus, if Bob could verify that the qubits he received are indeed encoded in the DFS, or at least estimate the rate of states that fall outside the DFS, then single qubit measurements and two-qubit entangled source would be sufficient to implement a rotationally-invariant encoded version of PBC00. However, a projection in the DFS might be difficult to implement experimentally. In the absence of such a projection, Eve could use the complementary space of the DFS to control the measurement outcomes obtain by Bob. A very simple and efficient eavesdropping attack would consist in performing the measurement

$\mathcal{M}_{\text{PBC00}}$ — whose elements have support outside the DFS —, and sending an exact copy of the result to Bob. By doing so, Eve could always predict Bob's measurement outcome and would get full information about the key without being detected.

To counter this attack, Bob can apply a random collective rotation $U^{\otimes 4}$ to all four qubits before performing the measurement $\mathcal{M}_{\text{comp}}$ (U is chosen randomly according to the uniform Haar measure on $SU(2)$). This is equivalent to choosing a random polarization basis in which to measure all four photon.[1] The random rotation $U^{\otimes 4}$ has no effect on the states encoded in the DFS as they are rotationally invariant by construction. However, as a consequence of the random rotation $U^{\otimes 4}$, any state outside the DFS will result in an invalid measurement outcome with good probability, i.e. an outcome with an unbalanced number of 0's and 1's. Hence, by performing the polarization measurement in a random basis, Bob will be able to estimate the rate at which Eve encodes information outside the DFS. In Sec. 4.4, we will demonstrate that this simple scheme is unconditionally secure.

[1] Moreover, instead of choosing a basis at random over the entire Bloch sphere, Bob can simply choose among an σ_x, σ_y, or σ_z measurement, which has obvious practical advantages [142].

4.3 Variation of the robust four-photon protocol with higher transmission rate

4.3.1 Transmission rate η^2

Alice prepares n pairs of photons in singlet states. She applies a random permutation to the photons and sends all of them to Bob. (We assume here that the photons are distinguishable by their temporal ordering.) Bob measures the polarization of all n photons in a randomly selected basis. Then, he publicly announces to Alice which photons were received, i.e. which of his detectors clicked. For $n > 2/\eta^2$, the probability that two pairs of photons have actually made it to destination will approach unity. Among the photons that were detected by Bob, Alice picks two pairs of photons that originally formed singlet states. (In the case where less than two pairs are detected by Bob, they discard the data and resume.) She then publicly announces the set of four photons that contain these two pairs.

At this stage, Alice and Bob have a simulation of the robust four-photon protocol. Observe that there are three possible arrangements of the order of these photons, corresponding to the diagrams of Eq. (4.2). Alice selects two of these three possible arrangements, one of which is corresponding to the actual arrangement. She announces these two arrangements publicly, the first one encoding the key bit 0, while the second encoding the key bit 1. In the light of this information and his measurement outcomes, Bob can discriminate between the two arrangements with some finite probability, and generate a key bit, exactly as he did in the four-photon protocol, i.e. via Eq. (4.4). The transmission rate of this protocol is thus of order η^2.

Notice that Bob performs his measurements before learning from Alice which set of four photons contain entangled pairs, which might appear to be a major departure from the original protocol where there were only four photons to start with. Similarly, in the original protocol, Alice announces the two possible photon arrangements before sending them to Bob, while here she only makes the announcement after the photon are measured by Bob. The crucial point is that the order in which these steps are performed has no effect whatsoever on the outcome of the measurements, so the two protocols are equivalent, except that that the choice of photon that composed a four-photon state is dependent on the eavesdropper action. In the subsection 4.4.3, we show that it does not affect the security of the protocol.

4.3.2 Transmission rate η

To reach a transmission rate η, Alice must have the ability to store quantum information and to perform Bell measurements. She must first create n qubits in singlet states, and send one qubit of each pair to Bob. The qubit that she keeps will never be transmitted to Bob, so it does not need to be encoded in a photon. The n singlet could for example consist of photon-atom pairs, which can be prepared using quantum electrodynamic cavities; the photon gets sent to Bob, and the atom remains in Alice's cavity. This would greatly simplify Alice's storage requirement.

Bob measures all photons in a randomly selected polarization basis. He then announces publicly which detectors have recorded a photon; there should be roughly $n\eta$ of them. Alice then discards all atoms whose photon partner was not properly detected by Bob. Alice randomly pairs up the remaining atoms, and performs a Bell measurement on each pair. She keeps only those pairs of atoms whose measurement outcome corresponds to the singlet state. The effect of this measurement is to transfer the correlations to the photon partners of the atoms; a procedure known as *entanglement swapping* [191]. Alice announces publicly a set of four photons whose atom partners yielded the singlet measurement outcome. When $n > 4/\eta$, the probability that at least two pairs of atoms yield this outcome approaches unity. If there are less than two such pairs, they resume the protocol.

At this stage, Alice and Bob have a simulation of the four photon protocol. They proceed as in the previous section. Again, the temporal order of some of the steps of the protocol has been modifies, but without affecting the outcome of the protocol. The transmission rate of this protocol is of order η.

4.4 Security of the robust four-photon QKD protocol

Following the method described in Chapter 3, the PBC00 protocol can be shown secure by reducing it to a EDP.[2] In that case, the phase and bit error rate are related by $e_{\text{phase}} \approx \frac{5}{4} e_{\text{bit}}$.

Since the robust four-photon protocol is an encoded version of PBC00, it also reduces to an (encoded) EDP. Thus, to establish its security, all we need to show is that *i*) Bob can verify that the received state lies in the code subspace (the DFS) and *ii*) Alice and Bob can estimate the rate of (encoded) bit and phase error. The next two subsections demonstrates how this is accomplished.

4.4.1 Projection onto the DFS

The effect of Bob's random collective rotation $U^{\otimes 4}$ to all four qubits — which in practice is equivalent to performing the polarization measurements in a random basis — is to transform the received state ρ into

$$\mathcal{E}(\rho) = \frac{1}{\mathcal{N}} \int U^{\otimes 4} \rho (U^\dagger)^{\otimes 4} dU \tag{4.5}$$

where dU is the uniform Haar measure over $SU(2)$ and $\mathcal{N} = \int dU$. This map is often referred to as the *twirling* operation. We know from representation theory, and in particular from the angular momentum composition theory, that this transformation breaks up the Hilbert space according to the irreducible representations (irrep) of $SU(2)$. In this case, the 16-dimensional Hilbert space $\mathcal{H}$ of four qubits can be decomposed as

$$\begin{aligned} \mathcal{H} &= (\mathcal{H}_0)^{\otimes 2} \oplus (\mathcal{H}_1)^{\otimes 3} \oplus \mathcal{H}_2 && (4.6) \\ &= \mathbb{C}^2 \otimes \mathcal{H}_0 \oplus \mathbb{C}^3 \otimes \mathcal{H}_1 \oplus \mathbb{C} \otimes \mathcal{H}_2 && (4.7) \end{aligned}$$

where $\mathcal{H}_j$ corresponds to the space of total angular momentum j and has dimension $2j+1$. Equation (4.7) thus indicates that the spin-0 irrep is 2-fold degenerate, the spin-1 irrep is 3-fold degenerate, and the spin-2 irrep is non-degenerate. A convenient basis choice for $\mathcal{H}$ is thus $|j, m, \mu\rangle$ where j labels the irrep; m corresponds to the total angular momentum along the z axis, so it takes the values $-j, -j+1, \ldots, j$ and labels the basis states in each irrep $\mathcal{H}_j$; and μ is a degeneracy index, labeling basis states in the $\mathbb{C}^{d_j}$ sectors. The label μ can be chosen to represent the different symmetries of the state under the possible permutation of the four qubits.

Define the projectors $P_j = \sum_{m,\mu} |jm\mu\rangle\langle jm\mu|$ that project onto the subspaces with fixed value of j. The effect of the map $\mathcal{E}$ on a state ρ can be described as follows. First,

[2] Based on the analysis given in Subsection 3.6, it is possible to improve the secure key rate the PBC00 protocol. We simplify our proof by putting aside those extra considerations.

ρ gets completely "decohered" in the j sector, i.e. $\rho \to \sum_j P_j \rho P_j$. Then, each $\mathcal{H}_j$ sector (cf. Eq. (4.7)) is put into a maximally mixed state. Formally, the overall process can be written

$$\mathcal{E}(\rho) = \rho_0 \oplus \rho_1 \otimes \frac{1\!\!1_{\mathcal{H}_1}}{3} \oplus \rho_2 \otimes \frac{1\!\!1_{\mathcal{H}_2}}{5} \tag{4.8}$$

where $\rho_j = Tr_{\mathbb{C}^{d^j}}(P_j \rho P_j)$, or in terms of matrix elements,

$$[\mathcal{E}(\rho)]_{jm\mu,j'm'\mu'} = \delta_{jj'}\delta_{mm'} \sum_{m=-j}^{j} \rho_{jm\mu,jm\mu'}. \tag{4.9}$$

The DFS corresponds to the zero angular momentum subspace, so it encodes on the $\mathbb{C}^2$ sector of Eq. (4.7). We will thus write $P_{\text{DFS}} = P_0$ for the projector onto the DFS and $P_{\text{DFS}}^{\perp} = P_1 + P_2$ for the projector onto its orthogonal complement. As expected, when the state is initially in the DFS, i.e. when $P_{\text{DFS}}\rho P_{\text{DFS}} = \rho$, then the map $\mathcal{E}$ has no effect as can be seen from Eq. (4.8). Moreover, observe that $[\mathcal{E}(\rho), P_j] = 0$ for all j. As a consequence, whether or not Bob performs the measurement $\{P_{\text{DFS}}, P_{\text{DFS}}^{\perp}\}$ will not affect the measurement outcomes of his subsequent photon measurements.

Indeed, one can think of $\mathcal{E}(\rho)$ of Eq. (4.8) as a statistical mixture of states that lie in a space with a fixed value of j. In other words, there are no *coherent* superpositions of different j values, only statistical mixtures. Thus, a measurement of $\{P_j\}$ would reveal the value of j, without modifying the state otherwise.

The security of the robust four-photon QKD protocol then follows from a variation of Theorem 6 of [73] about *tagging*.[3] Essentially, we can pessimistically assume that every time the state lays outside the DFS, Eve has complete information about the associated key bit. Thus, it is possible to generate a secure key using this protocol at the rate

$$S = p_{\text{conc}} \left[\frac{p_{\text{DFS}}}{p_{\text{conc}}} - h(e_{\text{bit}}) - \frac{p_{\text{DFS}}}{p_{\text{conc}}} h(e_{\text{phase}}^{\text{DFS}}) \right] \tag{4.10}$$

where e_{bit} is the bit error rate over all conclusive results, $e_{\text{phase}}^{\text{DFS}}$ is the phase error rate of states in the DFS, and p_{DFS} is the fraction of states that lie in the DFS. (In [73], projectors onto single and multi-photon states played the role of projectors onto and outside the DFS respectively. The derivation is otherwise identical.) This result holds independently of

[3] The *tagging* argument can be understood in terms of the *Shield*. In our situation, it is explained as follow. If we prove that there exists a system S that is not accessible to Eve, but that contains full information about which state received by Bob is projected inside or outside the encoding space, then Bob's could have use the information in that system S to correct the phase error (see Chapter 3). Therefore, for the purpose of phase error correction, Bob could correct the phase errors separately on the qubits that correspond to a projection inside or outside the DFS. Assuming the worst case where the phase error probability on the results that where projected outside the DFS is $\frac{1}{2}$, then we obtain directly Equation C.5.

Bob's practical ability to distinguish stated inside and outside the DFS, as long as he can estimate p_{DFS} and $e_{\text{phase}}^{\text{DFS}}$. We will demonstrate how this can be done shortly.

As a consequence of this result, the key ingredient required to reduce the robust four-photon protocol to the provably secure PBC00 — verifying that the state is actually in the DFS — does not need to be performed. After applying the random rotation to the qubits — which effectively projects the state randomly onto or outside the DFS — all Bob needs to do is to estimate the fraction p_{DFS} of states that lie inside the DFS and their phase error rate $e_{\text{phase}}^{\text{DFS}}$.

4.4.2 Estimating p_{DFS} and $e_{\text{phase}}^{\text{DFS}}$

First, observe that for each value of j, only the states with $m = 0$ yield a valid measurement outcome, i.e. one with a balanced number of 0's and 1's (see Equation. 4.4). The DFS is associated with the $j = 0$ subspace in which all states have $m = 0$, so states in the DFS always yield valid outcomes. The states outside the DFS however can take different values of m. In the $j = 1$ subspace, m takes the values $-1, 0, 1$. Moreover, the effect of $\mathcal{E}$ is to completely mix all three of these states, so they occur with the same probability. We conclude that if ρ had support on the $j = 1$ subspace, it would yield an invalid outcome with probability $2/3$. Similarly, a state ρ with support on the $j = 2$ subspace, it would yield an invalid outcome with probability $4/5$.

Let k_m be the number of Bob's results associated to the space span by $\{|jm\mu\rangle \mid \forall\mu, j\}$. For examples, $k_{m=2}$ is the number of Bob's results that are associated to state $|1111\rangle$ and $k_{m=1}$ is the number of of Bob's results that are associated to either one of the following states: $|1110\rangle$, $|1101\rangle$, $|1011\rangle$ and $|0111\rangle$. Results that are associated to $m \neq 0$ are always invalid. The results that are that are associated to $m = 0$ may or may not be conclusive. Let k_{conc} be the number of conclusive (i.e valid) results. Then $p_{\text{conc}} = \frac{k_{\text{conc}}}{n}$ where n is the total number of four-photon states sent by Alice. If losses in the channel are neglected then $n = \sum_{m=-2}^{2} k_{\text{conc}}$.

If Bob records $k_{m=1}$ and $k_{m=2}$ of invalid measurement outcomes corresponding to states with $m = 1$ and 2, he estimates that asymptotically, $3(k_{m=1} - k_{m=2})$ and $5k_{m=2}$ states where projected outside the decoherence-free subspace into the subspace corresponding to $j = 1$ and $j = 2$, respectively.[4] Alice and Bob discard the qubits that yielded invalid outcomes, so these are not counted as conclusive. Thus an unrecognizable $k_{m=1}$ remaining results should correspond to states outside the DFS. This information is enough to determine a lower bound on the fraction of states inside the DFS: $p_{\text{DFS}} \geq \frac{k_{\text{conc}} - \frac{2}{3}k_{m=1}}{n}$.

To estimate $e_{\text{bit}}^{\text{DFS}}$, we can again be pessimistic and assume that all bit errors are results that correspond to a projection into the DFS. This sets $e_{\text{bit}}^{\text{DFS}} = e_{\text{bit}}$ where e_{bit} is evaluated

[4] Asymptotically, we should have $k_m = k_{-m}$.

from the test bits (that are selected among the valid measurement outcomes). When the states are in the DFS, the protocol amounts to an encoded version of PBC00, so the relation $e_{\text{phase}}^{\text{DFS}} = \frac{5}{4} e_{\text{bit}}^{\text{DFS}}$ derived in [27] gives us all the required quantities.

We will now demonstrate with a specific example that our two pessimistic assumptions — that a fraction 1/3 of the states outside the DFS and associated to $m = 0$ yield valid measurement results and that all bit flip errors occur in the DFS — are actually tight. Suppose Eve performs a measurement in the computational basis $\mathcal{M}_{\text{comp}}$. Upon measurement outcomes 0011 or 1100, she send Bob the state $|1,0,1\rangle = \frac{1}{\sqrt{2}}(|0011\rangle - |1100\rangle)$, sends $|1,0,2\rangle = \frac{1}{\sqrt{2}}(|0101\rangle - |1010\rangle)$ for outcomes 0101 or 1010, and finally $|1,0,3\rangle = \frac{1}{\sqrt{2}}(|0110\rangle - |1001\rangle)$ for 0110 or 1001. As the notation indicates, these three states correspond to the three degenerate $m = 0$ states of the $j = 1$ irrep, so they have no support on the DFS which has $j = 0$. The effect of Bob's random rotation will be to take each of these $|1,0,\mu\rangle$ into an equal mixture of $|1,m,\mu\rangle$ for $m = -1,0,1$. With probability 1/3, this state will yield a valid measurement outcome.

When Bob's measurement outcome is valid, then it will always agree with the one found by Eve. This is due to the fact that the map $\mathcal{E}$ does not affect the μ sector, i.e. $\mathcal{E}$ preserve the symmetries of the states under qubit permutation. In other words, the only valid measurement outcome that Bob can get are those with a non-zero overlap with the $m = 0$ state of a given irrep. Since Eve prepares the states in a known irrep (as indicated by the label μ), she can predict Bob's measurement outcome. Thus, Eve has complete information about all valid measurement outcomes, and does not induce any error. Thus, all bit errors must be associated with states in the DFS, so $e_{\text{bit}}^{\text{DFS}} = e_{\text{bit}}$. This complete the security proof of the robust four-photon QKD protocol.

4.4.3 Security of robust protocols with transmission rate η and η^2

One may worry that the modified versions of the robust four-photon QKD with higher transmission rate might allow new strategies and give extra power to Eve since the selection of the four photons, each composed of qubits, is dependent on her attack. In this subsection, we explain why that is not the case and why the same equation for the lower bound on secret key generation applies to all versions of the robust four-photon QKD protocol.

Consider the entanglement-based version of the robust four-photon QKD protocol. Alice initially prepares $4n$ maximally entangled pairs $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. She then uses entanglement-swapping on her part of the pairs to encoded states for the protocol (Equation 4.2). Then she can wait for Bob to declare which qubits he received, or did not receive, before performing the entanglement-swapping. In the standard version of the protocol, entanglement-swapping is done within blocks of four consecutive pairs. However, Alice could instead have discarded all the qubits associated with losses in the channel, regrouped the remaining pairs in fours and performed the entanglement-swapping within those groups. The version of the robust four-photon protocol with transmission η reduces to this second version of the entanglement based protocol. Even if the encoding space was chosen after Bob's declaration, we can still use the upper bound for the bit and phase error rates in the encoded space along with the estimate of how many qubits have been projected outside the encoded space, which we derived in Subection 4.4.2. Therefore, the secret key generation rate is still given by Equation C.5. A similar argument can be used to prove the security of robust four-photon protocol with transmission η^2.

Chapter 5

Conclusion

5.1 Results

The equivalence between security proof based on the uncertainty principle and private state distillation follows from Theorem 2, which was strongly inspired by [130, 105]. In [86], it was proved that every QKD protocol could be reduced to a private state distillation protocol. When new security proofs [154, 50, 153, 152], not based on distillation arguments, improved the known lower bound for secure key rates, we wanted to find the private state distillation protocols that correspond to those higher key rates. We therefore constructed the quantum operations needed by Alice and Bob to perform the corresponding private state distillation protocol and obtained an alternative proof of a major result about privacy amplification (Theorem 7 [154, 152]). Our method is a direct generalization of [173, 105], which are also based on the use of two commuting linear error correction codes, also known as CSS codes [34, 175]. The main difference being that we exploit a modified version of the HSW theorem involving a family of two-universal hashing function. Our proof of Theorem 7 might not be simpler than the ones of [154, 152], but it is arguably more physical and intuitive.

On the other hand, as explained in Chapter 3, our method does provide a much simpler security proof for prepared-and-measured QKD protocols. One of the consequences of this simplification is that we were able to easily generalize an observation by [109] to improve the secret key generation rate of certain protocols. To achieve this highest key rate, we used the quantum de Finetti theorem. For lower key rate, we can use instead the Azuma's inequality for parameter's estimation. Considering finite QKD, the level of security is affected by the version of the proof that is used and it might be worthwhile to avoid the quantum de Finetti theorem when possible.

In Chapter 4, we introduced a QKD protocol using decoherence-free subspace. The main advantage of this protocol is that even though the encoding is done using multi-photon, it requires only single-photon measurement. It turns out that the measurements are not restricted to the encoding space, which implies that standard security proofs do not apply to the protocol. Using a *tagging* argument [73], we derive the security of the protocol. Although multi-photon QKD protocols are, in general, more affected by loss in the channel than single-photon ones, we observe that this problem can be circumvented in some interesting situations. We note that although this analysis was constructed for the robust four-photon protocol, it can be adapted for security proofs of other QKD protocols as well, particularly in situations in which Bob's measurement is not restricted to Alice's encoding space. For example, the robust three and two state protocols discussed in [24] and [25] could be shown secure using the method we have developed.

5.2 Open Questions

- Renner's bound for the security level using the quantum de Finetti theorem is not necessarily tight. There might be ways to achieve the highest key rate with a better security level.

- One of the huge successes of the security proof, based either on entanglement distillation or the uncertainty principle, is that it can tackle the problem of imperfect devices. We did a direct generalization of those proof methods. Is it possible to use our new tools to improve the known results relating to security proof of QKD with imperfect devices? In this thesis, we are not considering a particular case of imperfect devices. However, we have developed a framework of security proof where the concept of Shield is central. The next step would be to use the Shield to represent some imperfection in the devices. We have strong reason to believe that it is possible to use our results about parameter's estimation and how to build an extra Shield from dit error correction, to improve some known security proofs with imperfect devices.

- We don't know if single-photon protocol can achieve the channel capacity even by including two-way classical communication. Consider a depolarization channel such that $(1-p)\rho + \frac{p}{3}\sum_{a=x,y,z}\sigma_a\rho\sigma_a$. We know that there exists a single-photon protocol (the six-state protocol) that can achieve a positive secret key generation rate up to $p = 0.414$ [36]. We also know that the depolarization channel can be used for entanglement distillation up to $p = 0.5$ [20]. One well-known open question is if there exists a single-photon prepared-and-measured protocol that allows positive rate at higher p. A related question would be if there is a multi-photon prepared-and-measured protocol that allows positive rate for $p > 0.414$, and if it can be implemented experimentally.

Appendix A

Supplemental Information

A.1 Shannon Entropy and Classical Mutual Information

In a seminal paper on the theory of information [166], Shannon proposed a function to quantity information. Consider the problem of transmitting a string $\bar{s}$ built from n symbols represented by the alphabet Γ. It is assumed that the letters $\mu \in \Gamma$ follow the probability distribution p_μ and the different symbols in the string are identical and independently distributed. Shannon showed that the function $H(\Gamma) := \sum_{\mu \in \Gamma} p_\mu \log_2 p_\mu$ is asymptotically the maximum compression bit rate of the message. The term *assymptotic* refers to the limit $n \to \infty$. Shannon entropy of the whole string is defined as $H(\bar{s}) := H(\Gamma_1, \Gamma_2, ..., \Gamma_n)$ where Γ_i is the variable corresponding to picking the i^{th} symbol of the string $\bar{s}$. The function H was already well known in thermodynamics and statistical mechanics, but Shannon's innovation was to show that it was a measure of information. His result implies that the average amount of bits required to store the information of a string $\bar{s}$ is $H(\bar{s})$.

We can define the Shannon entropy of any variable X with probability distribution p_x to be

$$H(X) := \sum_x p_x \log_2 p_x. \tag{A.1}$$

Consider another variable Y with probability distribution p_y. The variable (X, Y) respects the joint probability distribution $p_{x,y}$. The conditional variable $X|Y$ represents the outcome X given Y and follows the probability distribution $p_{x|y}$. The Shannon entropy has the following important relation:

$$H(X, Y) = H(X) + H(Y|X). \tag{A.2}$$

Remark that X and Y are independent if and only if $H(X|Y) = H(X)$ (which is equivalent to $H(Y|X) = H(Y)$). In general, conditioning will never increase the Shannon entropy so that $H(X|Y) \leqslant H(X)$. Under a transformation described by $f(X)$, the entropy is also decreasing meaning that there is always equal or more information in the variable X than in $f(X)$. We therefore have that

$$H(f(X)) \leqslant H(X). \tag{A.3}$$

Another useful quantity is the mutual information

$$I(X,Y) := H(X) + H(Y) - H(X,Y). \tag{A.4}$$

It corresponds to the average amount of information that can be obtained about the variable X from another variable Y. It can be shown that

$$H(X,Y) := H(X) + H(Y) - H(X,Y) \geqslant 0 \tag{A.5}$$

Derivation of those properties and many more details about the Shannon entropy can be found in many great textbooks. Here is a suggested one [46].

A.2 Impossibility of Secret Key Distribution Using Classical Communication

The Shannon entropy can also be interpreted as a measure of uncertainty. Suppose that Alice and Bob share a string $\bar{s}$. From Eve's point of view, the values of that string follow some probability distribution that depends on Eve's a-priori information about that string. The Shannon entropy quantities the average amount of extra information that Eve would obtain by accessing the string $\bar{s}$. In other words, the Shannon entropy quantifies Eve's a-priori uncertainty about the string $\bar{s}$.

Consider the problem of distributing a one-time pad using classical communication. In light of the introduction to the Shannon entropy, one-time-pad is defined as a shared string of n-bits $\bar{s}$ between Alice and Bob such that $H(\bar{s})$ from Eve's perspective is maximal (i.e. $H(\bar{s}) = n$). The impossibility of distributing a one-time pad is an immediate consequence of the following theorem.

Theorem 9 *Alice and Bob cannot increase the Shannon entropy of their shared classical register using classical communication.*

Proof. Suppose that Alice and Bob share some string $\bar{s}$ of bits and that the Shannon entropy of $\bar{s}$ is $H(\bar{s})$ from Eve's perspective. In complete generality, suppose that after n rounds of two-way classical communication, Alice and Bob compute the string $\bar{x} = f(\bar{s}, M)$ where M is the message that Alice and Bob sent to each other and f is an arbitrary function. Eve can have a copy of M, so that the entropy of $\bar{x}$ from her perspective is $H(\bar{x}|M)$. Using Eq. A.2, A.3 and A.5, we obtain

$$\begin{aligned} H(\bar{x}|M) &= H(f(\bar{s}, M)|M) && \text{(A.6)} \\ &\leqslant H(\bar{s}, M|M) && \text{(A.7)} \\ &= H(\bar{s}, M) - H(M) && \text{(A.8)} \\ &\leqslant H(\bar{s}) && \text{(A.9)} \end{aligned}$$

■

This theorem has stronger consequences than just preventing the distribution of a perfect one-time pad (i.e. when $H(\bar{s}) = n$). Suppose that we are interested in distributing a key that respects some slightly weaker security conditions. For example, someone might want to obtain a key from which Eve has absolutely no information with very high probability, but she may have some or full information on the key with a very small probability. In this case, Alice and Bob must still increase the Shannon entropy of their shared register, which is impossible using classical communication by theorem 9.

Please note that Theorem 9 will fail if Alice and Bob are allowed to use quantum communication. QKD is one counter-example where Alice and Bob are able to increase the entropy of their shared register. One of the obvious reasons why the steps of the proof are not valid in the quantum case is that the quantum non-cloning theorem prevents Eve from making a copy of the message. Another intuitive argument on why the proof fails is that there exist pure quantum states that Alice and Bob could share, such that, even if Eve knew perfectly well what that state was, she would have no knowledge about the outcome of Alice and Bob's local measurements on that state. The state $|\Phi_d\rangle$ given by Equation 1.3 is such an example.

A.3 Schmidt Decomposition and Purification

A very useful tool in quantum information is the Schmidt decomposition:

Theorem 10 *Suppose $|\psi\rangle_{AB}$ is a pure state of a composite finite system, AB. Then there exist orthonormal states $|i\rangle_A$ for system A, and orthonormal states $|i\rangle_B$ for system B such that*

$$|\psi\rangle_{AB} = \sum_i \lambda_i |i\rangle_A |i\rangle_B. \tag{A.10}$$

where λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$.

A proof of this theorem can be found in [138]. We now use it to prove another theorem. Consider a quantum state ρ_A restricted to the finite Hilbert space of the system A. A *purification* of the state ρ_A is any pure state $|\psi\rangle_{AB}$ such that $\mathrm{Tr}_B[|\psi\rangle_{AB}\langle\psi|] = \rho_A$ for some system B.

Theorem 11 *Suppose that we have two purifications $|\psi\rangle_{AB}$ and $|\psi'\rangle_{AB}$ of the states ρ_A, those purifications are related by some unitary U_B which is restricted to the system B:*

$$|\psi'\rangle_{AB} = U_B|\psi\rangle_{AB} \tag{A.11}$$

Proof. Using the Schmidt decomposition theorem, we know that there exist the orthonormal states $|\mu_i\rangle_A$ and $|\mu_i\rangle_B$ such that

$$|\psi\rangle_{AB} = \sum_i \lambda_i |\mu_i\rangle_A |\mu_i\rangle_B. \tag{A.12}$$

for some coefficient $\lambda_i \geqslant 0$. Since

$$\mathrm{Tr}_B[|\psi\rangle_{AB}\langle\psi|] = \rho_A = \mathrm{Tr}_B[|\psi'\rangle_{AB}\langle\psi'|], \tag{A.13}$$

then

$$\mathrm{Tr}_B[|\psi'\rangle_{AB}\langle\psi'|] = \sum_i \lambda_i^2 |\mu_i\rangle_A\langle\mu_i|. \tag{A.14}$$

Since $|\psi'\rangle_{AB}$ is pure, then

$$|\psi'\rangle_{AB} = \sum_i \sqrt{\lambda_i} |\mu_i\rangle_A |\mu_i'\rangle_B \tag{A.15}$$

where $\sqrt{\lambda_i}|\mu_i'\rangle_B :=_A \langle\mu_i|\psi'\rangle_{AB}$. However, using the cycling property of trace,

$$\sqrt{\lambda_i \lambda_j}{}_B\langle \mu'_j | \mu'_i \rangle_B = \sqrt{\lambda_i \lambda_j} \mathrm{Tr}_B[|\psi'\rangle_{AB}\langle \psi' | \mu_j \rangle_A \langle \mu_i |] \tag{A.16}$$

$$= \sqrt{\lambda_i \lambda_j} \delta_{i,j}. \tag{A.17}$$

Until now, $|\mu'_i\rangle_B$ was well defined only if $\lambda_i \neq 0$. Note that $\{|\mu'_i\rangle_B \mid \lambda_i \neq 0\}$ is a set of orthonormal vectors that can be extended to an orthonormal basis $\{|\mu'_i\rangle_B \mid 0 \leqslant i \leqslant |B|-1\}$ of the system B. $|B|$ means the dimension of the finite system B. Consider also an orthonormal basis $\{|\mu_i\rangle_B \mid 0 \leqslant i \leqslant |B|-1\}$ that extends the set of orthonormal vectors $\{|\mu_i\rangle_B\}$ used for the Schmidt decomposition of the state $|\psi\rangle_{AB}$. Define $U_B = \sum_i |\mu'_i\rangle_B \langle \mu_i|$. Then,

$$|\psi'\rangle_{AB} = U_B |\psi\rangle_{AB} \tag{A.18}$$

■

A.4 Von-Neumann Entropy and Quantum Mutual Information

We are looking for a function that quantifies the information of a quantum state ρ. Since this function cannot depend on a choice of basis, then it must depend only on the eigenvalues λ_i of ρ. Thus, a direct generalization of the Shannon Entropy for quantum state can be given by

$$S(\rho) := \sum_i \lambda_i \log_2 \lambda_i. \tag{A.19}$$

It is called the Von-Neumann Entropy. To support this choice as a measure of information, it was shown [92, 162] that n copies of the quantum state ρ could be compressed with high fidelity (see Appendix A.5) into $nH(\rho)$ qubits. The quantum mutual information is defined in analogy to the classical case. Consider a quantum state ρ_{AB}:

$$I(A{:}B) := S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{A.20}$$

Suppose that the state in system A is classical (meaning that it can be written as $\rho_A = \sum_k p_k |k\rangle_A\langle k|$ for some non-negative real number p_k such $\sum_k p_k = 1$). In that case, ρ_{AB} is a classical-quantum state (i.e $\rho_{AB} = \sum_k p_k |k\rangle_A\langle k| \otimes \rho_B^{(k)}$ for some $\rho_B^{(k)}$) and

$$\begin{aligned} I(A{:}B) &= I(K{:}B) & \text{(A.21)}\\ &= H(p_k) + S(\rho_B) - S(\sum_k p_k|k\rangle_A\langle k| \otimes \rho_B^{(k)}) & \text{(A.22)}\\ &= S(\rho_B) - \sum_k p_k S(\rho_B^{(k)}). & \text{(A.23)}\end{aligned}$$

Holevo showed that if Alice and Bob share a cq-state $\rho_{AB} = \sum_k p_k|k\rangle_A\langle k| \otimes \rho_B^{(k)}$, where Alice's state is represented by the random variable K, and if Bob performs any quantum measurement on his part of the state with outcome Y, then

$$I(X:Y) \leqslant S(\rho_B) - \sum_k p_k S(\rho_B^{(k)}). \tag{A.24}$$

This is known as the *Holevo's bound*. It can be achieved in certain non-trivial situations, as implied by the *HSW theorem*:

Theorem 12 *([83, 164, 50]) Suppose that Alice and Bob share n copies of*

$$\sum_{k=0}^{d-1} p_k |k\rangle_A\langle k| \otimes \rho_B^{(k)}. \tag{A.25}$$

If Alice measures k and randomly picks a subset $X \in \{0,1,...,d-1\}^n$ such that $k \in X$ and $|X| \cong nS(\rho_B) - n\sum_k p_k S(\rho_B^{(k)})$ then there exists a measurement that Bob, given X, can perform to predict k with a probability $\epsilon_n \to 0$ exponentially as $n \to \infty$.

A.5 Fidelity

In this section, we introduce another measure of distance between two quantum states ρ and σ. The *fidelity* F is defined as

$$F(\rho, \sigma) := \mathrm{Tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}]. \tag{A.26}$$

Suppose that σ is a pure state (i.e. $\sigma = |\psi\rangle\langle\psi|$), then

$$F(\rho, |\psi\rangle) := \sqrt{\langle\psi|\rho|\psi\rangle}. \tag{A.27}$$

In contrast to the trace distance, F is not a metric. Here is a theorem [138] that connects the trace distance to the fidelity:

Theorem 13 *Consider any two density matrices ρ and σ, then*

$$\frac{1}{2}\mathrm{Tr}|\rho - \sigma| \leqslant \sqrt{1 - F(\rho, \sigma)^2}. \tag{A.28}$$

Here is another very important result in quantum information involving the fidelity and purification, the *Uhlmann's theorem*:

Theorem 14 *[138] Consider the density matrices ρ and σ in a quantum system Q. Introduce a second system R which is a copy of Q. Then*

$$F(\rho, \sigma) = max_{\{|\psi\rangle, |\phi\rangle\}} |\langle\psi|\phi\rangle| \tag{A.29}$$

where the maximization is over all purifications $|\psi\rangle$ of ρ and $|\phi\rangle$ of σ into RQ.

A.6 Generalized Pauli Operator

In this section, we define the operator X and Z in a d-dimensional space [71]. Consider an orthonormal computational basis $|k\rangle$ for $0 \leqslant k \leqslant d-1$. Define

$$Z := \sum_{k=0}^{d-1} e^{\frac{2\pi ki}{d}} |k\rangle\langle k| \tag{A.30}$$

and a complementary orthonormal basis

$$|\widetilde{x}\rangle := \sum_{k=0}^{d-1} Z^x |k\rangle. \tag{A.31}$$

Remark that $Z|\widetilde{x}\rangle = |\widetilde{x} + \widetilde{1}\rangle$, where sum modulo d is implicitly understood. Define

$$X := \sum_{x=0}^{d-1} e^{\frac{-2\pi xi}{d}} |\widetilde{x}\rangle\langle\widetilde{x}|. \tag{A.32}$$

Then

$$\begin{aligned}
|k\rangle &:= \sum_{x=0}^{d-1} |\widetilde{x}\rangle\langle\widetilde{x}|k\rangle && \text{(A.33)}\\
&= \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} (\sum_{k'=0}^{d-1} \langle k'|Z^{\dagger x}|k\rangle)|\widetilde{x}\rangle && \text{(A.34)}\\
&= \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} e^{-\frac{2\pi xki}{d}} |\widetilde{x}\rangle && \text{(A.35)}\\
&= \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} X^k |\widetilde{x}\rangle. && \text{(A.36)}
\end{aligned}$$

Therefore, $X|k\rangle = |k+1\rangle$.

Since $ZX|k\rangle = e^{\frac{2\pi i(k+1)}{d}} |k+1\rangle = e^{\frac{2\pi i}{d}} XZ|k\rangle$, then

$$ZX = e^{\frac{2\pi i}{d}} XZ. \tag{A.37}$$

A.7 Overview of Classical Error Correction

In this thesis, we are interested in classical error correction codes that are deterministic and require only one-way communication from Alice to Bob. In that case, we can describe the classical error correction procedure as follow.

Suppose that Alice want to send one classical message of size d^m to Bob through a memoryless noisy classical channel $\xi : \{0, ..., d-1\} \rightarrow \{0, ..., d-1\}$ where ξ is a probabilistic function. Alice can encode her message using a function $g : \{0, ..., d-1\}^m \rightarrow C$ where $C \subseteq \{0, ..., d-1\}^n$ (i.e. C is composed of n-dit strings), $n \geqslant m$ and $|C| = d^m$. The choice of the set C of codewords, with the recovery function, define the error correction code. Bob will receive $\xi(g(c))$, and will use the recovery function $r^{(C)} : \{0, ..., d-1\}^n \rightarrow C$. With high probability, $r^{(C)}(\xi(g(c))) = c$. The codes used in Appendix B.2 that are based on two-universal family of Hashing functions are examples of one-way classical error correction code.

Now we will see how classical error correction can be used to purify entangled state. Define

$$|\Phi^{j_1}\rangle_{A_1B_1} = \frac{X_B^{j_1}}{\sqrt{d}} \sum_{k=0}^{d-1} |k, k\rangle_{A_1B_1} \tag{A.38}$$

where X is defined in Appendix A.6. Suppose that Alice and Bob share n noisy copies of the maximally entangled state:

$$(\sum_{j_1=0}^{d-1} p_{j_1} |\Phi^{j_1}\rangle_{A_1B_1}\langle\Phi^{j_1}|)^{\otimes n}. \tag{A.39}$$

for some probability distribution p_{j_1}. Define the superoperator

$$\xi_B(\cdot) = \sum_{j\in\{0,1\}^n} p_j X_B^j \cdot X_B^j, \tag{A.40}$$

where $p_j = p^{j_1}...p^{j_n}$, $X^j = X^{j_1} \otimes ... \otimes X^{j_n}$ and $\mathbb{1}$ is the identity operator. Alice and Bob share

$$\xi_B(\rho_{AB}^{(0)}). \tag{A.41}$$

where $\rho_{AB}^{(0)} = (|\Phi^0\rangle_{A_1B_1}\langle\Phi^0|)^{\otimes n}$. Let $P_A^{(C)}$ be the projection into the codewords on a classical error correction code:

$$P_A^{(C)} = \sum_{c\in C} |c\rangle_A\langle c| \tag{A.42}$$

where $|c\rangle_A = |c_1\rangle_{A_1}...|c_n\rangle_{A_n}$. Since $P_A^{(C)}\rho_{AB}^{(0)}P_A^{(C)} = P_B^{(C)}\rho_{AB}^{(0)}P_B^{(C)}$, then

$$P_A^{(C)}\xi_B(\rho_{AB}^{(0)})P_A^{(C)} \quad = \quad \xi_B(P_B^{(C)}\rho_{AB}^{(0)}P_B^{(C)}) \tag{A.43}$$

By the definition of error correction code, there exists an recovery operator $R^{(C)}$ and a system B', such that

$$\mathrm{Tr}|\mathrm{Tr}_{B'}[R_{BB'}^{(C)}\xi_B(P_B^{(C)}\rho_{AB}^{(0)} \otimes |0\rangle_{B'}\langle 0|)R_{BB'}^{(C)\dagger}] - \frac{1}{d^n}\sum_{c\in C}|c,c\rangle_{AB}\langle c,c|| \leqslant d^{m-n}\epsilon_C \tag{A.44}$$

where ϵ_C is the probability of failure of the classical error correction code averaged over all the codewords in C. The recovery operation is an operator that follows the relation $R_{BB'}^{(c)}|x\rangle_B|0\rangle_{B'} = |r^{(C)}(x)\rangle_B|x - r^{(C)}(x)\rangle_{B'}$ for all $x \in \{0,...,d-1\}^n$.

Appendix B

Privacy Amplification and Private State Distillation

B.1 Two-Universal Family of Hashing Functions and Privacy amplification

A *two-universal family of hashing functions* is defined as a distribution of functions $f : X \rightarrow T$ that respect

$$Pr[f(x) = f(x')] \leq \frac{1}{|T|} \tag{B.1}$$

for any distinct $x, x' \in X$. It was proved that any two-universal family of hashing functions could be used for privacy amplification [154, 153, 152]. Privacy amplification proceeds as follows: if Alice and Bob share some uniformly-distributed error-free key k on which Eve possesses some information, then Alice and Bob can agree publicly on a random function f picked from a distribution of two-universal family of hashing functions and compute $f(k)$. If $|T|$ is chosen small enough depending on the amount of correlation Eve could have with the key, then the new key $f(k)$ is secure in the sense that $\exists \rho_E$ such that

$$\mathrm{Tr}[\sum_k |f(k)\rangle_{AB}\langle f(k)| \otimes \rho_E^{(k)} - \sum_k |f(k)\rangle_{AB}\langle f(k)| \otimes \rho_E] \leqslant 2\epsilon|X| \tag{B.2}$$

for some very small ϵ and where $|f(k)\rangle_{AB} := |f(k), f(k)\rangle_{AB}$. If $f(k)$ is uniformly-distributed, then we obtain the definition of security given in section 1.5.4. Suppose that $X = \{0, 1, ..., d-1\}^n$ and $T = \{0, 1, ..., d-1\}^m$ where $m < n$, here are two examples of two-universal family of hashing functions for which each f is uniformly-distributed:

1. The uniform distrubtion of all the functions $f : X \rightarrow T$ such that there is exactly d^{n-m} values of k such that $f(k) = c$, for all $c \in T$.

2. Consider all m by n matrices M of rank m, with entries $\{0, 1, ..., d-1\}$. Associate each M to a function $f_M(k) = Mk$, where the multiplication and sum are modulo d. The uniform distribution of those functions is two-universal.

The proof that the first example is a two-universal family of hashing functions is trivial. To see that the second example is also two-universal, consider two strings k and k' of n dits such that $k \neq k'$. Pick a random n dits string $s^{(1)}$ (with the restriction that $s^{(1)}$ is not the zero string). The probability that $s^{(1)} \cdot k = s^{(1)} \cdot k'$ is exactly $\frac{1}{d}$. Pick another string $s^{(2)}$, randomly over all strings that are linearly independent of the previous string $s^{(1)}$. Suppose that $s^{(1)} \cdot k \neq s^{(1)} \cdot k'$ then the probability that $s^{(2)} \cdot k = s^{(2)} \cdot k'$ is smaller than $\frac{1}{d}$. By strong induction, we can prove that if $s^{(i)} \cdot k \neq s^{(i)} \cdot k'$ for $1 \leqslant i \leqslant \ell$ and if $s^{(\ell+1)}$ is picked randomly over all strings that are linearly independent of the strings $s^{(i)}$ for $1 \leqslant i \leqslant \ell$, then the probability that $s^{(\ell+1)} \cdot k = s^{(\ell+1)} \cdot k'$ is smaller than $\frac{1}{d}$. Therefore, consider the matrix M where the i^{th} row is $s^{(i)}$ for $1 \leqslant i \leqslant m$. Then the probability that $Mk = Mk'$ is smaller than $\frac{1}{d^m} = \frac{1}{|T|}$. This concludes our sketch of the proof.

B.2 HSW theorem for Two-Universal Family of Hashing Functions

In this section, we generalize the HSW theorem. In [83, 164], they prove Theorem 12 using random code. We establish the same result using other possible codes that are related to the two-universal family of hashing functions. A relation between the following results and privacy amplification is given in appendix B.3 for some specific two-universal family of hashing functions.

Consider the following scenario. Let p_i be some probability distribution and σ_i some normalized density matrices for $i \in \{0, 1, ..., d-1\}$. Suppose Alice and Bob initially share the state

$$\sum_{x\in\{0,1,...,d-1\}^n} \hat{p}_x |x\rangle_A\langle x| \otimes \hat{\sigma}_B^x \tag{B.3}$$

where $\hat{\sigma}^x = \sigma_{x_1} \otimes \sigma_{x_2} \otimes ... \otimes \sigma_{x_n}$ and $\hat{p}_x = \Pi_{i=1}^n p_{x_i}$ for $1 \leq x_i \leq d$. Let $\hat{\sigma} := (\sum_{i=0}^{d-1} p_i\sigma_i)^{\otimes n}$. Alice measures her share of the state in the computational basis. With probability higher than $1-\epsilon$, the measurement result v is typical (i.e. v is typical iff $2^{-nH(p_i)-n\delta} \leq p_v \leq 2^{-nH(p_i)+n\delta}$) where $\epsilon = e^{-\frac{n\delta^2}{2}}$. This follows directly form the law of large numbers.

Let Q and Q^x be the projection into the typical subspace of $\hat{\sigma}$ and $\hat{\sigma}^x$, respectively. It follows that (see [83, 46] for details):

$$\mathrm{Tr}[\hat{\sigma}(I-Q)] \leqslant \epsilon, \tag{B.4}$$

$$\mathrm{Tr}[\hat{\sigma}^x(I-Q^x)] \leqslant \epsilon, \tag{B.5}$$

$$\mathrm{Tr}[Q\hat{\sigma}Q\hat{\sigma}] \leqslant ||Q\hat{\sigma}Q||_\infty \mathrm{Tr}[\hat{\sigma}] = ||Q\hat{\sigma}Q||_\infty, \tag{B.6}$$

$$Q^x \leqslant 2^{n\sum_i p_i S(\sigma_i)+n\delta}\hat{\sigma}^x, \tag{B.7}$$

$$\sum_{x \text{ typical}} \hat{\sigma}^x \leqslant 2^{nH(p_i)+n\delta}\hat{\sigma} \tag{B.8}$$

and

$$||Q\hat{\sigma}Q||_\infty \leqslant 2^{-S(\hat{\sigma})+n\delta}. \tag{B.9}$$

where $||M||_\infty$ is the infinite norm, and it is equal to the maximal eigenvalue of M. The next step consists of Alice randomly picking a function f from a distribution that is two-universal. The definition of a two-universal family of hashing functions is given by a distribution of functions $f : X \rightarrow T$ that respect $Pr[f(x) = f(x')] \leq \frac{1}{|T|}$ for any distinct $x, x' \in X$. She sends $f(v)$ to Bob through the public channel. Bob applies the pretty good measurement defined by Eq. 11 in [83] on his part of the quantum state. This measurement depends on a set of code words. In our case, the code words are all the *typical* strings x such that $f(x) = f(v)$. In general, the number of code words is not fixed.

A bound for the average error probability is given by Eq. 17 of [83]. Assuming v is typical, a similar formula for the conditional error probability — but without averaging over v— can be derived using exactly the same arguments as Holevo:

$$\begin{aligned} P_{er}(v) \quad &\leq \quad 3\mathrm{Tr}[\hat{\sigma}^v(I-Q)] + \mathrm{Tr}[\hat{\sigma}^v(I-Q^v)] \\ &\quad + \sum_{x:x\neq v, f(x)=f(v)} \mathrm{Tr}[Q\hat{\sigma}^v QQ^x] \end{aligned} \tag{B.10}$$

where the sum is made only over typical strings since those are our code words. In the unlikely case where v is not typical, we use $P_{er}(v) = 1$.

Suppose that v is typical, and using Equation B.5:

$$\begin{aligned} P_{er}(v) \quad &\leq \quad 3\mathrm{Tr}[\hat{\sigma}^v(I-Q)] + \epsilon \\ &\quad + \sum_{x:f(x)=f(v)} \mathrm{Tr}[Q\hat{\sigma}^v QQ^x]. \end{aligned}$$

We are however interested by the average error probability:

$$\begin{aligned} \langle P_{er}(v)\rangle_{v,f} \quad &\leq \quad 3\mathrm{Tr}[\hat{\sigma}(I-Q)] + 2\epsilon \\ &\quad + \langle \sum_{x:f(x)=f(v)} \mathrm{Tr}[Q\hat{\sigma}^v QQ^x]\rangle_{v,f} \end{aligned}$$

where an extra ϵ appear since the probability of v to be non-typical is less then ϵ. The sum over x is always restricted x typical. The average over v is made over all v, typical and not typical. Using Equation B.4:

$$\begin{aligned}
\langle P_{er}(v)\rangle_{v,f} &\leq 5\epsilon + \langle \sum_{x:f(x)=f(v)} \mathrm{Tr}[Q\hat{\sigma}^v QQ^x]\rangle_{v,f} \\
&= 5\epsilon + \langle \sum_{x \text{ typical}} \Pr[f(x)=f(v)]\mathrm{Tr}[Q\hat{\sigma}^v QQ^x]\rangle_v \\
&\leq 5\epsilon + \frac{1}{|T|}\langle \sum_{x \text{ typical}} Tr[Q\hat{\sigma}^v QQ^x]\rangle_v \\
&= 5\epsilon + \frac{1}{|T|} \sum_{x \text{ typical}} \mathrm{Tr}[Q\hat{\sigma}QQ^x] \\
&\leq 5\epsilon + \frac{2^{n\sum_i p_i S(\sigma_i)+n\delta}}{|T|} \sum_{x \text{ typical}} \mathrm{Tr}[Q\hat{\sigma}Q\hat{\sigma}^x] \\
&\leq 5\epsilon + \frac{2^{nH(p_i)+n\sum_i p_i S(\sigma_i)+2n\delta}}{|T|}\mathrm{Tr}[Q\hat{\sigma}Q\hat{\sigma}] \\
&\leq 5\epsilon + \frac{2^{nH(p_i)+n\sum_i p_i S(\sigma_i)+2n\delta}}{|T|}||Q\hat{\sigma}Q||_\infty \\
&\leq 5\epsilon + \frac{2^{nH(p_i)-nS(\hat{\sigma})+n\sum_i p_i S(\sigma_i)+3n\delta}}{|T|}
\end{aligned}$$

where we used Equations B.6, B.7, B.8 and B.9. The average error probability is exponentially small if

$$\begin{aligned}
|T| &\geq 2^{nH(p_i)-nS(\hat{\sigma})+n\sum_i p_i S(\sigma_i)+3n\delta} && \text{(B.11)} \\
&= 2^{nH(p_i)-nI(X_1:B_1S_1)+3n\delta}. && \text{(B.12)}
\end{aligned}$$

Remark that in Chapter 2, we are interested in the case where variable x is uniformly-distributed. Remark also that instead of sharing n copies of a state, we could suppose that Alice and Bob have only one copy (take $n=1$ in the above derivation). In that case, we would have found that

$$\begin{aligned}
\langle P_{er}(v)\rangle_{v,f} &\leq \frac{1}{|T|}\sum_x \mathrm{Tr}[\hat{\sigma}Q^x] && \text{(B.13)} \\
&\leq \frac{1}{|T|}\sum_x ||\hat{\sigma}||_\infty \mathrm{Tr}[Q^x] && \text{(B.14)} \\
&= \frac{2^{n\log_2 d-nS(\hat{\sigma})+n\sum_i p_i S(\sigma_i)}}{|T|} && \text{(B.15)} \\
&= \frac{2^{n\log_2 d-nI(X_1:B_1S_1)}}{|T|} && \text{(B.16)}
\end{aligned}$$

where Q^x is the projector into the support of $\hat{\sigma}^x$. An alternative way of writing Equation B.14 is

$$\begin{aligned} \langle P_{er}(v)\rangle_{v,f} &\leq \frac{1}{|T|}\sum_x ||\hat{\sigma}||_\infty \mathrm{Tr}[Q^x] && \text{(B.17)} \\ &= \frac{\sum_x 2^{S_0(\hat{\sigma}^x)-S_\infty(\hat{\sigma})}}{|T|}. && \text{(B.18)} \end{aligned}$$

In the last step, we used the definition of the *Rényi* entropy [154] (i.e. $S_0(\rho) := \log_2[\mathrm{Rank}(\rho)]$ and $S_\infty(\rho) := -\log_2[\max_i(\lambda_i)]$ where λ_i are the eigenvalues of ρ).

B.3 Proof of Theorem 5

Consider a purification of the state given in Equation 2.20:

$$|\psi\rangle_{ABTSE} = |\psi\rangle^{\otimes n}_{A_1B_1T_1S_1E_1}, \tag{B.19}$$

where

$$|\psi\rangle_{A_1B_1T_1S_1E_1} = \frac{1}{\sqrt{d}} \sum_{k_1=0}^{d-1} |k_1\rangle_{A_1} |k_1\rangle_{B_1} |0\rangle_{T_1} |\phi^{(k_1)}\rangle_{S_1E_1}. \tag{B.20}$$

The system T contains extra ancillas initialized to the zero state. We can write

$$|\psi\rangle_{A_1B_1T_1S_1E_1} = \frac{1}{d} \sum_{x_1=0}^{d} |\widetilde{x_1}\rangle_{A_1} Z^{-x_1}_{B_1} \sum_{k_1=0}^{d} |k_1\rangle_{B_1} |0\rangle_{T_1} |\phi^{(k_1)}\rangle_{S_1E_1}. \tag{B.21}$$

where the basis $|\widetilde{x}\rangle$ is defined by Equation 2.4. Z is the generalized pauli operator with eigenvectors $\{|k\rangle\}$ for $0 \leqslant k \leqslant d-1$. Note that $Z^{-x_1} = Z^{d-x_1}$. Alice and Bob agree on a random linear function f of rank m (i.e. for each f, there exists a m by n matrix M of rank m such that $f_M(k) = Mk$ for all k; see the second example in appendix B.1 for more details). For each M, there exists n by $n-m$ matrices $M^{\perp}$ of rank $n-m$ such that $MM^{\perp} = 0$. Alice chooses at random one such matrix $M^{\perp}$. Note that, after averaging over all f, the distribution of functions

$$g(\widetilde{x}) := (M^{\perp})^{\dagger}\widetilde{x} \tag{B.22}$$

is two-universal.

The following measurements are not actually done by Alice and Bob in the real privacy amplification protocol. However, we assume that they actually performed them, and at the end of the section, we will explain why those extra operations are not necessary. Suppose that Alice measures the observables $X^{M^{\perp}_{1j}} \otimes ... \otimes X^{M^{\perp}_{ij}} \otimes ... \otimes X^{M^{\perp}_{nj}}$ for $1 \leqslant j \leqslant n$ where $X := \sum_{x_1=0}^{d-1} e^{\frac{-2\pi i x_1}{d}} |\widetilde{x_1}\rangle\langle\widetilde{x_1}|$. This is similar to measuring $g(\widetilde{x})$. Suppose that Alice sends her outcome c_A to Bob. By the modified HSW theorem for two-universal family of Hashing function that we proved in appendix B.2, we obtain that Bob can guess $\widetilde{x}$ with a probability of error

$$P_{er} \leqslant d^m 2^{-nI(X_1:B_1S_1)} + 3n\delta. \tag{B.23}$$

Since any POVM can be extended to an unitary on a larger space (i.e. in this case, we extend the space with the system T), $\exists$ an unitary W_{BTS} such that

$$\frac{1}{d^{m+\frac{n}{2}}} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}_{TSE}[\ {}_B\langle\widetilde{x}|W_{BTS}Z_B^{-x}(\sum_{k_1=0}^{d-1} |k_1\rangle_{B_1}|0\rangle_{T_1}|\phi^{(k_1)}\rangle_{S_1E_1})^{\otimes n}] \ \geq \ 1 - P_{er} \tag{B.24}$$

where the sum is over the $x \in \{0, 1, ..., d-1\}^n$ outcomes that are compatible with Alice's announcement c_A, and $Z^{-x} = Z^{-x_1} \otimes ... \otimes Z^{-x_n}$ (Z is the generalized pauli operator on a space of dimension d with eigenkets $|k\rangle$ for $0 \leqslant k \leqslant d-1$).

Define

$$\begin{aligned} P_{AB} &:= \sum_{k \in \{0,...,d-1\}^n} |k, k\rangle_{AB}\langle k, k| && \text{(B.25)} \\ &= \sum_{x \in \{0,...,d-1\}^n} |\widetilde{x}, -\widetilde{x}\rangle_{AB}\langle \widetilde{x}, -\widetilde{x}|, && \text{(B.26)} \end{aligned}$$

so

$$P_{AB}W_{BST}P_{AB} = \sum_{x \in \{0,...,d-1\}^n} |\widetilde{x}, -\widetilde{x}\rangle_{AB}\langle \widetilde{x}, -\widetilde{x}| \otimes V_{ST}^{(x)} \qquad \text{(B.27)}$$

where the $V_{ST}^{(x)} =_B \langle -\widetilde{x}| W_{BST} |-\widetilde{x}\rangle_B$ are not necessarily unitary. However, $W_{BST}^\dagger P_{AB} W_{BST}$ is a projector. Therefore,

$$P_{AB}W_{BST}^\dagger P_{AB}W_{BST}P_{AB} \qquad \text{(B.28)}$$

must be diagonal. It implies that $V_{ST}^{(x)\dagger}V_{ST}^{(x)}$ is diagonal (and $\leqslant \mathbb{1}$) for every x. Similarly, $V_{ST}^{(x)}V_{ST}^{(x)\dagger}$ is diagonal (and $\leqslant \mathbb{1}$) for every x. Define $S_{1,x}$ and $S_{2,x}$ as the smallest space that support $V_{ST}^{(x)\dagger}V_{ST}^{(x)}$ and $V_{ST}^{(x)}V_{ST}^{(x)\dagger}$, respectively. Therefore, we can write

$$V_{ST}^{(x)} = \sum_i \alpha_i^{(x)} |i\rangle_{S_{2,x}} \;{}_{S_{1,x}}\langle i| \qquad \text{(B.29)}$$

for some complex number $\alpha_i^{(x)}$, and orthonormal states $\{|i\rangle_{S_{1,x}}\}$ and $\{|i\rangle_{S_{2,x}}\}$ that depend on x. Let $S_{j,x}^\perp$ be the space orthogonal to $S_{j,x}$ such that $S_{j,x}^\perp \bigcup S_{j,x} = ST$. By counting the dimensions, it is easy to see that there exists an invertible transformation $Q_{ST}^{(x)} : S_{1,x}^\perp \rightarrow S_{2,x}^\perp$ such that $Q_{ST}^{(x)\dagger}Q_{ST}^{(x)} = \mathbb{1}_{S_{1,x}}$. Let $R_{ST}^{(x)} : S_{1,x} \rightarrow S_{2,x}$ such that $R_{ST}^{(x)} = \sum_i \frac{\alpha_i^{(x)}}{|\alpha_i^{(x)}|} |i\rangle_{S_{2,x}} \;{}_{S_{1,x}}\langle i| - V_{ST}^{(x)}$. Consequently, $\hat{V}_{ST}^{(x)} := V_{ST}^{(x)} + R_{ST}^{(x)} + Q_{ST}^{(x)}$ is unitary.

Define

$$|\Phi\rangle_{AB} := \frac{1}{\sqrt{d^m}} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} |\widetilde{x}, -\widetilde{x}\rangle_{AB} \qquad \text{(B.30)}$$

and

$$|\phi\rangle_{STE} :=_{AB} \langle \Phi| \sum_{x \in \{0,...,d-1\}^n} |\widetilde{x}, -\widetilde{x}\rangle_{AB}\langle \widetilde{x}, -\widetilde{x}| \otimes \hat{V}_{ST}^{(x)} |\psi\rangle_{ABSTE}. \qquad \text{(B.31)}$$

Our next step is to find a bound on $\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|]$:

Lemma 2

$$\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|] \geqslant 1 - P_{er} \tag{B.32}$$

Proof.

$$\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|] = \frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[{}_B\langle -\widetilde{x}|\hat{V}_{ST}^{(x)} Z_B^{-x}|\psi_x\rangle_{BSTE}\langle\psi_x|Z_B^x \hat{V}_{ST}^{(x)\dagger}|-\widetilde{x}\rangle_B] \tag{B.33}$$

where $|\psi_x\rangle_{BSTE} := Z_B^x \ {}_A\langle\widetilde{x}|\psi\rangle_{ABSTE}$.

$$\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|] \ = \ \frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[\hat{V}_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|\hat{V}_{ST}^{(x)\dagger}] \tag{B.34}$$

where $|\psi_{x,0}\rangle_{STE} :=_B \langle\widetilde{0}|\psi_x\rangle_{BSTE}$ is not normalized. Using the facts that

$$\mathrm{Tr}[Q_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|Q_{ST}^{(x)\dagger}] = \mathrm{Tr}[|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|Q_{ST}^{(x)\dagger}Q_{ST}^{(x)}] \geqslant 0, \tag{B.35}$$

$$\mathrm{Tr}[Q_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|(V_{ST}^{(x)} + R_{ST}^{(x)})_{ST}^{\dagger}] = \mathrm{Tr}[|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|(V_{ST}^{(x)} + R_{ST}^{(x)})_{ST}^{\dagger}Q_{ST}^{(x)}] = 0 \tag{B.36}$$

and

$$\mathrm{Tr}[(V_{ST}^{(x)} + R_{ST}^{(x)})|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|Q_{ST}^{(x)\dagger}] = \mathrm{Tr}[|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|Q_{ST}^{(x)\dagger}(V_{ST}^{(x)} + R_{ST}^{(x)})] = 0, \tag{B.37}$$

we obtain

$$\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|] \ \geqslant \ \frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[(V_{ST}^{(x)} + R_{ST}^{(x)})|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|(V_{ST}^{(x)} + R_{ST}^{(x)})^{\dagger}] \tag{B.38}$$

Since the $R_{ST}^{(x)\dagger}V_{ST}^{(x)}$ and $V_{ST}^{(x)\dagger}R_{ST}^{(x)}$ are diagonal with real non-negative entries, then

$$\mathrm{Tr}[V_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|R_{ST}^{(x)\dagger}] = \mathrm{Tr}[|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|R_{ST}^{(x)\dagger}V_{ST}^{(x)}] \geqslant 0, \tag{B.39}$$

and

$$\mathrm{Tr}[R_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|V_{ST}^{(x)\dagger}] = \mathrm{Tr}[|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|V_{ST}^{(x)\dagger}R_{ST}^{(x)}] \geqslant 0. \tag{B.40}$$

Therefore,

$$\mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|] \geqslant \frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[V_{ST}^{(x)}|\psi_{x,0}\rangle_{STE}\langle\psi_{x,0}|V^{(x)\dagger}]. \tag{B.41}$$

We conclude by observing that

$$\begin{aligned} &\frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[V_{ST}^{(x)}|\psi\rangle_{STE}\langle\psi|V^{(x)\dagger}] \\ &\quad = \frac{1}{d^m} \sum_{x \ s.t. \ g(\widetilde{x})=c_A} \mathrm{Tr}[{}_B\langle -\widetilde{x}|W_{BST}Z_B^{-x}|\psi\rangle_{BSTE}\langle\psi|Z_B^{x}W_{BST}^{\dagger}|-\widetilde{x}\rangle_B] \\ &\quad \geqslant (1-P_{er}) \end{aligned} \tag{B.42}$$

where we used Equation B.24. ■

Define

$$\rho := Tr_{STE}[\hat{W}_{ABST}|\psi\rangle_{ABSTE}\langle\psi|\hat{W}_{ABST}^{\dagger}]$$

where

$$\hat{W}_{ABST} := \sum_{x\in\{0,\ldots,d-1\}^n} |\widetilde{x},-\widetilde{x}\rangle_{AB}\langle\widetilde{x},-\widetilde{x}| \otimes \hat{V}_{ST}^{(x)}.$$

By the Uhlmann's Theorem 14, $F(\rho, |\Phi\rangle_{AB}\langle\Phi|) = \max_{\{|\psi'\rangle,|\phi'\rangle\}}|\langle\psi'|\phi'\rangle|$ where $|\psi'\rangle$ is a purification of ρ and $|\phi'\rangle$ is a purification of $|\Phi\rangle_{AB}$. Since all purifications of a state are related by unitaries (Theorem 11) and that the Fidelity is invariant under unitaries (i.e. $F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho,\sigma)$ for any unitary U [138]), then we might as well fix $|\psi'\rangle$ and maximize over $|\phi'\rangle$ only.

Consider $F(|\Phi\rangle_{AB},\rho)$ and fix $|\psi'\rangle := \hat{W}_{BST}|\psi\rangle_{ABSTE}$. By the Uhlmann's theorem 14, $F(\rho,|\Phi\rangle_{AB}\langle\Phi|) = \max_{\{|\phi'\rangle\}}|\langle\psi'|\phi'\rangle|$. Since $|\Phi\rangle_{AB}$ is a pure state, then the maximum is obtained at $|\phi'_{max}\rangle = |\Phi\rangle_{AB}|u\rangle_{STE}$ for some pure state $|u\rangle_{STE}$. Since each $\hat{V}_{ST}^{(x)}$ is unitary over its domain, it is therefore possible to extend them to unitaries $\hat{V}_{ST}^{\prime(x)}$ over larger domains, say the combined system ST. Define

$$\hat{W}'_{ABST} := \sum_{x\in\{0,\ldots,d-1\}^n} |\widetilde{x},-\widetilde{x}\rangle_{AB}\langle\widetilde{x},-\widetilde{x}| \otimes \hat{V}_{ST}^{\prime(x)}$$

so that

$$\begin{aligned} &\mathrm{Tr}[|\hat{W}_{ABST}^{\prime\dagger}|\Phi\rangle_{AB}|u\rangle_{STE}\langle u| \ {}_{AB}\langle\Phi|\hat{W}'_{ABST} - |\psi\rangle_{ABSTE}\langle\psi||] \\ = &\quad \mathrm{Tr}[||\Phi\rangle_{AB}|u\rangle_{STE}\langle u| \ {}_{AB}\langle\Phi| - \hat{W}'_{ABST}|\psi\rangle_{ABSTE}\langle\psi|\hat{W}_{ABST}^{\prime\dagger}|] \\ = &\quad \mathrm{Tr}[||\Phi\rangle_{AB}|u\rangle_{STE}\langle u| \ {}_{AB}\langle\Phi| - \hat{W}_{ABST}|\psi\rangle_{ABSTE}\langle\psi|\hat{W}_{ABST}^{\dagger}|]. \end{aligned}$$

where we used the fact that the trace distance is invariant under the action of a joint unitary. Using Theorem 13,

$$\begin{aligned}\mathrm{Tr}[||\Phi\rangle_{AB}|u\rangle_{STE}\langle u|\ {}_{AB}\langle\Phi| - \hat{W}_{ABST}|\psi\rangle_{ABSTE}\langle\psi|\hat{W}^{\dagger}_{ABST}|] \\ &\leqslant 2\sqrt{1 - F(|\Phi\rangle_{AB}|u\rangle_{STE}, \hat{W}_{BST}|\psi\rangle_{ABSTE})^2} \\ &\leqslant 2\sqrt{1 - F(|\Phi\rangle_{AB}, \rho)^2} \\ &= 2\sqrt{1 - \mathrm{Tr}[|\phi\rangle_{STE}\langle\phi|]} \\ &\leqslant 2\sqrt{P_{er}}. \end{aligned} \tag{B.43}$$

Remark that $\hat{W}'^{\dagger}_{BST}|\Phi\rangle_{AB}|u\rangle_{STE}$ is a private state. Therefore, it is possible to extract an $\sqrt{P_{er}}$-secure key from $|\psi\rangle_{ABSTE}$. From Equation B.23,

$$P_{er} \leqslant d^m 2^{-nI(X_1:B_1S_1)+3n\delta}. \tag{B.44}$$

for some arbitrary factor δ. Pick $m = \lceil \frac{nI(X_1:B_1S_1)-4n\delta}{\log_2 d} \rceil$ so that

$$P_{er} \leqslant 2^{-n\delta+1}. \tag{B.45}$$

For $n \to \infty$, $P_{er} \to 0$ exponential fast. We just showed how to achieve the secret key rate of $nI(X_1 : B_1S_1)$ bits. The fact that this rate is optimal follows from the Holevo's bound (Theorem A.24). If a higher rate was achievable, Alice and Bob would be able to distill with exponential reliability a private state associated to a secret key larger then $nI(X_1 : B_1S_1)$. From theorem 2, it would imply that Bob could distinguish between more then $nI(X_1 : B_1S_1)$ different states with exponential reliability. By the HSW-theorem, it would allow Alice to transmit more than $nI(X_1 : B_1S_1)$ bits of information to Bob. This is impossible as a direct consequence of the Holevo's bound.

In the actual privacy amplification protocol, Alice and Bob computes $f(k)$ and this give them the secret key. Alice won't perform any measurement in the X-bases. Computing $f(k)$ is the same as measuring the observables $Z^{M_{i1}} \otimes ... \otimes Z^{M_{ij}} \otimes ... \otimes Z^{M_{in}}$ for $1 \leqslant i \leqslant m$. Any observable that commute with those, as $X^{M^{\perp}_{1j}} \otimes ... \otimes X^{M^{\perp}_{ij}} \otimes ... \otimes X^{M^{\perp}_{nj}}$ for $1 \leqslant j \leqslant n$ for $1 \leqslant i \leqslant n - m$, will not perturb the state $|f(k)\rangle$ or change any information about it. Supposing Alice performs the measurements in the X-bases and she sent her results to Bob through an insecure channel, the extra information $g(\widetilde{x})$ that Eve could obtain could only help her. In practice, Alice does not have to perform the measurements in the X-bases and send $g(\widetilde{x})$ to Bob since those operations do not affect the value of $f(k)$ and could not decrease Eve's information about it. This argument is identical to the one given in [173] to explain why phase error correction does not have to be performed to obtain a secure key. A similar observation explains why we can allow Bob to use the shield S to guess Alice's state $|\widetilde{x}\rangle$: Eve does not have more information about $f(k)$ supposing that Bob does or does not have access to the shield.

Appendix C

Other Results Relevant to QKD

Until now, we have presented — with some extra material—most of the results given in [24, 27, 148]. In this appendix, we present more work completed while I was a Ph.D student. The first two papers are about the proposal and the physical implementation of a two-photon QKD protocol that do not required a reference frame. Most of the results given in Chapter 4 can be adapted to this protocol and an overview of the unconditional security proof was given in the second paper. In the third paper, we discuss and analyze in details for some QKD protocols the rather obvious fact that if Alice and Bob can be confident that all or a fraction of the noise in the detector is not caused by an eavesdropper, then they can achieve a much higher secret key rate and longer distances.

C.1 Robust Quantum Communication Using A Polarization-Entangled Photon Pair

J.-C. Boileau, R. Laflamme, M. Laforest, C. R. Myers
Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada.

Abstract: Noise and imperfection of realistic devices are major obstacles for implementing quantum cryptography. In particular, birefringence in optical fibres leads to decoherence of qubits encoded in photon polarization. We show how to overcome this problem by doing single qubit quantum communication without a shared spatial reference frame and precise timing. Quantum information will be encoded in pairs of photons using “tag” operations which corresponds to the time delay of one of the polarization modes. This method is robust against the phase instability of the interferometers despite the use of time-bins. Moreover synchronized clocks are not required in the ideal no photon loss case as they are only necessary to label the different encoded qubits.

Quantum mechanics allows the distribution of cryptographic keys whose security is based on the laws of physics instead of the difficulty of solving mathematical problems

[189, 17]. Turning this idea into practical technologies brings exciting challenges. The first prototype for quantum cryptography was built more than ten years ago over a distance of 30 cm in free space [16] and used the photons' polarization as qubits of information. Since, many quantum key distribution (QKD) experiments have been realized through air and optic fibres [69]. One of the obstacles to improve the fibre based prototypes is the birefringence effects due to geometric asymmetries and tension fluctuations which are a major impediment for polarization based-coding experiments [66]. When the coherence time of the photon is large compared to the delay caused by polarization mode dispersion, the birefringence can be represented by a time dependent unitary transformation $U(t)$ that acts on the polarization space. The time dependance comes from the mechanical variations in the fibre over time and its rate varies with the environmental conditions.

A possible solution to this problem is the application of active feedback [63]. Tomography on some predetermined polarization states could be used to approximate U for a certain time interval [90, 158]. By applying his approximation of $U^\dagger$ before his measurements, Bob (the receiver) could recover the states sent by Alice (the sender). However, this technique is practical only if the rate of change of U is relatively low. For this reason, the most successful QKD experiments were not based on polarization coding, such as the phase based experiment proposed by Bennett *et al.* using unbalanced interferometer [15, 185, 87]. However, a good control of the polarization modes is necessary to obtain a better visibility since some components like phase modulators are polarization dependent and the temperature of the interferometers must be stabilized since very small fluctuations between the two arms cause phase shifts that corrupt the quantum states.

Another very important example of a successful QKD protocol is the Plug-and-Play set-up [133, 178]. Using a Faraday mirror [129], the photons sent by Bob are reflected back in the fibre by Alice, who in turn encodes information in their phase. By travelling back in the fibre, the birefringence is reversed and, as it can be shown, the polarization state received by Bob are orthogonal to the original one. Since Bob controls the polarization state of the photon, he can make use of a polarized beamsplitter which increases the interference visibility. Although the Plug-and-Play set-up has very interesting characteristics, it is not compatible with a non-Poissonian source which could get rid of the multi-photons per pulse problem. Another disadvantage is that the use of two-way quantum cryptography is more vulnerable to a certain kind of eavesdropping strategy: the Trojan attack. An eavesdropper (i.e. Eve) could send photons in Alice's lab, catch them after they were reflected by the Faraday mirror and get some information about Alice's set-up without being detected.

To circumvent the threat of the Trojan attack and the instability of the interferometers, Walton *et al.* [187] proposed a one-way protocol based on decoherence free-subspaces in which each qubit is encoded in the time and phase of a pair of photons. In this Letter, we propose a new way to protect qubits encoded in polarization states of a photon pair from birefringence effects in optical fibre.

The idea is to take advantage of the fact that birefringence can be well approximated by a collective error model as long as the photons travel inside a time window small compared to the variation of the birefringence. Thus, if the effect of birefringence on one photon is $U(t)$, on n photons it is $U(t)^{\otimes n}$. This latter operator can be interpreted as a rotation of the reference frame axis and our protocol reduces to the problem of developing a strategy to do quantum communication without a shared reference frame.

In a recent paper [10], Bartlett *et al.* showed it should be possible to "*communicate with perfect fidelity without a shared reference frame at a rate that asymptotically approaches one encoded qubit per transmitted qubit.*" In particular, they proposed a method to encode a qubit using four photons in a decoherence-free-subspace of the collective noise model. However this required having full control of the states of qubits. This is out of reach of today's technology. More recently, two realistic QKD protocols that do not require any shared reference frame have been proposed [24]. These protocols do not require a general state of a qubit but only a set of non-orthogonal states. It encodes qubits in both three and four photon states, which makes the protocol more sensitive to photon loss. For these reasons, we will describe a two photon protocol robust against phase instability of the interferometer without the need for a shared spatial reference frame or synchronized clocks. If we neglect dispersion and discard relativistic situations then we are close to having no need for a shared reference frame at all.[1]

To explain our protocol we need to introduce the "tag" operation T_i which delay the photons in the state $|i\rangle$ by a specific amount of time. Experimentally it can be implemented using a polarized beamsplitter to separate polarization modes in arms of different length before recombination in the same optical path.

Suppose Alice inputs a two-photon state of the form $\alpha|HV\rangle + \beta|VH\rangle$ where H and V correspond to the horizontal and vertical polarization state of a photon. The time delay between the two photons Δt_p, must be fixed by Alice and known by Bob. It must be large enough such that Bob's apparatus can differentiate between the two photons and that "tag" operation will never change their order of arrival. If Alice applies the "tag" operation T_V on the initial state then she will have $\alpha|HV_T\rangle + \beta|V_TH\rangle$, where subscript T denotes the delay. Suppose some collective noise $U^{\otimes 2}$ (that includes a change of reference frame) is applied to this state when it travels to Bob and suppose also that Bob applies the "tag" operation $T_{H'}$ when he receives it. Up to a global phase, the state is then mapped

[1] For reasons we will explain later, Bob needs to know the relative rate of time flow in Alice's reference frame.

to

$$\begin{aligned}
&\frac{\alpha}{2}(|H'_T V'_T\rangle - |V' H'_{TT}\rangle + \delta_1(|H'_T V'_T\rangle + |V' H'_{TT}\rangle) \\
&+\delta_2(|H'_T H'_{TT}\rangle + |V' V'_T\rangle) + \delta_3(|H'_T H'_{TT}\rangle - |V' V'_T\rangle)) \\
&+\frac{\beta}{2}(|V'_T H'_T\rangle - |H'_{TT} 1'\rangle + \delta_1(|V'_T H'_T\rangle + |H'_{TT} 1'\rangle) \\
&+\delta_2(|H'_{TT} H'_T\rangle + |V'_T V'\rangle) + \delta_3(|H'_{TT} H'_T\rangle - |V'_T V'\rangle))
\end{aligned} \tag{C.1}$$

where $|H'\rangle$ and $|V'\rangle$ notation is used since the state is now defined in Bob's reference frame. We used the fact that the anti-symmetric state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ is invariant under collective noise and that $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle)$ will be mapped to a superposition of the triplet Bell states for which the δ's represent the relative weights and phases and follow the equality $||\delta_1||^2 + ||\delta_2||^2 + ||\delta_3||^2 = 1$. For later convenience, we define $|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|HH\rangle \pm |VV\rangle)$ and we will drop the apostrophe notation for simplicity.

The last operation is to project onto the states subspace in which the photons are separated in time by exactly Δt_p, i.e. both have been subjected to one tag operation. This operation does not require synchronized clocks, since Bob just needs to compare the arrival time of both photons. If the interval of time between a pair of photons is not Δt_p, then he discards these qubits, which happens $1 - ||\frac{(1+\delta_1)}{2}||^2$ of the time if we neglect photon loss. Otherwise, Bob will obtain Alice's initial state $\alpha|H_T V_T\rangle + \beta|V_T H_T\rangle$ with certainty. As it could have been showed using simple calculations, the final result is independent of the phase coherence instability between both arms of the interferometer in a way similar to the qubits encoded in the Walton *et al.* protocol [187].

To check if the communication is efficient, $||\frac{(1+\delta_1)}{2}||^2$ must be estimated. If the collective noise is averaged uniformly[2] over all possible values of $U(t)^{\otimes 2}$, then $\langle||\frac{(1+\delta_1)}{2}||^2\rangle = \frac{1}{3}$, which means Bob will obtain Alice's state with a probability of $\frac{1}{3}$. Yet, this result supposes that the unitary matrix U will average uniformly over all possible values during the communication time. To make the protocol independent of the environment, Bob could apply a random unitary matrix $B^{\otimes 2}$ on the photon polarization states just before making his "tag" operation[3].

An improved version of the scheme exploiting some partial knowledge of the shared reference frame to modify the transformation B to approximate the transformation $U^\dagger(t)$

[2] We assume that the randomness of the birefringence is such that the distribution of U over a large amount of time is uniform. The Haar measure over the space of unitary matrices is then used to calculate the average $\langle\langle\psi|T_1^\dagger U^{\otimes 2} T_1|\psi\rangle\rangle$ which equals $\frac{1}{3}$ independently of $|\psi\rangle$. Consequently, $\langle||\frac{(1+\delta_1)}{2}||^2\rangle = \frac{1}{3}$.

[3] The distribution of the operator B should correspond to the normalized Haar measure. Experimentally, B could be implemented with Pockels cells the same way as Franson and Jacobs in their 1995 experiment [63].

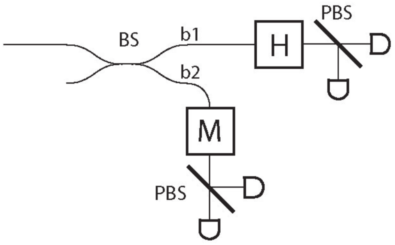


Figure C.1: After receiving the two photons and applying his "tag" operation, Bob can use this circuit to measure the qubit $\alpha|HV\rangle + \beta|VH\rangle$ in any basis by adjusting the gate M with a success probability of at least $\frac{1}{8}$. We refer to the text for more details.

would increased the ratio of useful encoded qubits. Depending on the efficiency of the active feedback mechanism and the rate of change of $U(t)$, the ratio could converge to 1.

To measure the qubit in a particular basis, Bob could use a normal symmetric beamsplitter and consider the result when each photon goes through a different branch, as shown in figure C.1. Define p such that $p = 0$ if the first photon goes through branch b1 and 1 if it is the second photon. Remark that the two photons arrive at the beamsplitter at different times and that Bob can differentiate them. At the end of branch b1, Bob measures in his diagonal $\{|+\rangle, |-\rangle\}$ polarization basis. Define k such that $k = 0$ if the outcome is $|+\rangle$ and 1 if it is $|-\rangle$. The photon on the other branch b2 must then be in the state $X^p Z^k(\alpha|H\rangle + \beta|V\rangle)$ where X and Z are the corresponding Pauli operators. Using Pockels cells (M) on the second branch and a polarized beam splitter, Bob can measure the qubit in any specific basis with a chance of success reduced by a factor of at most 8, since at the very least the measurement is successful when each photon exits from a different branch and p=k=0. Measurement in some bases will be successful more often than others.

We have described a technique to encode a robust qubit against collective noise and to measure it in any basis. We now show how this could be useful for a realistic QKD implementation. First, we describe the well known QKD protocol BB84 [17]. This protocol uses a set of four quantum states consisting of two maximally conjugate basis states $|0\rangle, |1\rangle$ and $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Alice randomly chooses which basis she will use to encode qubits to send Bob, who, upon arrival of a qubit, also chooses at random in which of the two basis he will perform a measurement. After repeating the protocol for a string of random bits, they publicly share what basis they used for each qubit. The bits for which they have used the same basis is used to build the *sifted* key. Since Eve has no prior knowledge of which basis Alice and Bob will use, any attempt of eavesdropping will disturb the states and induce errors in the *sifted* key with high probability. A portion of the *sifted* key is used to detect possible eavesdropping. If the error rate is lower than some given threshold, the left over bits will be transformed to the final secret key by using error correction and privacy amplification [130, 173].

To implement a protocol similar to BB84, Alice needs to encode the states $|HV_T\rangle$, $|V_TH\rangle$, $\frac{1}{\sqrt{2}}(|HV_T\rangle + |V_TH\rangle)$ and $\frac{1}{\sqrt{2}}(|HV_T\rangle - |V_TH\rangle)$ using parametric down conversions, filters and polarized beamsplitters as shown in figure C.2. We have to note that the measurement procedure described earlier works only if the state received by Bob after post-selection was of the form $\gamma_1|HV\rangle + \gamma_2|VH\rangle$ where $\gamma_i \in \mathbb{C}$ respecting a normalizing condition. This condition may no longer be true if sources of noise other than collective noise are considered or if we suppose that Eve altered the state sent to Bob. In the latter case, Bob's state after post-selection would look like $\gamma_1|HV\rangle + \gamma_2|VH\rangle + \gamma_3|VV\rangle + \gamma_4|HH\rangle$. To implement the provenly secure BB84 protocol, Bob must be able to project that state into the subspace in which Alice has encoded her space i.e. the space spanned by $|HV\rangle$ and $|VH\rangle$. If Bob wants to measure in the computational basis $(\{|VH\rangle, |HV\rangle\})$, then immediately after his "tag" operation he simply needs to measure the $|H\rangle$ or $|V\rangle$ polarization of each photon. In this case, he will also distinguish and be able to discard the states $|HH\rangle$ and $|VV\rangle$. The measurement in the diagonal basis $|\Psi^{\pm}\rangle$ is not as straight forward. Suppose Bob applies an extra Hadamard gate on both photons before measuring the polarization states. If $\gamma_3 = \gamma_4 = 0$, then he measures $|\Psi^+\rangle$ if both photons have the same polarization and $|\Psi^-\rangle$ if they have different polarization. In general, $\gamma_3 = \gamma_4 \neq 0$, but the uniformly distributed random rotation B performed by Bob (unknown to Eve) when he received the state will destroy any phase coherence between the states $\gamma_1|HV\rangle + \gamma_2|VH\rangle$, $|HH\rangle$ and $|VV\rangle$ from Eve's perspective. Intuitively, this means if Eve used the space spanned by $\{|VV\rangle, |HH\rangle\}$ it would be the same as if she randomly sent one of $|\Psi^-\rangle$ or $|\Psi^+\rangle$ to Bob, giving her no advantage. The complexity of the QKD security proof which includes coherentattacks restrains our argument, but the authors conjuncture that our protocol is unconditionally secure with the same error threshold as BB84. As a last remark, we note that only the qubits that have survived the post-selection are used to build the *sifted* key to estimate the error rate and construct the final secret key.

Earlier we discussed the possibility of using a feedback mechanism to increase the success rate of the post-selection. It could also be used in the QKD implementation discussion above, but Bob must be careful with whatever mechanism he uses since he must ensure the phase coherence between the three states $\gamma_1|HV\rangle + \gamma_2|VH\rangle$, $|HH\rangle$ and $|VV\rangle$ be lost from Eve's perspective. A final random phase gate would be enough since it does not affect the success probability of the post selection, but will destroy the coherence between these states.

The advantages of our protocol over the Plug -and-Play one are that this protocol is one-way, so there is no need to be as worried with the Trojan attack. Moreover, it does not

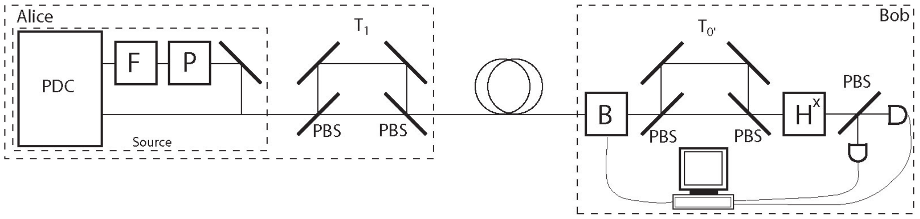


Figure C.2: Implementation of a modified version of BB84 protocol based on qubits robust against collective noise. Quantum states are generated through parametric down converters (PDC) supplemented by filters (F) and phase shifter (P). Alice and Bob do their "tag" operation using polarized beamsplitters (PBS). The B operator is randomized uniformly or determined by using a smart feedback mechanism. Bob measure the state in the computational or the diagonal basis depending if he applied the identity ($x = 0$) or the Hadamard gate ($x = 1$).

require interferometer stability like in the Walton *et al.* protocol (by using decoherence-free subspace). Although our protocol has similarities to the latter protocol, it is distinct for the following reasons:

First, synchronized clocks are necessary in our protocol only to label the different photon pairs. In the Walton *et al.* protocol, Bob must be able to distinguish between photons that have been delayed once, twice and not at all. Our protocol just needs to compare the delay between the two photons and not their particular time of arrival. Consequently, it requires a much smaller order of timing precision. For example, parametric down-conversion sources with long pulse length no longer induce errors caused by uncertainty in the emission time since both photons are always created simultaneously. Remark that if the number of events in which simultaneous dark counts on different detectors occur is negligible, extra timing precision would not help Alice and Bob to reduce the noise caused by the detector's dark counts and is therefore not necessary to our protocol.

Second, in the Walton *et al.* protocol, there is a $\frac{1}{4}$ chance, independent of the birefringence, that the photons will be measured in the phase basis and a $\frac{3}{4}$ chance of measuring in the time basis. However, the optimal efficiency for the ideal implementation of BB84 is a probability of measurement equal to $\frac{1}{2}$ in each basis. For this reason, Walton *et al.* indicate that the intrinsic efficiency of their scheme was $\frac{1}{4}$. In the case where B is chosen from a uniform distribution, our protocol would have an intrinsic efficiency ratio of $\frac{1}{6}$ since only a third of the photon pairs is not discarded. However, depending on the feedback mechanism, the intrinsic efficiency ratio could be higher than $\frac{1}{6}$, up to $\frac{1}{2}$.

Third, the final state Bob uses is encoded in polarization, not in time and phase. A good control of the polarization states allows Bob to get ride of the noise caused by the polarization dependance of some experimental components, like phase modulators.

In this paper, we have given a realistic robust scheme to do single qubit communication using two-photon states per encoded qubit. This technique goes around the problem of birefringence in optical fibre, the requirement of high precision synchronized timing and also the interferometer phase coherence instability. The protocol could be slightly modified to exploit partial information about a spatial reference frame to increase the bit rate by using active feedback. We also explained how to implement a slightly modified version of BB84 using the previously mentioned methods.

We would like to conclude with some problems that could make an experimental implementation of our schemes more difficult. Depolarization could be a serious distance limitation for our protocol, forcing us to use sources with longer coherence times [69]. To prevent chromatic dispersion from affecting the time delays between the photons, the average wavelength of the photons should be chosen according to the zero chromatic dispersion of the optical fiber [69, 176, 62]. Finally, since our protocol encoded each qubit with two photons, attenuation and detector's inefficiencies have a more significant affect on its efficiency compared to one-photon protocols. Nevertheless our proposal is in reach of experimental implementation and provides an elegant solution to the problem of birefringence in optical fibres.

We wish to thank Konrad Banaszek, Joseph Emerson, Simon Leblond and Rob Spekkens for helpful discussions that inspired this work. J.C.B., M.L., R.L. acknowledge support NSERC and R.L. from ARDA.

C.2 Experimental Quantum Communication without Shared Reference Frame

T.-Y. Chen[1], J. Zhang[1], J.-C. Boileau[2], X.-M. Jin[1], B. Yang[1], Q. Zhang[1], T. Yang[1], R. Laflamme[2], and J.-W. Pan[1,3]

[1] Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China.

[2] Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, N2L 3G1,Canada.

[3] Physikalisches Institut, Universität Heidelberg, Philosophenweg 12, 69120 Heidelberg, Germany.

Abstract: We present an experimental realization of a robust quantum communication scheme [Phys. Rev. Lett. **93**, 0220501 (2004)] using pairs of photon entangled in polarization and time. The scheme overcomes errors due to collective rotation of the polarization modes (for example, birefringence in optical fiber or misalignment), is insensitive to phase's fluctuation of the interferometer and does not require precise timing. No shared reference frame is required except from the need to label the different photons. We use this scheme to implement a robust variation of the

Bennett-Brassard 1984 quantum key distribution protocol (BB84) over 1km of optical fiber. We conclude by discussing and solving the unconditional security of our protocol.

Quantum cryptography [17], whose security is based on the fundamental principles of quantum mechanics, is a fast expanding field of quantum information both theoretically and experimentally [69]. Recently, many quantum key distribution (QKD) experiments have been realized through optical fiber and free space using weak-coherent source or entangled photon pairs. The maximum distances of free space QKD using weak-coherent source and entangled photons are 23.4km by Kurtsiefer *et al.* [111] and 13km by Peng *et al.* [140], respectively. Their aim was to try to validate the feasibility of quantum communication with satellites. Despite some security flaws, fiber-based QKD over 100km has been achieved [70].

Polarization and phase-time are most common coding methods to implement QKD. Although polarization can be suitable for free space QKD, it is generally not suitable for fiber-based QKD because of the time and wavelength dependences of birefringence which will depolarize the photons. Experimentally, active feedback or self-compensation could be applied to solve these problems [63], but it is efficient only when the thermal and mechanical fluctuations are rather slow. A popular alternative to polarization coding is phase-time coding using unbalanced interferometers [15, 87, 183]. However, phase-time coding can be very sensitive to the phase's fluctuations between the two arms of the interferometers and requires thermal stability. Some ingenious tricks like two-way communication [133, 178] are insensitive to phase's fluctuation, but have themselves disadvantages like being incompatible with perfect single photon sources and being sensitive to backscattering light.

To overcome the problems mentioned above, Walton *et al.* proposed a scheme based on decoherence-free subspace (DFS) which required encoding qubits using phase and time entanglement between two photons [187]. Then Boileau *et al.* [25] proposed a variation of that protocol that use a combination of time bins and polarization modes for coding. These schemes are insensitive to phase's fluctuations of the interferometer and robust against collective rotation induced by birefringence or misalignment. In single photon QKD protocol, a precisely synchronized clock is necessary to reduce the time window to minimize the contribution of dark counts. However, it is not the case for coding schemes using photon pairs, because the photons simultaneously originating from the pair can provide precise time references for each other. The fact that no synchronized clock is necessary could be useful if the arrival time of photon fluctuates.

The obvious disadvantages of two photon schemes are that they are much more sensitive to photon loss and seem more inefficient than the single photon schemes. However, it would be possible to reduce the qubit losses to a level comparable to single photon schemes by using post-selection, entanglement swapping and quantum memory devices [26]. As a step

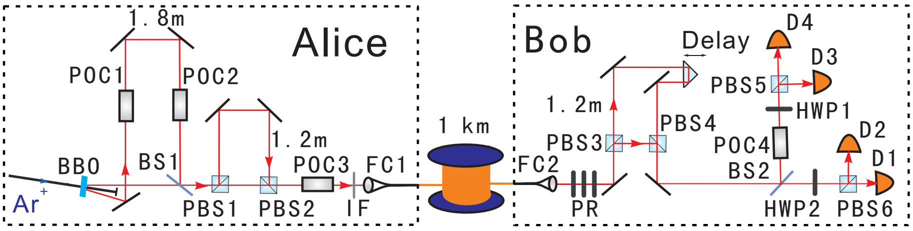


Figure C.3: Experiment setup for robust QKD. A 400 mW Argon-ion laser beam (Ar^+) of 351 nm passes through a 2 mm beta-barium-borate (BBO) crystal and produces polarization-entangled photon pairs of 702 nm. POC1-POC4 are four Pockel cells; BS1 and BS2 are beam-splitters; PBS1-PBS6 are six polarized beam-splitters; FC1 and FC2 are two fiber couplers; HWP1 and HWP2 are two half-wave plates (HWP); a HWP inserting between two quarter-wave plates (QWP) constitutes a polarization rotator (PR); IF is a interference filter; D1-D4 are four single photon detectors.

forward in that direction, we implemented a variation of the BB84 protocol based on the robust scheme of Boileau *et al.* [25], and realized an efficient quantum communication without any shared reference frame.

Our experimental scheme is illustrated in Fig. C.3. On Alice's side, polarization-entangled photon pairs are generated via type-II spontaneous parametric down-conversion (SPDC) [112]. The two photons of each pair are labelled by passing two arms with 1.8m length difference. In the long arm, two Pockel cells (POC1 and POC2), driven by high voltage pulse generators gated with random number signal, are used to produce the four states similar to that of BB84: $|HV\rangle + |VH\rangle$, $|HV\rangle - |VH\rangle$, $|HV\rangle + i|VH\rangle$ and $|HV\rangle - i|VH\rangle$, where H and V stand for horizontal and vertical polarization mode, respectively.

The two entangled photons can be combined into the same path in the beam-splitter BS1 with a probability of 1/4. Afterward, the vertically polarized photons are firstly tagged with a delay T in an unbalanced interferometer composed of two polarizing beam-splitters (PBS1 and PBS2) with a 1.2m difference between the two pathes' length. The POC3 is added to perform a random collective rotation of polarization before the photons are collected into the fiber coupler (FC1). Then, the photons are sent to Bob directly or through a 1 km single mode optical fiber. A polarization rotator (PR) is used to simulate the collective rotation noise. On Bob's side, the received horizontally polarized photons are tagged with the same delay T. To make the two timing tags exactly the same, a right-angled prism (Delay in Fig. C.3) is used to adjust the path length precisely.

Using notation introduced in Ref. [25] and supposing that the initial state was of the form $\alpha|HV\rangle + \beta|VH\rangle$, the resulting state can be written as:

$$((\delta_1+1)/2)(\alpha|H'_T V'_T\rangle + \beta|V'_T H'_T\rangle) + ((\delta_1-1)/2)(\alpha|V' H'_{TT}\rangle + \beta|H'_{TT} V'\rangle)$$
$$+((\delta_2+\delta_3)/2)(\alpha|H'_T H'_{TT}\rangle + \beta|H'_{TT} H'_T\rangle) + ((\delta_2-\delta_3)/2)(\alpha|V' V'_T\rangle + \beta|V'_T V'\rangle),$$

where $|H'\rangle$ and $|V'\rangle$ represent Bob's polarization basis frame which can be different from Alice's one. The subscripts T and TT mean that the photon has been tagged once and twice, respectively. The δ_j's are parameters that depend directly on the collective rotation of the polarization mode. They satisfy the following relation: $||\delta_1||^2 + ||\delta_2||^2 + ||\delta_3||^2 = 1$. Giving the arrival time of the photons, the final state is projected to the original state with a probability $p_s = ||\frac{\delta_1+1}{2}||^2$. The p_s could be anything between 0 and 1. To make p_s independent of the environment or any misalignment of the reference frame, Alice or Bob could apply a random unitary transformation $B^{\otimes 2}$ between the two tagging operations. If B is chosen from the uniform distribution over $U(2)$, then p_s is in average equal to $\frac{1}{3}$ whatever is the collective rotation. Because it's difficult to realize random transformation B over the whole $U(2)$ space experimentally, we simplify the experimental set-up by using only one POC (POC3 in Fig. C.3). Making it do nothing half of the time, and a bit-flip operation otherwise, p_s could also average to a non-zero value, $\frac{1}{4} \leqslant p_s \leqslant \frac{1}{2}$.

The received photons are split at BS2. The two half-wave plates HWP1 and HWP2 are set such that they performs as Hadamard gates on the polarization. By switching POC4 such that it do nothing or act as a QWP at 90°, we can select a random measurement basis (either $\{|H'_TV'_T\rangle + |V'_TH'_T\rangle, |H'_TV'_T\rangle - |V'_TH'_T\rangle\}$ or the $\{|H'_TV'_T\rangle + i|V'_TH'_T\rangle, |H'_TV'_T\rangle - i|V'_TH'_T\rangle\}$). By post-selecting the cases where each of two photons exit from different outputs of BS2 and their arrival time difference is 6ns (which is related to the 1.8m time label), the states are differentiated according to their polarization (the same or different) [25]. The detection events within the 3ns coincident time window are recorded to generate quantum key bits.

For the measurement to succeed, it is crucial to observe two photons interference after the timing tags. It requires to match accurately the difference of the path's lengths of the two interferometers by adjusting the prism on Bob's side (see Fig. C.3). The curve in Fig. C.4 shows an interference fringe with a visibility of above 95%. The fact that interference is observed over a large length interval (of about one hundred micrometers) clearly implies that the interference is robust against the phase instability of the interferometers as claimed in Ref. [187, 25].

In order to demonstrate the robustness of the protocol in principle, we first use a 4m optical fiber to implement the QKD protocol. Approximately $12,000$Hz polarization-entangled pairs are detected behind a interference filter (IF) of 1.6nm FWHM. The entangled photon pairs are transferred to one of the four states randomly and sent to Bob. Due to the photon losses in the BSs and the fiber connecters, only a maximum of 140Hz coincidences can be registered on Bob's side after calibrating the PR. We then rotate the angle of the first QWP of the PR to simulate the degree of collective rotation noise. In the experiment, five settings are selected for particular angles of the QWP. The first setting

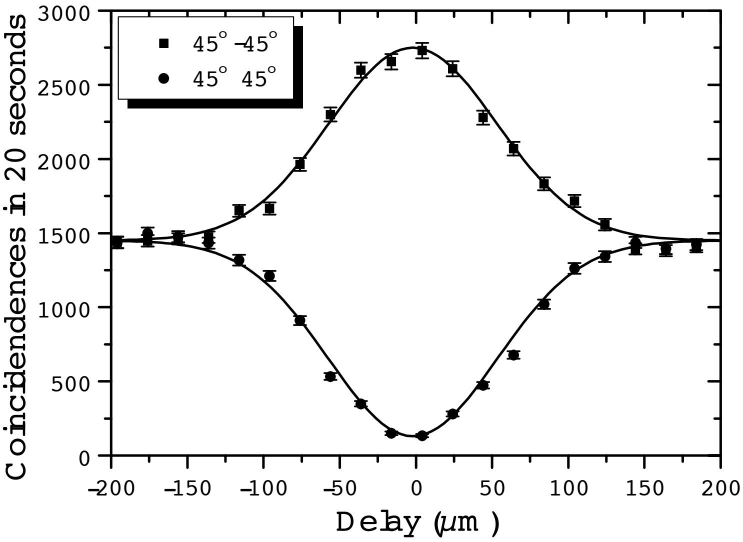


Figure C.4: Interference pattern observed for the state $|H'_TV'_T\rangle - |V'_TH'_T\rangle$ after applying a Hadamard gate on each photon and by measuring the coincident counts. The zero delay position corresponds to the maximum interference visibility.

corresponds to the case where there is no collective rotation and coincidence is maximal. The last setting corresponds to a collective bit-flip. The other settings are chosen via rotating the angle of QWP with equal intervals between the best and the worst settings. We investigated the change of error rates and coincidences under these conditions with or without random rotation implemented by POC3.

As shown in Fig. C.5, the coincidence without random rotation is very dependent of the collective noise. When the angle of rotation increases, the coincidence decreases to a minimum while the error rate increases close to 50%. As predicted, using random rotation makes the coincidence and the error rates much less dependent on the collective noise. We later show that the all the error rates obtained with the random rotations are suitable for secured QKD.

We also constructed a practical QKD system over 1km single mode fiber, whose attenuation for each photon at 702 nm is 4.8 db/km. Due to the photon loss in the fiber and an additional fiber connecter, the maximal coincidence detected in Bob's side dramatically drops to 1.4 Hz. We measured the error rates under the same collective noise settings as used in the experiment with short fiber. The results of the experiment are shown in Fig. C.6. Due to the photon loss, the QBER observed in the experiment is a bit higher than the cases with short fiber, but is still well below the lower bound for secure key

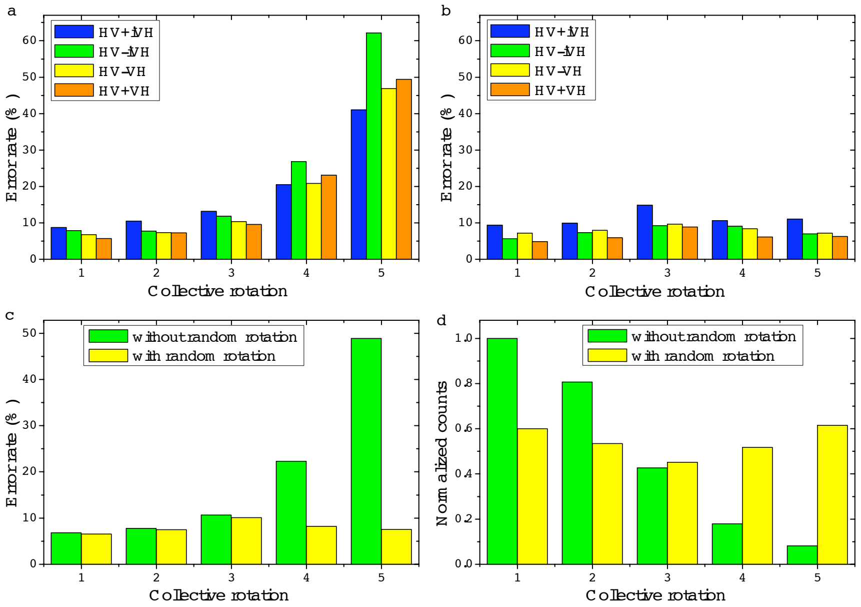


Figure C.5: Quantum bit error rates (QBER) measured under different collective rotations and with a 4m single mode fiber. The first collective rotation correspond to the identity and the fifth, to a bit-flip. The angles in between are chosen has described in the text. The QBER of each states was measured in 20 minutes without (a) or with random rotations (b). The average QBER over all states with and without the random rotations are compared in (c). In (d), we give the normalized coincidence counts in function of the angle of the collective rotation with or without random rotations.

distribution.

In fact, the error rates are mostly come from the imperfection of state preparation and accidental coincidence. The 2000 Hz single count rate of each detector and the 3ns coincidence window lead to an accidental coincidence of 0.024Hz. When the collective noise changes, the accidental coincidence will induce an error rate from 1.7% to 50% when no random rotation is applied. However, with the help of random rotation it will only cause an error rate varied from 3.4% to 6.8%, which is much more independent of the collective noise. The imperfect state source will also cause an error rate of 4% . The average error rate is observed to be 10.2% ± 0.3% (see Fig. C.6-c). The following security proof will show that our experimental method can successfully distribute quantum key over a practical collective noise channel without shared reference frame.

It should be noted that our experimental measurements are not exactly the same as the ones required by the standard BB84 protocol. However, if Bob was able to do a projection into the space S spans by $|H'V'_T\rangle$ and $|V'_T H'\rangle$, then both measurements would

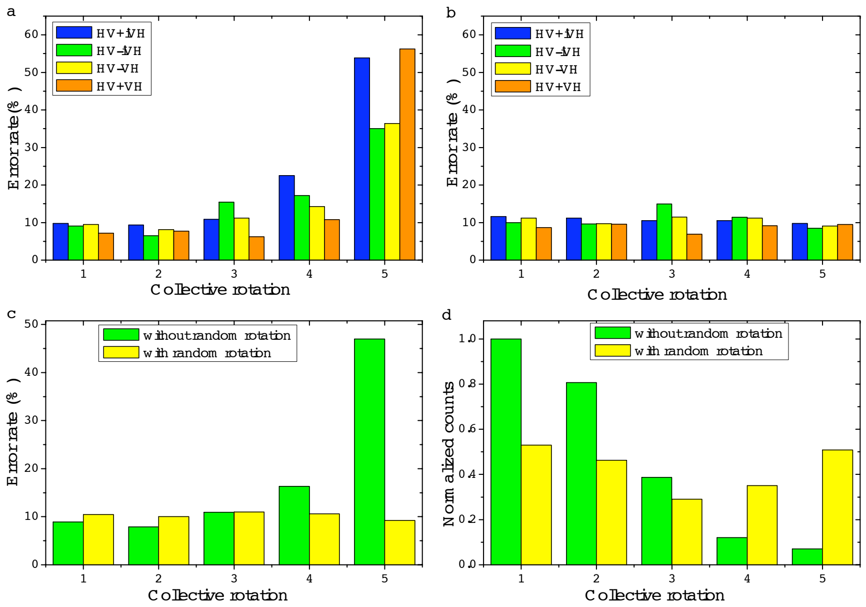


Figure C.6: QBER measured under different collective rotations (as in Fig. C.5) and with a 1km single mode fiber. The QBER of each states was measured in 3 hours without (a) or with random rotations (b). The average QBER over all states with and without the random rotations are compared in (c). In (d), we give the normalized coincidence counts in function of the angle of the collective rotation with or without random rotations.

be identical. Instead of doing this projection directly, suppose that, just before his tag operation, Bob apply on the polarization modes of the photon pairs a random phase shift $M_\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}^{\otimes 2}$ with ϕ chosen between 0, $\frac{\pi}{2}$, π or $\frac{3\pi}{2}$.

Because of the post-selection, the only states that we need to consider are the ones of the form $a_1|H'V'_T\rangle + a_2|V'_TH'\rangle + a_3|H'H'\rangle + a_4|V'_TV'_T\rangle$ for some complex numbers a_i. Consider the density matrix ρ of an arbitrary mixture of these states. Simple calculations show that the elements of $\rho' = \sum_{\phi=0,\frac{\pi}{2},\pi,\frac{3\pi}{2}} M_\phi \rho M_\phi$ that corresponds to $|V'_TV'_T\rangle\langle H'V'_T|$, $|H'H'\rangle\langle H'V'_T|$,$|V'_TV'_T\rangle\langle V'_TH'|$, $|H'H'\rangle\langle V'_TH'|$ and their transposes vanish. In other word, sometime the state is projected into or outside S. In the first case, the measurement reduce to one of the two von Neumann measurement used in the standard BB84. In the other case, the state may be projected outside S and the measurement fails. If Alice and Bob were able to know exactly which pairs were projected inside or outside S then they could reject those pairs that correspond to the wrong projection and could perform the standard BB84 protocol with the remaining pairs. However, they don't have this information. Instead, they can measure with a good accuracy the ratio of the states that are projected inside S, which we refer to as p^S. This can be achieved by selecting a random sample of photon's pairs on which the measurement corresponding to $\{|H'_T H'_T\rangle, |H'_T V'_T\rangle, |V'_T H'_T\rangle, |V'_T V'_T\rangle\}$ is applied immediately after the tagging operation, and by counting the results corresponding to either $|H'_T V'_T\rangle$ or $|V'_T H'_T\rangle$. If that sample is large enough, the measured ratio of states projected inside and outside S will be very close to p^S. Remark that the number of pairs required for the measurement of p^S is asymptotically negligible.

To obtain a secure key, it is necessary and sufficient to bound Eve information about the key after bit error correction since privacy amplification [19, 154] can be used to reduce asymptotically that information to zero with a key's lost proportional to Eve's information. For the qubits projected outside S, Alice and Bob assume the worst case scenario and suppose that Eve has full information about the results corresponding to these states. For the qubits projected inside S, Shor and Preskill's proof [173] bound Eve's information after bit error correction by $H(e_x^{\mathrm{s}})$ where H is the Shannon entropy, and e_x^{s} is the bit error rate restricted to the states projected inside S. Consequently, the secret key generation rate is at least

$$p_{\mathrm{s}} - H(e_x) - p_{\mathrm{s}} H(e_x^{\mathrm{s}}) \tag{C.2}$$

of the conclusive results, where $H(e_x)$ is the fraction of results lost asymptotically because of the bit error correction. e_x and e_x^{s} are the bit error rate over all conclusive results and over the conclusive results that were projected inside S, respectively. A conclusive result is defined as any measurement that give a bit to the key before error correction and privacy amplification.

Consequently, to be able to obtain a secret key, Alice and Bob simply need an upper bound for e_x^{s}, e_x being measured directly by using a sample of test bits and p_{s}, by the procedure outline above. e_x^{s} can be estimate from the equation

$$e_x = p_{\mathrm{s}} e_x^{\mathrm{s}} + (1 - p_{\mathrm{s}}) e_x^{\bar{\mathrm{s}}}, \tag{C.3}$$

where $e_x^{\bar{\mathrm{s}}}$ is the error rate of the conclusive results corresponding to states projected outside S. However, $e_x^{\bar{\mathrm{s}}} = \frac{1}{2}$ asymptotically. This is a consequence of Bob's random choice of ϕ for M_ϕ and the fact that the coefficients of the density matrix ρ' corresponding to $|V'_T V'_T\rangle\langle H'H'|$ and $|H'H'\rangle\langle V'_T V'_T|$ are zeros.

Therefore, in order to obtain a secret key Alice and Bob simply need to get a lower bound for p^S, and an upper bound for e_x. In the experiment, we measured p^S using the method explained above. In the case with 4m fibre, p^S is measured to be 97% and in the case with 1km fibre p^S is 91%. In our experiment with random rotations, the observed e_x's (see Figs. C.5 and C.6) are sufficient to guarantee secure key distribution.

Any coherent attack from an eavesdropper was considered in our security analysis. However, we assumed perfect state preparation and measurements, and that Eve's has no access whatsoever to Alice and Bob's lab. For a more realistic security analysis, considerations as the ones treated in Ref. [73, 67] would be necessary.

In summary, we have realized one of the first efficient quantum communication protocols without shared spatial and reference frame including no time reference, except to label the qubits. It could be useful for free-space transmission in the case where the receiver and the sender are moving relative to each other. It could also be useful to avoid birefringence effect in optical fiber and would be a possible solution to the phase instability of interferometers. Our experiment is a first step toward more efficient robust quantum communication since it is only an example of a series of more complex quantum communication schemes exploiting decoherence-free subsystem of the collective-noise and time tags [9]. We also showed the unconditional security of a robust quantum key distribution protocol based of BB84. We conclude with the remarks that technological advances of entangled photon sources and quantum memories would greatly enhance our results, there exist other interesting alternatives to quantum communication without alignment [174] and some other experiments exploiting the decoherence-free subspace of the collective-noise [29].

This work was supported by the CAS, the NNSFC, the NSERC and the PCSIRT. This work was also supported by the Alexander von Humboldt Foundation, the Marie Curie Excellence Grant of the EU and the government of Ontario.

C.3 Higher Security Thresholds for Quantum Key Distribution by Improved Analysis of Dark Counts

J.-C. Boileau, R. Laflamme, M. Laforest, C. R. Myers
[1]Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada.
[2]Perimeter Institute for Theoretical Physics, 35 King Street North, Waterloo, ON, N2J 2W9, Canada.

Abstract: We discuss the potential of quantum key distribution (QKD) for long distance communication by proposing a new analysis of the errors caused by dark counts. We give sufficient conditions for a considerable improvement of the key generation rates and the security thresholds of well-known QKD protocols such as Bennett-Brassard 1984, Phoenix-Barnett-Chefles 2000, and the six-state protocol. This analysis is applicable to other QKD protocols like Bennett 1992. We examine two scenarios: a sender using a perfect single-photon source and a sender using a Poissonian source.

The goal of quantum key distribution (QKD) is to extend a shared secret key for use as a one-time pad to encode classical messages. The advantage of QKD is that its security is based on the laws of quantum mechanics and not on the unproven complexity of a mathematical problem as in classical cryptography. These last few years, many encouraging experiments demonstrated QKD, some spanning more than a hundred kilometers through optical fibers [70, 96]. The main source of errors is usually due to dark counts from the detectors. A dark count is when a detector fires independently (or in the absence) of a qubit state encoded by the sender, Alice. If qubit losses are considerable, then the receiver, Bob, will receive many empty pulses, and dark counts from his detectors will induce a high error rate.

In this paper, for simplicity, we refer specifically only to four different QKD protocols: Bennett 1992 (B92), Phoenix-Barnett-Chefles 2000 (PBC00), Bennett-Brassard 1984 (BB84), and the six-state protocol, which are two, three, four, and six state protocols, respectively [15, 145, 17, 33]. In B92, Alice encodes random bits using two non-orthogonal states, say $|\psi_1\rangle$ and $|\psi_2\rangle$, and sends them to Bob. He makes the measurement corresponding to the Positive Operator-Valued Measure (POVM) $\{\alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1|, \alpha|\overline{\psi}_2\rangle\langle\overline{\psi}_2|, \mathbb{1} - \alpha|\overline{\psi}_1\rangle\langle\overline{\psi}_1| - \alpha|\overline{\psi}_2\rangle\langle\overline{\psi}_2|\}$, where $|\overline{\psi}_j\rangle$ is orthogonal to $|\psi_j\rangle$ and α equals $\frac{1}{1+|\langle\psi_1|\psi_2\rangle|}$. Bob's measurement either determines which state Alice did not send (from which Bob can deduce the encoded bit) or is inconclusive. PBC00 is similar to B92 but uses three non-orthogonal states, say $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\psi_3\rangle$, that form an equilateral triangle in the X-Z plane of the Bloch sphere. She encodes her random bits using random bases from either $\{|\psi_1\rangle, |\psi_2\rangle\}$, $\{|\psi_2\rangle, |\psi_3\rangle\}$, or $\{|\psi_3\rangle, |\psi_1\rangle\}$. Bob performs the POVM $\{\frac{2}{3}|\overline{\psi}_1\rangle\langle\overline{\psi}_1|, \frac{2}{3}|\overline{\psi}_2\rangle\langle\overline{\psi}_2|, \frac{2}{3}|\overline{\psi}_3\rangle\langle\overline{\psi}_3|\}$. After Bob measures all of the qubits, Alice declares publicly which basis she used for each. By deduction, Bob can sometimes retrieve Alice's state. Alice and Bob discard the other results. It can be shown that, neglecting the qubit losses, the rate of conclusive results is $\frac{1}{2-e_x}$ where e_x is the bit error rate. A *conclusive* result corresponds to any pair of qubits not discarded by Alice and Bob.

To implement BB84, Alice encodes a random bit in either $\{|0\rangle, |1\rangle\}$ or its conjugate basis $\{|+\rangle, |-\rangle\}$. For each qubit, Bob randomly measures in one of these bases. They only keep results for which they used the same basis. The six-state protocol is identical to BB84 except that Alice and Bob choose from three different bases: $\{|0\rangle, |1\rangle\}$,$\{|+\rangle, |-\rangle\}$, and $\{\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\}$. We can modify BB84 and the six-state protocol by choosing bases with non-equal probabilities, increasing the chance of agreement [121]. The rate of results for which identical bases are used converges asymptotically to 1. Below, we calculate the key generation rates of BB84 and the six-state protocol using this asymptotic result.

Mayers [130] produced the first unconditional security proof of BB84. Shor and Preskill proposed a simpler proof based on ideas from Lo and Chau [173, 120]. Their security proof has been generalized to other protocols including B92, PBC00, and the six-state protocol [180, 182, 27, 118, 150]. We improve the secret key generation rate of these QKD protocols by proposing a slight modification of these proofs. Our main idea is based on a variation of a theorem proved in Ref. [73]. We assume that an eavesdropper, Eve, can perform any attack consistent with quantum mechanics, but cannot get any information about Alice's or Bob's labs or control their apparatus. We discuss later how realistic these assumptions are and how it is possible to slightly relax them. We study two cases: one where Alice's source can create a single photon on demand, and another where it follows a Poisson distribution. For simplicity, we give details only about Shor and Preskill's security proof of BB84 and not other protocols.

At the end of this paper, we compare the updated error rate thresholds and key generation rates of BB84, PBC00, and the six-state protocol with previous results. The same arguments could improve other QKD protocols, including B92. However, B92's phase estimation bound depends on qubit losses in the channel and the number of inconclusive results, complicating the analysis. Since our goal is to describe a general technique to improve security thresholds, we only treat the simpler cases as examples.

The Shor and Preskill proof first shows the security of an entanglement distillation protocol (EDP) for QKD, and subsequently reduces the EDP to BB84. For convenience, we define $|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle \pm |1\rangle|1\rangle)$ and $|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle \pm |1\rangle|0\rangle)$.

The structure of the **EDP that can be reduced to BB84** in Shor and Preskill's proof is as follows:

1. Alice creates n pairs of the form $|\Phi^{+}\rangle$ and sends the second half of each pair to Bob after randomly applying the identity or the Hadamard gate on it.

2. After Bob confirms that he has received all of Alice's states, Alice publicly declares the random rotation that she used on each qubit. Bob undoes the transformations on the corresponding qubits.

3. With no eavesdropping or channel noise, Alice and Bob will share n perfect pairs of the form $|\Phi^{+}\rangle$. They can now measure their qubits in the same basis to share a secret key. However, noise and eavesdropping induce errors. If the bit and the phase error rates are low enough, then error correction can be applied to obtain m perfect pairs of the form $|\Phi^{+}\rangle$ where $m \leqslant n$.

4. Alice and Bob can estimate the bit error rate by comparing bit measurements from a sample of pairs, called test bits. A bit (or X) error on a pair occurs when Alice and Bob share either $|\Psi^{+}\rangle$ or $|\Psi^{-}\rangle$. A phase (or Z) error corresponds to $|\Phi^{-}\rangle$ or $|\Psi^{-}\rangle$. A Y error corresponds to $|\Phi^{-}\rangle$ or $|\Psi^{+}\rangle$. Y error estimation could provide information about the correlation between bit and phase errors. Because Alice randomly applies the identity or Hadamard gate, it can be shown that the bit error rate, e_x, and the phase error rate, e_z, are approximately equal, independent of channel noise and Eve's strategy. In BB84,

Alice and Bob have no information about Y errors.
5. Depending on the bit error rate measured on the test bits, Alice and Bob apply error correction on the other pairs. If we suppose one-way error correction using CSS codes [34, 175], a lower bound for generation rate $\frac{m}{n}$ for the perfect pairs is given asymptotically by

$$S = p_c[1 - H(e_x, e_z)] \tag{C.4}$$

where H is the Shannon entropy ($H(e_x, e_z) = H(e_x) + H(e_z|e_x)$ is the entropy of the bit-phase error pattern) and p_c is the rate of conclusive results. For simplicity, we assume that the proportion of test bits is negligible. □

Shor and Preskill showed that this EDP, and thus BB84, were unconditionally secure with a key generation rate given by Eq. C.4. Since $H(e_x)$ is asymptotically the fraction of bits sacrificed for bit error correction, it implies that $H(e_z|e_x)$ is an upper bound on the fraction of information that Eve has about the key after bit error correction. A consequence is that privacy amplification, as introduced in Ref. [19], can be used to simplify the post-processing of the QKD protocol. As shown in Ref. [154], privacy amplification can generate a secret key by sacrificing a number of bits asymptotically proportional to Eve's information.

The reduction of the EDP to BB84 assumes that Alice uses a source which emits a single photon on demand. In a more realistic situation, Alice's source would emit a photon pulse following a Poisson distribution. Unfortunately, when Alice sends two or more photons containing the same quantum information at the same time, Eve can measure one to gain information about the key without detection. Accounting for this attack (but assuming Eve has no information about the random phase of the signal emitted by a coherent light source), a more general equation of the secret key generation rate, combining results from Ref. [73] and Ref. [118], and using the improvement suggested in Ref. [119], is given asymptotically by

$$S = p_c[\omega_0 + \omega_1 - H(e_x) - \omega_1 H(e_z^1|e_x)] \tag{C.5}$$

where ω_1 is the fraction of the conclusive results corresponding to single-photon pulses, ω_0 is the fraction of the conclusive results corresponding to empty pulses (the presence of background noise, for example), and e_x^1 (e_z^1) is the bit (phase) error rate restricted to conclusive results from single-photon pulses. e_x (e_z) is still the bit (phase) error rate over all conclusive results. If Alice has a source that emits a single photon on demand, then $\omega_0 = 0$, $\omega_1 = 1$, $e_j^1 = e_j$ for $j \in \{x, y, z\}$, and $S = p_c[1 - H(e_x, e_z)]$ as expected.

To prove Eq. C.5, it was argued that since Alice and Bob want an identical key and cannot differentiate multi-photon from single-photon pulses, they must correct all bit errors, asymptotically losing a fraction $H(e_x)$ of the results in the process. To apply privacy amplification on the remaining bits and obtain a secret key, Alice and Bob must upper bound Eve's information. If we assume that the phase of the signal is random[4], there is no coherence between states with different photon numbers. Thus, we can categorize each bit of the resulting key as being associated with an empty, single-, or multi-photon pulse. Assuming the worst case, Eve has full information about the results associated with multi-photon pulses. On the other hand, she has no information about Alice's bits corresponding to empty pulses. By the Shor-Preskill's arguments discussed earlier, the fraction of information that Eve could extract from the results corresponding to single-photon pulses is upper bounded by $H(e_z^1|e_x)$. Consequently, Eve's information about Alice's remaining key is upper bounded by $(1-\omega_0-\omega_1)+\omega_1 H(e_z^1|e_x)$. After privacy amplification, Eve has no information about Alice's key. The same is true of Bob's key since it is identical to Alice's. Therefore, the secret key generation rate is given by Eq. C.5.

Similarly, since Shor-Preskill's proof can be adapted to B92, PBC00 and the six-state protocol [180, 182, 27, 118], these protocols can be shown unconditionally secure with a key generation rate given by Eq. C.5.

The above argument does not differentiate between a single photon emitted by Alice that is successfully measured by Bob and a single photon that is lost in the channel (or taken by Eve) followed by a dark count measured by Bob. However, these cases may be analyzed separately. Consider the following four types of conclusive results.

1. Successful measurement of a *qubit state* (physically corresponding to a photon received from the channel) that originated from a single-photon pulse. Note that the qubit state could have been manipulated by Eve.

2. Successful measurement of a qubit state that originated from a multi-photon pulse.

3. Empty pulses from Alice followed by a successful measurement of a qubit state by Bob (ie. Eve may send a qubit state to Bob even if Alice emits nothing).

4. Dark count events: Bob doesn't receive a qubit state, but one of his detectors fires.

The dark count events are independent of Alice's or Eve's actions. We define p_c^{emp}, p_c^{sq}, and p_c^{mq} as the rate of conclusive results corresponding to qubit states, received by Bob, associated with empty pulses, single-photon pulses, and multi-photon pulses, respectively. We define p_c^{dk} as the rate of conclusive results associated with dark counts. Note that

$$p_c = p_c^{emp} + p_c^{sq} + p_c^{mq} + p_c^{dk}. \tag{C.6}$$

[4] Recently, it was shown the Eve could use extra information about the phase of the signal to her advantage [124], though the extent is unknown.

We remark that the background noise has two different contributions: intrinsic and extrinsic. The intrinsic contribution is caused by elements from Bob's lab while the extrinsic contribution is from external sources. The sun and backscattering light in two-way QKD are examples of external sources of background noise. Based on our assumptions, Eve may control the external sources of background noise, but not the ones inside Bob's lab. Following our previous definitions, the only contribution to p_c^{dk} is intrinsic. Any external sources will contribute to p_c^{emp}, p_c^{sq}, and p_c^{mq} since they correspond to Bob receiving a qubit state from the channel. For convenience, in this paper, *dark counts* always refer to the intrinsic contribution of background noise. We assume for simplicity that dark counts are independent of other measurement results.

We now explain how it is possible to achieve a better bound for the secret key generation rate than Eq. C.5. As before, a fraction $H(e_x)$ of the results are lost due to bit error correction. Assuming again that the phase of the signal is random from Eve's perspective, each bit of the resulting key corresponds to one of the four types of conclusive results described above. From previous arguments, Eve has a fraction $H(e_z^{sq}|e_x)$ of information about conclusive results from Category 1 and, in the worst case scenario, full information about those from Category 2. e_x^{sq} and e_z^{sq} are defined as the bit and phase error rates on the conclusive results restricted to Category 1. When Alice emits an empty pulse and it is followed by a successful measurement of a qubit state by Bob, we assume that the qubit state was created by Eve. A conservative assumption is that Eve has full information about Bob's results from Category 3.[5] Supposing dark count rates are the same in all detectors and independent of Eve and other measurement results, Bob's results from Category 4 are completely random and Eve has no information about them[6]. Consequently, the fraction

[5] In the case of B92, it is easy to show that this assumption is necessary, but it might be too strict for other protocols like PBC00, BB84, and the six-state protocol.

[6] For simplicity, we suppose that the dark count rates are uniform over all detectors and that they are independent of other measurement results. If dark count rates differ from detectors, we suggest two options. In one, Bob uses a random transformation to switch the role of the detectors in the measurement. For example, in BB84, Bob could apply, at random, an extra Y operation on the received qubits to switch the role of the detectors when measuring in the $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ bases. A second option is to bound Eve's information from an estimate of the probability that a detector fires relative to the others in the case of a dark count. Assuming dark counts are independent of other measurement results, in BB84 and the six-state protocol, with only two detectors, Eve's information is bounded by $1 - H(q)$ where q is the probability that the first detector fires in the case of a dark count. It is interesting to note that if Eve has some control over the probability q and could change it from one dark count event to another, then, by entropic concavity, Eve's information is bounded by $1 - H(q_{worst}^{ave})$, where q_{worst}^{ave} is the worst estimate of the average of q. Determining the value of q_{worst}^{ave} can be very hard, but it is related to the level of confidence that Alice and Bob have on their ability to counter or detect Eve if she tries to change the properties of the detectors. Similarly, if dark counts are correlated to other measurement results, we can upper bound Eve's information with restrictions on the correlations.

of information that Eve has on Bob's key after bit error correction is upper bounded by $\frac{1}{p_c}(p_c^{emp} + p_c^{mq} + p_c^{sq} H(e_z^{sq}|e_x))$. Therefore, the secret key generation rate is lower bounded by

$$S_b = p_c^{sq} + p_c^{dk} - p_c H(e_x) - p_c^{sq} H(e_z^{sq}|e_x). \tag{C.7}$$

We emphasize that it is not necessary for Alice and Bob to know which events correspond to each class of conclusive results.

In the derivation of Eq. C.7, we bounded Eve's information about Bob's key. However, we could have instead bounded Eve's information about Alice's key. In this case, Eve has no information about the bit chosen by Alice when she sends a vacuum states. But she could have some information about Alice's portion of the key corresponding to dark counts (unless Alice sent an empty pulse). Using similar arguments, we obtain

$$S_a = p_c^{sq} + p_c \omega_0 - p_c H(e_x) - p_c^{sq} H(e_z^{sq}|e_x). \tag{C.8}$$

Combining Eq. C.7 and Eq. C.8, we obtain a new lower bound for the secret key generation rate,

$$S \;=\; \max[S_a, S_b]. \tag{C.9}$$

Remark that the concavity of entropy and $\omega_1 e_z^1 = \frac{p^{sq}}{p^c} e_z^{sq} + (\omega_1 - \frac{p^{sq}}{p^c}) e_z^{dk}$ imply that $\omega_1 H(e_z^1|e_x)) \geqslant \frac{p^{sq}}{p^c} H(e_z^{sq}|e_x) + (\omega_1 - \frac{p^{sq}}{p^c}) H(e_z^{dk}|e_x)$. We can rewrite this as $\omega_1(1 - H(e_z^1|e_x)) \leqslant \frac{p^{sq}}{p^c}(1 - H(e_z^{sq}|e_x))$, since it can be argued that $e_z^{dk} = \frac{1}{2}$. Therefore, the secret key generation rate given by Eq. C.8 (and Eq. C.9) is always greater than or equal to the one given by Eq. C.5.

To evaluate Eq. C.9, Alice and Bob must be able to determine all quantities involved in it. For this purpose, we study two different situations: Alice has a source that emits a single photon on demand or one that follows a Poisson distribution.

In both situations, e_x is estimated from test bits, and p_c^{dk} can be calculated from the predetermined dark count probability C of the detectors and the number of empty pulses not associated with dark counts that Bob receives. If C is not fixed, Bob might block his detection units randomly and estimate p_c^{dk} from these results. For this to be true, it is important that Eve is not allowed to reduce the dark count probability without being detected. But is this a valid assumption? In practice, Eve could try to cool down the detectors or send bright pulses to disable them at will. Furthermore, there might be some uncertainty in the measurement of p_c^{dk}, even in the absence of an eavesdropper. Since a dark count could be interpreted as Eve sending a random state to Bob, we remark that lower bounds for C and p_c^{dk} are sufficient to obtain a better key generation rate using Eq. C.9. Establishing a high level of confidence on a lower bound for p_c^{dk} seems very hard in practice. However, it might be possible through experimental research and tests on reducing dark count rates of detectors.

If Alice has a source that emits single photons, $\omega_0 = 0$ and $p_c^{mq} = 0$, then Eq. C.9 reduces to Eq. C.7 and $e_x = \frac{1}{p_c}(p_c^{sq} e_x^{sq} + p_c^{dk} e_x^{dk})$, where e_x^{dk} is the bit error rate over conclusive events associated with dark counts. $e_x^{dk} = \frac{1}{2}$ which implies that Bob can estimate e_x^{sq} from the value of e_x measured on test bits. $H(e_z^{sq}|e_x) = H(e_z^{sq}|e_x^{sq})$ can be evaluated depending on the protocol used. It can easily be shown that, for the six-state protocol, $e_x^{sq} = e_y^{sq} = e_z^{sq}$ [118]. For BB84, $e_x^{sq} = e_z^{sq}$ and $0 \leqslant e_y^{sq} \leqslant 2e_x^{sq}$ [173]. For PBC00, it was shown that $e_z^{sq} = \frac{5}{4} e_x^{sq}$ and $\frac{1}{4} e_x^{sq} \leqslant e_y^{sq} \leqslant \frac{9}{4} e_x^{sq}$ [27].

In the absence of errors due to dark counts, $p_c^{dk} = 0$. By solving $S(e_x) = 0$, we find that the bit error rate threshold is 12.6% for the six-state protocol, 11.0% for BB84, and 9.81% for PBC00. If we now suppose that e_x^{sq} is fixed, then the bit error rate threshold increases as shown in Tab. C.1. Note that the bit error rate threshold depends on the contribution of errors not associated to dark counts.

Table C.1: Bit error rate thresholds for BB84, PBC00, and the six-state protocol using a single-photon source and assuming fixed values of e_x^{sq}, which is the bit error rate of the results not associated with dark counts.

	$e_x^{sq} = 0$	$e_x^{sq} = 0.01$	$e_x^{sq} = 0.1$
PBC00	50%	43%	-
BB84	50%	44%	13%
Six-State Protocol	50%	46%	19%

Tab. C.1 reflects the potential of a special analysis for dark counts. For any of the previous QKD protocols, if the errors are only caused by dark counts ($e_x^{sq} = 0$), then the bit error rate threshold is $\frac{1}{2}$, which implies there is no bound on the distance for communication. However, we must keep in mind that this result is derived using many special conditions. In practice, e_x^{sq} is non-zero and, since there is decoherence in the channel and extrinsic sources of background noise, e_x^{sq} usually increases with the distance of communication. We also assumed that Alice and Bob perfectly know the dark counts rates of their detectors, that they are the same for all detectors, that they are independent of other measurements, and that Eve cannot lower them. However, even if one or more of these assumptions are not respected, it is still possible to slightly modify Eq. C.9, as we explained earlier, and obtain an improvement over Eq. C.5.

In Fig. C.7, we observe that the new method of calculating the key generation rate, using Eq. C.9, improves the achievable distance for PBC00, BB84, and the six-state protocol assuming a single-photon source. For simplicity, we suppose that the dark count probability, C, is the same for all detectors and that e_x^{sq} is fixed and independent of distance. We assume no qubit losses at $l = 0$, where l is the length of the channel, and neglect events when two different detectors fire simultaneously. Under these conditions, for BB84 and the six-state protocol, $p_c^{sq} \approx \eta$ and $p_c^{dk} \approx 2C(1 - \eta)$, where $\eta = e^{-Al}$ is the probability that a photon successfully travels through the channel and A is the attenuation in the fiber. For PBC00, $p_c^{sq} \approx \frac{1}{2-e_x}\eta$ and $p_c^{dk} \approx 2C(1 - \eta)$. Note that, since $\frac{p_c^{dk}}{p_c}$ is always equal or higher in PBC00 than for BB84 or the six-state protocol, PBC00's maximum achievable distance is lower for the same bit error rate.

We now consider the case where Alice uses a source that follows a Poisson distribution ($p_c^{mq} \neq 0$). We only provide the result for BB84, but our arguments are valid for other QKD protocols, including B92, PBC00, and the six-state protocol.

Decoy states [88] could be used to evaluate p_c^{sq} and e_x^{sq} precisely. Ref. [188, 122] explain how Alice could randomly vary the average photon number, μ, of her source to obtain, from statistics, precise estimates of the rate of conclusive results associated with single-photon pulses, $p_c\omega$, and the corresponding bit error rate, e_x^1. p_c^{sq} and e_x^{sq} can be easily derived from the following two relations: $p_c\omega = p_c^{sq} + 2Ce^{-\bar{\mu}}\bar{\mu}(1 - \eta)$ and $e_x^1 = e^{-\bar{\mu}}\bar{\mu}(\eta e_x^{sq} + 2C(1 - \eta)e_x^{dk})/(p_c\omega)$, where $\bar{\mu}$ is the global average photon number. Fig. C.8 shows that the decoy state method can also be improved by using Eq. C.9.

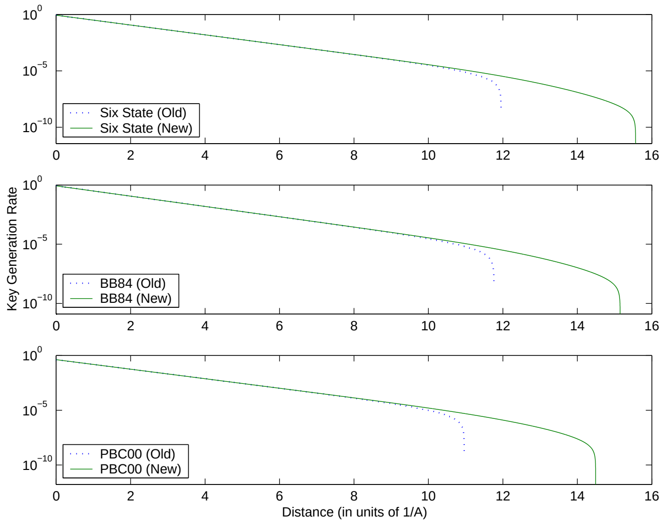


Figure C.7: Semi-log graph of the key generation rate of PBC00, BB84, and the six-state protocol as a function of distance, l, for $e_x^{sq} = 0.01$ and $C = 10^{-6}$ calculated using the old method (Eq. C.5) and the new one (Eq. C.9) assuming a perfect single-photon source ($p_c^{mq} = 0$).

If we don't use decoy states, a worst case estimate of p_c^{sq} and e_x^{sq} is possible. However, Eq. C.9 provides only a small improvement since, without decoy states, multi-photon pulses are usually a much more important limiting factor than dark counts.

In this paper, we showed that a high confidence in the stability of the dark counts of the detectors against the possible attack of an eavesdropper implies a significant increase of the robustness of most QKD protocols against dark counts, one of most important contributors of noise in quantum communication. We studied particularly the cases of PBC00, BB84 and the six-state protocol. We explained how to get an improvement of the secret key generation rate and of the achievable distance in some non-ideal situations, including when Alice uses a Poissonian photon source, when Alice and Bob know only a lower bound for the dark count rates of their detectors, and when the dark count rates are not uniform over the detectors. Further improvements to the secret key generation rate might come from using two-way error correction [72] and by artificially adding some errors in the key [110].

Our results benefitted from discussions with Daniel Gottesman and Hoi-Kwong Lo, whose contributions are greatly appreciated. We thank Nicolas Gisin who proposed using reverse reconciliation. We also thank Tony Anderson for his assistance. J.-C.B. and R.L. acknowledge support from the Government of Ontario, J.B. and R.L. from NSERC, and R.L. from CIAR, MITACS and ARDA.

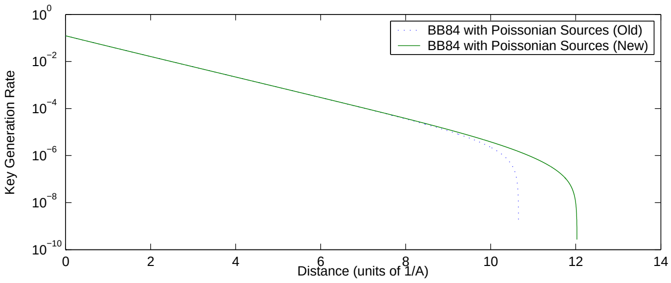


Figure C.8: Semi-log graph of the key generation rate of BB84 as a function of distance, l, for $e_x^{sq} = 0.01$ and $C = 10^{-6}$ assuming a Poissonian source and combined with the decoy state method with $\bar{\mu} = 0.5$. We compare the key generation rates calculated using Eq. C.5 and Eq. C.9.

Appendix D

Some Results Related to NMR Quantum Information Processing

In this appendix, we present two papers for which I have done related work.

D.1 Experimental Implementation of Discrete Time Quantum Random Walk on an NMR Quantum Information Processor

C.A. Ryan, M. Laforest, J.-C. Boileau, and R. Laflamme
Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada.

Abstract: We present an experimental implementation of the coined discrete time quantum walk on a square using a three qubit liquid state nuclear magnetic resonance (NMR) quantum information processor (QIP). Contrary to its classical counterpart, we observe complete interference after certain steps and a periodicity in the evolution. Complete state tomography has been performed for each of the eight steps making a full period. The results have extremely high fidelity with the expected states and show clearly the effects of quantum interference in the walk. We also show and discuss the importance of choosing a molecule with a natural Hamiltonian well suited to NMR QIP by implementing the same algorithm on a second molecule. Finally, we show experimentally that decoherence after each step makes the statistics of the quantum walk tend to that of the classical random walk.

D.1.1 Introduction

The idea of exploiting the quantum mechanical behaviour of a device to gain power in simulating quantum systems was first introduced by Richard Feynman [60]. The field of quantum computing has since grown enormously with the discovery of two algorithmic pillars: Shor's factoring algorithm [171] and Grover's search algorithm [75]. Both of these demonstrate a clear speed-up over their classical counterparts. Following in this path, many other quantum algorithms have been developed that provide a speed-up [97, 168, 56]. A more recent addition to the family of quantum algorithms which demonstrate an exponential speed-up, are those based on the quantum random walk - the quantum version of the successful classical random walk [39].

There is however, a need to explore more than the simple computational properties of the algorithms. They must also be experimentally tested in real devices and their relative ease of implementation compared and considered. In particular, in QIP devices where we are controlling the natural Hamiltonian, it is important to choose a system where the Hamiltonian is amenable to automatic and systematic control. This can be explored by implementing the same algorithm in different molecules and contrasting the performance. Although many different implementation schemes have been proposed for the quantum random walk algorithm, using for example trapped ions [186], an optical lattice [53], cavity QED [159], or an optical cavity [99], these have not been tested. The only experimental test of a quantum walk is the continuous time version of a quantum walk on a square using a two qubit nuclear magnetic resonance (NMR) quantum information processor (QIP) [54]. This work showed the contrast between a classical and quantum random walk and showed the influence of entanglement on the probability distribution of the quantum walk. Here, we present an experimental proof of principle experiment of a discrete time quantum walk on a square. The effects of decoherence on the quantum random walk has been investigated by several authors and indeed, it may offer some benefits [94, 32]. Therefore, we also explored the quantum to classical transition of our walk under the addition of decoherence to the coin. Furthermore, we compared and contrasted two different control schemes and molecules by implementing the algorithm on two molecules.

D.1.2 Quantum Random Walks

In the development of deterministic classical randomized algorithms, the methods of Markov chains and random walks have played a fundamental role [132]. These algorithms can be divided into two categories: continuous time random walks when the walker has a probability per unit time to make a move; and, discrete time random walks where the walker moves at defined time-steps. Since these processes are stochastic, it is not surprising that they have quantum counterparts. The quantum versions however, show remarkable differences with their classical analogues. The continuous time quantum walk (CTQW)[59] has been shown to provide an exponential speed-up in propagation through a graph [40, 39]. The discrete time quantum walk (DTQW) [5] plays an important role in the speed up of a quantum algorithm design for spatial searching [170, 1, 6].

One step of a classical discrete time random walk on a circle with n nodes, denoted by $\{0, \ldots, n-1\}$, is performed by repetition of the following two steps: (1) the walker first flips a coin and then (2) moves either clockwise or counterclockwise depending on the outcome of the coin toss.

If we perform a quantum mechanical treatment of the situation, we can label the nodes with a mutually orthonormal set of state vectors $\{|i\rangle\}_{i=0}^{n-1}$. The coined DTQW on the circle can be seen as "quantumly" flipping a coin degree of freedom using a unitary operation and then coherently moving the walker position degree of freedom clockwise, or counterclockwise, conditioned on the state of the coin [3]. For a Hadamard walk, the coin flipping operation is simply the Hadamard gate described by the matrix,

$$\hat{H} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Now, the conditional shift operator is defined as

$$\begin{aligned} \hat{S}|H\rangle|i\rangle &= |H\rangle|i\ominus 1\rangle \\ \hat{S}|T\rangle|i\rangle &= |T\rangle|i\oplus 1\rangle, \end{aligned}$$

where $\oplus$ and $\ominus$ are understood to be addition and subtraction modulo n and $|H\rangle$ and $|T\rangle$ describe the two basis states of the coin. Therefore, if the walker is in position $|i\rangle$, he will move clockwise to the position $|i\ominus 1\rangle$ if the coin in the state $|H\rangle$, or counterclockwise to $|i\oplus 1\rangle$ if the coin in the state $|T\rangle$. We can write this operator as

$$\hat{S} \;=\; \sum_{i=0}^{n-1} (|H\rangle\langle H|H \otimes |i\ominus 1\rangle\langle i\ominus 1|i + |T\rangle\langle T|T \otimes |i\oplus 1\rangle\langle i\oplus 1|i)\,.$$

Then, one step of the DTQW is defined as applying the operator

$$\hat{W} \;=\; \hat{S}(\hat{H}\otimes \mathbb{1}).$$

On a circle, this type of algorithm shows destructive interference effects and a probability distribution that is periodic in time. The contrasting dynamics for the classical and quantum random walks are shown in Fig. D.1. As opposed to the classical walk where the probability is always spread out, the quantum walk has steps where the probability amplitudes interfere such that all the probability comes back to one node. Furthermore, this walk is periodic in that after eight steps, the corresponding propagator is equal to the identity and the system comes back to its original state.

In our experimental setup we have three qubits available, which allows one qubit to describe the coin state and two for the position state. Thus, we have $n = 4$, and we are performing a discrete quantum walk on a square. The shift operator defined in Eq. D.1 would require a complicated quantum circuit involving a Toffoli gate. We can simplify the circuit required by using a shifting operator that moves the walker along a direction vector, i.e. horizontally or vertically (this also is analogous to the random walk on the hypercube [131]). Therefore, if we label the corners of the square as shown in Fig D.2, the shift operator on the three qubit register becomes,

$$\begin{aligned} \hat{S} &= \hat{P}_H \hat{X}^2 + \hat{P}_T \hat{X}^3 \\ &= (\hat{P}_H \hat{X}^2 + \hat{P}_T)(\hat{P}_T \hat{X}^3 + \hat{P}_H) \\ &= (\hat{X}^1 Cnot^{1,2} \hat{X}^1) Cnot^{1,3}, \end{aligned}$$

with X denoting the standard σ_x Pauli matrix, $P_{H/T}$ are the projectors on the two coin states, and the superscript indicates on which of the qubits the action is performed. Here, it is understood that the first qubit represents the coin and the second and third, the position register. The resulting probabilities for each step are shown in Table D.1.

	Classical				Quantum			
Corner	0	1	2	3	0	1	2	3
Step 0	1	—	—	—	1	—	—	—
Step 1	—	0.5	—	0.5	—	0.5	—	0.5
Step 2	0.5	—	0.5	—	0.5	—	0.5	—
Step 3	—	0.5	—	0.5	—	—	—	1
Step 4	0.5	—	0.5	—	—	—	1	—
Step 5	—	0.5	—	0.5	—	0.5	—	0.5
Step 6	0.5	—	0.5	—	0.5	—	0.5	—
Step 7	—	0.5	—	0.5	—	1	—	—
Step 8	0.5	—	0.5	—	1	—	—	—

Table D.1: Probability to be in each of the corner states as denoted in Fig. D.2. While in the classical random walk the probability always remains spread out between two corners, in the quantum random walk all the probability returns to one corner at certain time steps.

D.1.3 Liquid state NMR quantum information processing

The basic principles

A liquid state NMR QIP consists of an ensemble of roughly 10^{20} identical molecules dissolved in a liquid solvent. Due to the fast tumbling motion of the molecules, they are essentially decoupled from each other; ideally all the molecules have the same evolution. We can think of the quantum register made of qubits that correspond to the spin $\frac{1}{2}$ nuclei within each molecule. The sample is placed in a strong homogeneous magnetic field which provides the quantization axis and causes the spins to precess around the axis of the field. It is possible to implement single qubit gates using radio-frequency (r.f.) pulses resonant with the precession frequency, which can effect a rotation about any axis orthogonal to the axis of the external field. Two qubit gates are effected through the coupling from the natural Hamiltonian, which produces a controlled phase gate[113].

If the molecule used contains n distinguishable nuclei and the magnetic field is aligned along the z-axis, then the system Hamiltonian is approximated by,

$$\hat{\mathcal{H}} = \pi \sum_{i=1}^{n} \nu_i \hat{Z}_i + \frac{\pi}{2} \sum_{i>j} J_{ij} \hat{Z}_i \otimes \hat{Z}_j,$$

where ν_i is the Larmor frequency of spin i in Hz, J_{ij} is the coupling strength between spin i and j in Hz and Z in the conventional Pauli operator σ_z. The interaction part of the Hamiltonian can be approximated to the above Ising form (weak coupling regime or secular approximation) only if the difference between any two nuclei Larmor frequencies is much greater than the coupling between the nuclei. We can also turn off the coupling between any two spins as needed by applying refocussing r.f. pulses.

Implementing dephasing in NMR

We can apply a controllable amount of decoherence to selected spins using gradient techniques in NMR. Consider only one nucleus with state ρ and suppose we work in the rotating frame of that spin. On a NMR spectrometer, it is possible to apply a gradient to the external magnetic field. During the time that the gradient is applied, the spins will precess at different frequencies depending on their position in the sample. The state of the ensemble will then be given by an average over the observable sample,

$$\rho' \;=\; \frac{1}{2a} \int_{-a}^{a} e^{-i(\alpha' \gamma t z/2)\hat{Z}} \rho e^{i(\alpha' \gamma t z/2)\hat{Z}} dz,$$

where $2a$ is the length of the sample, t is the interval of time the gradient is being applied and $\alpha' = \alpha/\hbar$ and $\gamma = \nu/B_z$, the gyro-magnetic ratio of the nucleus. If we compute the integral, it can be shown that,

$$\begin{aligned}\rho' &= (1-p)\rho + pZ\rho Z, \\ &\quad p = \frac{1}{2}\left(1 - \frac{1}{\alpha'\gamma t a}\sin\left(\alpha'\gamma t a\right)\right),\end{aligned}$$

which is the exact form of a z-dephasing decoherence. The amount of dephasing can be controlled by the strength and time of the gradient pulse. Particular spins can be protected from the applied decoherence by applying a 180 degree rotation and applying a second gradient of the same strength and time. This second gradient will reverse the dephasing of the rotated spins and double it on the spins that were not rotated.

D.1.4 The experiment

We implemented the quantum walk algorithm on two molecules: trans-crotonic acid and trichloroethylne (TCE). This allowed us to compare the quality of two different methods of control and the merits of the two molecules.

Implementation on crotonic acid

The seven qubit molecule trans-crotonic acid (four carbons, two hydrogens and one methyl group) has been used in experimental demonstrations of quantum algorithms, such as quantum error correction [100, 28], and quantum simulations[135]. In this experiment, we used the carbon back-bone of labeled trans-crotonic acid in a solution of deuterated acetone. The hydrogen nuclei were decoupled using standard heteronuclear decoupling techniques[165]. We used C_3 as the coin and C_2 and C_4 as the position register (see Fig. D.3). C_1 was used as a labeling spin to ease the creation of the initial state. On a Bruker DRX Avance 600 NMR spectrometer, the molecule has the Hamiltonian parameters shown in Fig. D.3.

Since our system is homonuclear, the control of individual qubits is achieved through soft gaussian-like r.f. pulses at the Larmor frequency of the target nucleus. The length of the soft pulses is of the order of the inverse of the smallest chemical shift difference with the other nuclei. In our experiment the length of the selective pulses on C_1 and C_2,C_3,C_4 were 192μs and 704μs respectively.

Initial state preparation The experiment required the initial state

$$\rho_{in} = * \otimes |000\rangle\langle 000|000 = *(\mathbb{1}+\hat{Z})(\mathbb{1}+\hat{Z})(\mathbb{1}+\hat{Z}).$$

We created the labeled pseudo-pure state $Z000$ (using the notation $C_1C_2C_3C_4$) following the spatial averaging technique elaborated in [101].

Pulse sequence implementation The unitary of one step of the DTQW from Eq D.1 was translated to a sequence of pulses and coupling gates as shown in Fig. D.4. Although many pulse sequences are possible through the use of commutation rules, this particular one was designed to be the most efficient due to the cancellations possible during multiple step sequences. Moreover, the ZZ gates are achieved simultaneously, which shortens the overall pulse sequence, thus reducing decoherence effects. Commutation rules were also used to cancel pulses between the final step and the readout pulses. *Since the coupling constants between C_2 and C_3 is slightly lower then that between C_3 and C_4, the simultaneous $\pi/2$ couplings have been achieve by first letting those three qubits evolve freely until the coupling between C_3 and C_4 reach $\pi/2$. Thereafter, C_4 is refocus during the interval needed to C_2 and C_3 to terminates there coupling.* The ideal pulse sequence of rotations and couplings was then input into a pulse sequence compiler which numerically optimized the timing and phases of the pulses. The compiler pre-simulates the selective r.f. pulses using an efficient pair-wise simulation and then decomposes the simulated unitary into phase and coupling errors sandwiching the ideal selective pulse. These errors can then be taken into account by the refocussing scheme and phase of the pulses, so that the overall unitary is as close to the desired one as possible.

Since we are concerned with the final state of only three qubits in this experiment, complete state tomography is still feasible. On a three qubit system in NMR, only seven different readout pulses are required to rotate each term of the density matrix into observable simple single coherences [1]. And, since we were operating on a homonuclear system, observing the signal from all spins in one experiment was possible, with some post-processing to adjust for the correct phase of each individual rotating frame. The coupling between the labeling spin C_1 and the other three qubits is resolvable and so the presence of the labeling spin does not interfere with the tomography of C_2, C_3 and C_4.

Experimental results For the state tomography each of the peaks in the spectra were fitted using absorption and dispersion Lorenzian peaks. The full density matrix was then reconstructed by appropriately summing up the corresponding Pauli terms. Where two experiments gave values for the same density matrix terms, the values were simply averaged. As we observed only C_2, C_3 and C_4, the term $ZIII$ could not be determined. A suitable amount of that term was subsequently added to the density matrix so as to make the initial state as close to $Z000$ as possible. This amount was then kept constant for the density matrix reconstruction in subsequent experiments.

To quantify the success of our experiments, we computed the fidelity of the experimental density matrix to both the ideal and simulated results. In NMR, all states are nearly

[1] Readout pulses yII,IIy,IIx,yyI,Ixx,yyy,xxx are sufficient

completely mixed and the fidelity measure introduced in [61] is appropriate. We can compare one density matrix to another using the following formula,

$$F^{A,B} = \frac{Tr(\rho^A \rho^B)}{\sqrt{Tr((\rho^A)^2)}\sqrt{(Tr((\rho^B)^2)}}.$$

We made two comparisons. Firstly we compare the experimentally determined density matrix to the theoretically expected result. The theoretical result is achieved by multiplying the ideal initial state by the ideal propagator. To investigate how well we understand our control of the system, we also compare the fidelity of the results from a simulation of the experiment to the theoretical result.

	Experimental	Simulated
Step0	98 ± 5	—
Step1	97 ± 5	98
Step2	98 ± 5	98
Step3	92 ± 5	98
Step4	99 ± 5	98
Step5	94 ± 5	97
Step6	96 ± 5	97
Step7	96 ± 5	97
Step8	87 ± 4	97

Table D.2: Fidelities (in percent) of experimental and simulated results. The first column gives the fidelity of experimental density matrix determined from the tomography, with respect to the theoretical expected density matrix. The second column gives the fidelity of the simulation results. Errors are estimated from the fitting procedure. Note that since computer simulation of the spatial averaging that occurs during the pseudo-pure preparation is difficult and inaccurate, the initial state for the simulation was the experimental pseudo-pure state determined from the tomography. The fluctuations observed in the fidelity come from uncertainties in the fit and instabilities in the spectrometer over the course of the experiment.

The fidelities of simulated and experimental results are compared in Table D.10 *and Fig. D.5-a give a sample of the fully reconstructed density matrix for* C_2, C_3 *and* C_4. The loss of fidelity in our experiment, over and above that of the simulated control errors is explained from three sources which are not taken into account by the simulation. We have losses from T2 relaxation. Although our pulse sequence is short compared with the T2 relaxation times, during the quantum walk algorithm, the state is often in high coherences, which decay much faster than the simple T2 time. Inhomogeneities in the strong magnetic field also cause extra relaxation and dephasing. Further losses come from inhomogeneities of the r.f. field used to implement rotations and pulse angle mis-calibration.

Addition of decoherence on the coin In a subsequent experiment, we added dephasing decoherence to the entire qubit register using the technique described in section D.1.3. We expect the behaviour of the quantum walk should converge to the classical walk as the decoherence becomes complete after each step. To demonstrate this claim experimentally, we implemented the quantum random walk for four steps, adding decoherence of a certain strength between each step of the walk. The differences between quantum and classical walk is manifested in the different probabilities of being in each of the corners after each step. The results are shown in table D.3 for gradient strengths corresponding to no, partial and full decoherence.

	Quantum Walk with Decoherence											
	None				Partial				Full			
Corner	0	1	2	3	0	1	2	3	0	1	2	3
Step 0	100	0	4	−2	100	0	4	−2	100	0	4	−2
Step 1	2	57	−1	43	2	58	−2	44	0	59	−2	45
Step 2	57	1	44	−1	50	7	40	5	51	4	460	1
Step 3	7	14	3	78	2	14	1	84	−1	53	−3	53
Step 4	15	−1	84	3	19	1	78	4	50	0	53	−1

Table D.3: Classical versus estimate of quantum probability to be in each corner of the square for one through four steps (c.f. Table D.1). The quantum results were obtained for gradient strengths corresponding to no, partial and full decoherence. The experimental results were obtained by reconstructing the density matrix using the same fitting software used before and then applying the position measurement projectors to the reconstructed density matrix.

The divergence between the classical and quantum walk shows most clearly in steps three and four. Whether the walk is classical or quantum, steps one and two yield the same measurement probabilities for the position (however the quantum version with decoherence will have coherent superposition states). Analyzing the data from step 3 and 4, one can see that the quantum interference present so clearly in the quantum walk with no decoherence, is less obvious as the amount of decoherence increases. Instead of the probability all collecting in one corner it remains spread out between two opposite corners - the same as in the classical walk. *This can also be seen in Fig. D.5-b by the reduction of the off-diagonal* *terms of the density matrix as the decoherence become stronger.*

The probabilities even with zero gradient strength do not correspond perfectly to the ideal quantum walk. We believe these errors come from two sources. Because the gradient does not commute with any pulses we were not able to use commutation rules to reduce the number of pulses during multiple step experiments. Furthermore, gradient methods are hampered by diffusion and multiple gradients may lead to a return of signal that was "erased" by a previous gradient.

Comparison with the TCE molecule

For comparison purposes and to show the importance of choosing a molecule with good characteristics in liquid state NMR quantum information processing, we show our results from our initial attempt to implement the quantum walk on the molecule tri-chloroethylene (TCE) - a molecule with which we have much less control due to the presence of strong coupling. The molecule has been used for some initial demonstrations of quantum algorithms [45]. A diagram of the molecule and the parameters of its Hamiltonian are shown in Fig. D.6.

Pseudo-pure state preparation Since the TCE molecule contains only three qubits, we are unable to create the labeled pseudo-pure state that we used in the crotonic acid experiments. Instead, we chose to use temporal averaging and add three separate experiments to achieve the initial state $|000\rangle$. The three different initial states we used are,

$$\begin{aligned}\rho_1 &= \hat{Z} \otimes (\mathbb{1} + \hat{Z}) \otimes (\mathbb{1} + \hat{Z}) \\ \rho_2 &= \mathbb{1} \otimes \hat{Z} \otimes (\mathbb{1} + \hat{Z}) \\ \rho_3 &= \mathbb{1} \otimes \mathbb{1} \otimes (\mathbb{1} + \hat{Z})\end{aligned}$$

If we add the results of these three experiments, it is equivalent to having performed the algorithm on the initial state

$$\begin{aligned}\rho_{in} &= \rho_1 + \rho_2 + \rho_3 \\ &= (\mathbb{1} + Z) \otimes (\mathbb{1} + Z) \otimes (\mathbb{1} + Z) \\ &= |000\rangle\langle 000|000\end{aligned}$$

Since there is only one hydrogen nucleus in the molecule, we can use broadband hard pulses to control it. One useful property of the TCE molecule in a 600Mhz spectrometer, is that the J-coupling between the two carbons is almost exactly 10.5 times smaller than the difference in chemical shift ($\Delta\nu$). Therefore, during the time for a $\frac{\pi}{2}\hat{Z}\hat{Z}$ coupling gate between the two carbons ($\Delta t = \frac{1}{2J_{C_1C_2}}$), the relative chemical shift evolution of C_2 with respect to C_1 will be $\theta = 2\pi\Delta\nu\Delta t = \frac{\pi\Delta\nu}{J_{C_1C_2}} = -10.5\pi = -\pi/2 \text{ mod} 2\pi$. Therefore, in the reference frame rotating at the Larmor frequency of C_1, every time there is a $\pi/2$ coupling between the carbons, an extra $R_z^{-\pi/2}$ is naturally performed on C_2.

The chemical shift difference between the two carbons is small and the coupling between them large, so selective pulses were impossible to achieve using the same technique of gaussian-shaped pulses used in the crotonic acid experiments. These pulses would be very long (roughly 5 ms) and the large coupling errors that would occur during the pulse would be difficult to refocus. Instead, it was possible to perform selective pulses using hard pulses and the chemical shift evolution. To illustrate the technique, we demonstrate how to perform a selective $\frac{\pi}{2}$ rotation of C_2. If we use a reference frame rotating at the Larmor frequency of C_1, then, during a time $\tau = \frac{1}{4\Delta\nu}$, the spin C_1 will not precess while C_2 will undergo a rotation of $-\pi/2$ around the z-axis. Since, $\frac{1}{4\Delta\nu}$ is much less than the coupling time $\frac{1}{2J_{C_1C_2}}$, we can ignore the coupling between the two carbons and refocus only the hydrogen. Using this selective z-rotation combined with hard pulses which rotate the two carbons together, we can perform a $\frac{\pi}{2}$ rotation with phase ϕ on only C_2 as follows:

$$\begin{aligned}(R^{\pi/2}_{\phi-\pi/2} \quad &\otimes \quad R^{\pi/2}_{\phi-\pi/2})(\mathbb{1} \otimes R_z^{-\pi/2})(R^{\pi/2}_{\phi+\pi/2} \otimes R^{\pi/2}_{\phi+\pi/2}) \\ &= \quad \mathbb{1} \otimes R^{\pi/2}_{\phi-\pi/2} R_z^{-\pi/2} R^{\pi/2}_{\phi+\pi/2} \\ &= \quad \mathbb{1} \otimes R^{\pi/2}_{\phi}\end{aligned}$$

Similar pulse sequences can be derived to perform a π rotation on C_2 and selective pulses on C_1. Because of the different form of selective pulses used, the pulse sequences were written and optimized by hand. This required a different pulse sequence implementation of the quantum walk unitary which avoided as much as possible selective pulses and z-rotations where possible. The one z-rotation used, is a natural outcome of the $C_1 - C_2$ coupling gate as described above. This alternative pulse sequence is shown in Fig. D.7.

Experimental Results Fidelity results, similar to those calculated for the crotonic acid experiments are shown in Table D.4 *and a sample of reconstructed density matrices can be sen in Fig. D.5-c.* Clearly this experiment was not as successful as the implementation on the crotonic acid molecule. There are two main reasons for this loss of fidelity. Firstly, the chemical shift difference between the two carbons is very small. Because of this, the secular approximation no longer holds and thus the coupling between the two carbon spins can no longer be approximated by the Ising form $Z_{C_1} \otimes Z_{C_2}$. Indeed, it has to take all the strong coupling terms into account, i.e $\vec{S}_{C_1} \cdot \vec{S}_{C_2} = X_{C_1} \otimes X_{C_2} + Y_{C_1} \otimes Y_{C_2} + Z_{C_1} \otimes Z_{C_2}$.

Unfortunately, this strong coupling renders our ideal ZZ gates much less precise. Every coupling gate performed added XX and YY error terms which we could not refocus. This

	Experimental	Simulated
Step 0	98 ± 6	—
Step 1	85 ± 5	96
Step 2	82 ± 4	94
Step 3	70 ± 4	93
Step 4	80 ± 4	90
Step 5	76 ± 4	89
Step 6	65 ± 4	86
Step 7	53 ± 4	84
Step 8	43 ± 4	83

Table D.4: Experimental and simulated fidelities (in percent) for the implementation of 8 steps of the DTQW on the molecule TCE. Again, simulation of the pseudo-pure state was not performed.

coupling also caused problems during our selective carbon rotations. Although the coupling is small, there is an unrefocusable coupling of $\frac{\pi J_{C_1C_2}}{4\Delta\nu} \approx 4.27°$. Our only way to minimize these errors was to optimize the delay times analytically and from numerical simulations. However, these did not correspond well to the experimentally determined optimal values. This point also clearly demonstrates the second reason for the less satisfactory results on TCE. We were unable to use the numerical optimization of the pulse sequence compiler used for crotonic acid. The compiler provides a systematic and reliable way to produce pulse sequences which implement unitaries with high fidelity and is clearly superior to writing and optimizing pulse sequences by hand. These experiments also showed the limits of our simulator. For the crotonic acid experiments, where only soft pulses were used, the r.f. power applied changed slowly and the simulator was faithful to what r.f. power the spins were experiencing. In TCE, where control was achieved only through short hard pulses, other effects such as phase transients enter and the spins might experience an r.f. field much different from the ideal square pulse simulated. To fully understand the issues surrounding hard pulse control a much more detailed study of the probe response must be undertaken. This underlines a key point: control of a more complex and strongly coupled system could be obtained through sophisticated control techniques such as strongly modulating pulses [61] ; however, it seems prudent to invest the effort in a wise choice of molecule.

D.1.5 Conclusion

We have presented the first experimental implementation of a coined discrete time quantum walk. It showed a clear difference with the classical coined quantum walk, since the DTQW possesses destructive interference and periodicity in its evolution. A proof of principle like this lays down the path to more elaborate experiments using discrete quantum walks, such as the database search, walks on hypercube or N-nodes circle or a more profound study of the effect of decoherence on the walk. This paper also demonstrates the importance of choosing a natural Hamiltonian well suited to automated control in the context of quantum information processing.

D.1.6 Acknowledgments

C. R, M. L, J.-C. B. and R. L. would like to thank M. J. Ditty for technical support with the work on the spectrometer and C. Negrevergne for helpful discussion. This work has been supported by NSERC, ARDA and CFI.

D.2 Using error correction to determine the noise model

M. Laforest, D. Simon, J.-C. Boileau, J. Baugh, M. J. Ditty and R. Laflamme
Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada.

Abstract: Quantum error correcting codes have been shown to have the ability of making quantum information resilient against noise. Here we show that we can use quantum error correcting codes as diagnostics to characterise noise. The experiment is based on a three-bit quantum error correcting code carried out on a three-qubit nuclear magnetic resonance (NMR) quantum information processor. Utilizing both engineered and natural noise, the degree of correlations present in the noise affecting a two-qubit subsystem was determined. We measured a correlation factor of $c = 0.5 \pm 0.2$ using the error correction protocol, and $c = 0.3 \pm 0.2$ using a standard NMR technique based on coherence pathway selection. Although the error correction method demands precise control, the results demonstrate that the required precision is achievable in the liquid-state NMR setting.

D.2.1 Introduction

The idea of using quantum mechanical systems as information processing devices was proposed more than two decades ago [60], and yet, experimental realization of such devices remains a challenge. Ultimately, all physical realizations are faced with the presence of decoherence, or noise, caused by uncontrollable interactions with the environment [192].

To prevent the loss of coherence in the quantum mechanical processor, the theory of quantum error correction (QEC) has been developed [172, 175, 20]. QEC works by encoding the state of a system lying in a certain Hilbert space into a state in a larger Hilbert space. The encoding is designed to make possible the recovery of the original information after noise has acted on the overall system, through decoding and syndrome measurement, as long as the the noise level falls below a certain threshold [102, 103]. Many quantum error correction codes (QECC) have been developed for specific classes of noise models. For example, there are codes that can correct arbitrary single qubit errors [172, 175, 20, 115] but fail to correct multi-qubit errors. This work shows how this failure can be used to extract information about the noise of the system.

Most QECC are developed for independent, or uncorrelated, error models, meaning that the errors happening on one qubit are assumed to be independent of the errors on other qubits. Clearly, knowing whether or not there exist correlations in the noise model plays an important role in choosing the best QECC for a given system.

The noise model can be determined exactly by performing process tomography [42, 144]. However, the number of experiments required for complete tomography grows exponentially in the number of qubits. Often, process tomography is not needed and important (but partial) information about the noise can be extracted from fewer experiments. Here, we demonstrate the use of a three-qubit QECC to extract the two-qubit noise correlation factor under a transverse relaxation process in nuclear magnetic resonance (NMR) (e.g. T_2 relaxation).

Transverse relaxation is the main source of decoherence in liquid state NMR. NMR was used to perform the first experimental implementation of QEC [45], where it was shown that the three-qubit QECC could correct single-qubit phase flip errors caused by T_2 relaxation. Here, we will first briefly review the fundamentals of NMR, then model the noise present in such systems and show how noise correlations can affect the fidelity of the QEC protocol. We will describe a series of natural and engineered noise experiments for determining the two-spin noise correlations present for the ^{13}C subsystem of acetyl chloride (dissolved in deuterated chloroform). The experimental results are in agreement with those obtained using the standard NMR technique of coherence selection. In light of the exquisite sensitivity of these experiments to control imperfections, the results also demonstrate the high degree of precision attainable in controlling nuclear spins in liquid state NMR.

The results demonstrate that QEC can not only be used for correcting the effects of decoherence, but can also help to characterize the nature of those errors. Moreover, as QEC is a requirement for scalable quantum information processing (QIP), this methodology is universal for probing noise correlations in physical systems suitable for QIP.

D.2.2 NMR quantum computing

Liquid state nuclear magnetic resonance has proven to be a useful system for experimentally benchmarking small-scale quantum information processing devices [114, 139, 45, 101, 134]. A NMR quantum information processor consists of an ensemble containing of order 10^{20} molecules with spin-$\frac{1}{2}$ nuclei dissolved in a liquid solvent. Placed in a strong homogeneous magnetic field, nuclear spins precess about the direction of the field, defined conventionally as the z-axis. The rate at which the spins precess is the Larmor frequency, and nuclei with distinct Larmor (resonance) frequencies can be mapped to qubits. In the liquid state, picosecond-scale rotation and translation of the molecules causes spins on separate molecules to effectively decouple on the NMR timescale. Therefore, to a very good approximation, all molecules of the same type experience identical environments and the Hilbert space of the nuclear spin ensemble can be taken as that of a single molecule. Moreover, the rotational degree of freedom causes the internal dipolar interaction between the spins on each molecule to vanish. At thermal equilibrium, the Boltzmann distribution gives a slight excess of spins pointing along the $+z$ direction, so that an average magnetization is present along $+z$.

Control of the qubits is achieved by a radio-frequency (RF) Hamiltonian in which the frequency, phase and duration of the RF can be controlled externally. Single-qubit rotations are performed using RF pulses resonant with the Larmor frequency of the targeted qubit. By varying the RF duration and phase, rotations of arbitrary angle can be generated about any axis in the xy-plane. Two-qubit operations additionally use the natural coupling terms present in the internal Hamiltonian, which will be elaborated below.

D.2.3 Hamiltonian and noise model

The nuclear spin Hamiltonian in liquid-state NMR is composed of two types of terms, one corresponding to the single-spin Zeeman interaction (the term leading to precession) and bilinear terms corresponding to the scalar spin-spin coupling (J-coupling). For a molecule with N spin-1/2 nuclei, the weak coupling Hamiltonian is given by

$$\hat{\mathcal{H}} \;=\; \frac{1}{2}\sum_{i=1}^{N} 2\pi\nu_i \hat{Z}_i + \frac{\pi}{2}\sum_{i<j} J_{ij}\hat{Z}_i\hat{Z}_j$$

where ν_i are the Larmor frequencies, J_{ij} is the coupling strength between spins i and j, and $\hat{Z}_i$ is the z Pauli matrix for spin i. Note that when the condition $|\nu_i - \nu_j| >> J_{ij}/2$ does not hold (strong coupling regime), the scalar coupling operator takes the more general form $\vec{\sigma}_i \cdot \vec{\sigma}_j = \hat{X}_i\hat{X}_j + \hat{Y}_i\hat{Y}_j + \hat{Z}_i\hat{Z}_j$.

Despite the motional averaging that occurs in the liquid state, the $\hat{Z}\hat{Z}$ part of the intermolecular dipolar Hamiltonian is capable of creating relaxation, while the $\hat{X}\hat{X} +$

$\hat{Y}\hat{Y}$ part still averages to zero due to the weak coupling approximation[117]. This $\hat{Z}\hat{Z}$ interaction couples with the molecular motion and give rise to rapidly fluctuating local magnetic field, which effectively presents itself as a variation of the Larmor frequencies. This process is known as transversal relaxation.

Consider a single spin qubit surrounded by an environment $\mathcal{E}$ consisting of N other spins -$\frac{1}{2}$. The dipolar coupling between the qubit and its environment is described by the unitary evolution

$$\hat{U} = \prod_{j\in\mathcal{E}} e^{-ib_j\hat{Z}\hat{Z}_j}$$

where b_j is the interaction strength between the qubit and the j^{th} spin of the environment for a certain amount of time.

The global system can be assumed to initially be in the state

$$\rho_{glob} = \rho_{ini} \otimes \rho_{\mathcal{E}}$$

where ρ_{ini} represents the initial state of the qubit and $\rho_{\mathcal{E}}$ is the state of the environment. After the interaction with its environment, the final state of the qubit will be given by partial tracing the environment system. Moreover, the interaction strengths have a certain distribution of value $q(\vec{b})$, so that the state affected by the noise have the form

$$\rho_f = \int d\vec{b}q(\vec{b}) \sum_{a\in\{0,1\}^N} \langle a|\rho_{\mathcal{E}}|a\rangle e^{-i\xi_a\hat{Z}}\rho_{ini}e^{i\xi_a\hat{Z}},$$

where we have defined $\xi_a = \sum_m b_m(-1)^{a_m}$, a_m being the m^{th} digit of a. In room temperature liquid state NMR, the deviation of the state of the environment from the completely mixed state is negligible, so that $\langle a|\rho_{\mathcal{E}}|a\rangle = 1/N$. Because the environment is isotropic, the distribution $q(\vec{b})$ is a symmetric function of the b_j's. Therefore, the summation over a in Eq. D.1 can be absorbed in a new distribution of $\vec{b}$ and by letting $a = 0^{\otimes N}$.

The final state is then represented by

$$\rho_f = \int d\alpha p(\alpha)e^{-i\alpha\hat{Z}}\rho_{ini}e^{i\alpha\hat{Z}}$$

where $\alpha = \sum_m b_m$ and $p(\alpha)$ is the distribution of α which takes into account the new distribution of the b_j's. The interaction of the qubit with the environment causes an incoherent averaging of z rotations, which is equivalent to a variation of the Larmor frequency of the qubit. In liquid state NMR, N is a large number and the central limit theorem indicates that α has a gaussian distribution. For a M qubit system, this model generalizes to

$$\rho_f = \int d\vec{\alpha}p(\vec{\alpha})e^{-i\vec{\alpha}\cdot\vec{\hat{Z}}}\rho_{ini}e^{i\vec{\alpha}\cdot\vec{\hat{Z}}}$$

where $\vec{\alpha} = (\alpha_1, \ldots, \alpha_M)$, $\vec{\hat{Z}} = (\hat{Z}_1, \ldots, \hat{Z}_M)$ and $p(\vec{\alpha})$ is the multivariate gaussian distribution [184]

$$p(\vec{\alpha}) = \frac{1}{\sqrt{(2\pi)^M |\hat{\Sigma}|}} e^{-\frac{1}{2}\vec{\alpha}^T \cdot \hat{\Sigma}^{-1} \cdot \vec{\alpha}}.$$

$\hat{\Sigma}$ is the covariant matrix, or the correlation matrix, which takes the form

$$\begin{aligned} \hat{\Sigma}_{ii} &= \langle \alpha_i^2 \rangle \\ &= \sigma_i^2 \\ \hat{\Sigma}_{ij} &= c_{ij}\sqrt{\sigma_i \sigma_j} \end{aligned}$$

where σ_i^2 is the variance of α_i. c_{ij} is the correlation factor between α_i and α_j, which has value

$$c_{ij} = \frac{\langle \alpha_i \alpha_j \rangle}{\sqrt{\langle \alpha_i^2 \rangle \langle \alpha_j^2 \rangle}}.$$

From the Cauchy-Schwarz inequality, $0 \leq c_{ij} \leq 1$. For a single qubit, such a noise model will affect the state as

$$|k\rangle\langle k|l \rightarrow e^{\frac{-\sigma^2}{2}(1-\delta_{kl})}|k\rangle\langle k|l$$

From empirical results of transverse relaxation in NMR, the state of a single spin decays in time as

$$|k\rangle\langle k|l \rightarrow e^{-(1-\delta_{kl})\gamma_2 t}|k\rangle\langle k|l$$

where $1/\gamma_2 = T_2$ is the relaxation time constant. The variance of the distribution of the interaction strength of a qubit with its environment can thus be related to its relaxation time constant by

$$\sigma^2 = 2\gamma_2 t.$$

For two qubits, the noise correlation factor will affect the decay of their mutual state as follow:

$$\begin{aligned} |km\rangle\langle km|ln \rightarrow\ & e^{-(1-\delta_{kl})\gamma_2^{(1)}t-(1-\delta_{mn})\gamma_2^{(2)}t-2c_{12}t\eta_{kl}\eta_{mn}\sqrt{\gamma_2^{(2)}\gamma_2^{(2)}}} \\ & \times|km\rangle\langle km|ln, \end{aligned}$$

where $\eta_{ij} = \frac{1}{2}((-1)^i - (-1)^j)$. If correlations in the noise affecting two qubits is present, the transverse relaxation will be faster for a two spin double quantum coherence (e.g. $|00\rangle\langle 00|11$

and $|11\rangle\langle 11|00)$ and slower for a two spin zero quantum coherence (e.g. $|01\rangle\langle 01|10$ and $|10\rangle\langle 10|01)$.

The correlation in the noise on two qubits can be understood through distinguishability. If two nuclei precess at the same Larmor frequency, they are magnetically equivalent and thus see the same environment. The two spins will interact identically with the environment, thus yielding a correlation factor of 1. Two spins of different nuclear specie are distinguishable and the environment will act differently on each of them. No correlation is expected the respective noise i.e. $c_{12} = 0$. If we consider two nuclei of the same species with slightly different Larmor frequency, they are distinguishable enough to perform independent control, but they are chemically "near indistinguishable". The effect of the environment is thus partially correlated, i.e $0 < c_{12} < 1$.

D.2.4 Engineering the noise for two qubits

By explicitly expanding Eq. D.1 for two qubits, the noise model takes a discrete Kraus form,

$$\rho_f = \sum_i p_i \hat{U}_i \rho_{ini} \hat{U}_i^\dagger$$

where the unitary Kraus operators $\hat{U}_i$ and their coefficients p_i are given in Table D.5. One can thus engineer the noise on two qubits with a series of six separate experiments, each of them implementing a different Kraus operator, and then adding the results with the corresponding coefficient. This Kraus decomposition demonstrates that the transversal relaxation in NMR is equivalent to a phase flip error, where the qubits undergo a phase flip given by the operator in the first column of Table D.5 with a probability given by the second column.

$\hat{U}_i$	p_i
$\mathbb{1}$	$\frac{1}{4}\left(1+e^{-\gamma_2^{(1)}t}+e^{-\gamma_2^{(2)}t}+e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}}}\right)$
$\hat{Z}_1$	$\frac{1}{4}\left(1-e^{-\gamma_2^{(1)}t}+e^{-\gamma_2^{(2)}t}-e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}}}\right)$
$\hat{Z}_2$	$\frac{1}{4}\left(1+e^{-\gamma_2^{(1)}t}-e^{-\gamma_2^{(2)}t}-e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}}}\right)$
$\hat{Z}_1Z_2$	$\frac{1}{4}\left(1-e^{-\gamma_2^{(1)}t}-e^{-\gamma_2^{(2)}t}+e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}}}\right)$
$e^{-i\frac{\pi}{4}(\hat{Z}_1+\hat{Z}_2)}$	$\frac{1}{2}e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t}\sinh(2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}})$
$e^{i\frac{\pi}{4}(\hat{Z}_1+\hat{Z}_2)}$	$\frac{1}{2}e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t}\sinh(2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}})$

Table D.5: Kraus decomposition for the correlated noise on two qubits

D.2.5 Determining the correlation factor

This subsection will explain how the noise correlation factor between two spins can be extracted using standard NMR techniques and how quantum error correction can be used to achieve similar results. The details and results of the experimental implementation, as well as a summary of the advantages of this new technique will then conclude this subsection.

NMR techniques

In NMR, measurement of transversal relaxation times (T_2's) is a standard techniques and is implemented through single coherence decay and spin echo [78]. The same technique is applicable to double coherences to extract the noise correlation factor between two spins. Consider the following pulse sequence:

$$\frac{\tau}{2} \rightarrow \pi_1 \pi_2 \rightarrow \frac{\tau}{2}.$$

where τ is a certain time delay and π_i correspond to a π pulse on nuclei i around any axis in the xy-plane. They are used to refocus the field inhomogeneities via spin echo. If we apply such a pulse sequence to a state of the form

$$\rho_{ini} = |00\rangle\langle 00|11,$$

which can be created using standard NMR techniques of coherence selection such as phase cycling [23] or field gradients, the noise model developed earlier predicts that the amplitude of such a state should decay as

$$|00\rangle\langle 00|11 \quad \rightarrow \quad e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}}}|00\rangle\langle 00|11.$$

In NMR, only single coherence state can be detected. A final $\frac{\pi}{2}$ pulse is thus needed on one of the spin to detect such a state. By repeating the experiment for various value of τ, one obtain a decay curve. Once the values of T_2 are measured using single coherence decay experiments, it is possible to deduce the value of c_{12} .

Three qubit quantum error correction code

The three qubit quantum error correction code [45] can protect one qubit of information $|\psi\rangle$ against single qubit errors about one Pauli axis. The quantum circuit for this code can be found in Fig. D.8. If errors happen during the noise period, it can be shown that this code corrects any single qubit phase error, i.e. errors of the form $\hat{Z}_1, \hat{Z}_2$ or $\hat{Z}_3$, but fails at correcting multiple phase errors, i.e. $\hat{Z}_1\hat{Z}_2, \hat{Z}_1\hat{Z}_3, \hat{Z}_2\hat{Z}_3$ and $\hat{Z}_1\hat{Z}_2\hat{Z}_3$.

As seen above, the natural noise present in NMR consists of a phase flip. A valid measure to quantify the effect of the noise on the system is to consider the fidelity of entanglement F_E [163], which corresponds to averaging the state-correlation for the density matrix states $\hat{X}$, $\hat{Y}$ and $\hat{Z}$. In other words, the state-correlation $f_{\hat{A}}$ for an initial state $\hat{A}$ consist on the amount of polarization in the output relative to the input. The fidelity of entanglement is then given by

$$F_E = \frac{1}{4}(1 + f_{\hat{X}} + f_{\hat{Y}} + f_{\hat{Z}}).$$

If the decoherence is caused solely by the transversal relaxation, the fidelity of entanglement over time of such a protocol is given by

$$\begin{aligned} F_E = \frac{1}{4}\Big[&2 + e^{-\gamma_2^{(1)}t} + e^{-\gamma_2^{(2)}t} + e^{-\gamma_2^{(3)}t} \\ &-e^{-\gamma_2^{(1)}t-\gamma_2^{(2)}t-\gamma_2^{(3)}t}\cosh(2c_{12}t\sqrt{\gamma_2^{(1)}\gamma_2^{(2)}})\Big], \end{aligned}$$

where it has been assumed that the noise affecting qubit 3 was uncorrelated with the other qubits (because qubit 3 is represented by a separate nuclear species in our experiment). The correlation factor can be extracted from the deviation of the fidelity from unity, due to the failure of the code.

As demonstrated in Sec. D.2.4, it is possible to engineer the correlated noise on two spins using six different experiments. If we want to engineer the noise for a third uncorrelated qubit, it is done with twelve experiments, using the union of two sets of Kraus operators/coefficients given by

$$\{\hat{U}'_k, p'_k\} = \{\hat{U}_i^{1,2}, (1-q)p_i^{1,2}\}\bigcup\{\hat{U}_j^{1,2}\hat{Z}_3, qp_j^{1,2}\}$$

for $k = 1 \dots 12$ and $i, j = 1 \dots 6$ and where $q = \frac{1}{2}(1 - e^{\gamma_2^{(3)}t})$ corresponds to the probability of the uncorrelated qubit to undergo a phase flip and $\hat{U}_i^{1,2}$ and $p_i^{1,2}$ are the correlated noise Kraus operators/coefficients given in Table D.5.

Therefore, we can implement the QECC using those twelve noise operators and obtain the fidelity decay for various value of c_{12} and t .

The experiment

The theory laid down in the previous subsection assumed that the system is composed of two noise correlated qubits and one uncorrelated qubit. As seen in subsection D.2.3, such a system can be found in a molecule containing two spins of the same species with different Larmor frequency and one of a different kind. For this experiment, we have chosen the ^{13}C-labeled acetyl chloride dissolved in deuterated chloroform and used a 700 MHz Bruker

Avance NMR spectrometer with dual inverse cryoprobe. The structure, chemical shifts and J-coupling strengths of the molecule are given in Fig. D.9. For this molecule, the assumption of weak coupling used throughout subsection D.2.3 is fulfilled due to the large chemical shift difference between the two carbons.

The T_2's for each nuclei have been determined using a series of spin echo experiments for various delays and their values are given in Table D.6. To implement the quantum error correction code on this molecule, the circuit in Fig. D.8 was first converted into gates implementable in NMR, which consist of single qubit rotations about any axis in the xy-plane or around the z-axis, and J-coupling evolutions. A J-coupling of length $\tau = \frac{1}{2J}$ is locally equivalent to a C-NOT. Moreover, the z-rotations can be done instantaneously by changing the phase of subsequent pulses. This ideal NMR pulse sequence was then fed into a homemade compiler which estimates the first order phase and coupling errors during the pulses and then tracks the phase of the subsequent pulses and optimize the refocusing scheme and J-coupling delays to minimize overall coupling errors [101]. Spatial averaging [101] was used to initialized the states $|00\rangle\langle 00|00_{12} \otimes X_3$, $|00\rangle\langle 00|00_{12} \otimes Y_3$ and $|00\rangle\langle 00|00_{12} \otimes Z_3$, from the thermal state of a liquid state NMR system.

The quantum error correction code was first implemented using engineered errors in twelve experiments. The purpose of analyzing engineered noise is to be able to generate different fidelity of entanglement decay curves corresponding to different value of correlation factors. It is done by adding each experiments weighted by the corresponding coefficient given in Table D.5. Once a fidelity decay curve is obtained for the natural noise, it is possible to extract the correlation factor by comparing to which engineered noise fidelity decay curve the natural noise curve correspond to.

The natural noise fidelity decay curve was obtained by implementing the identity map during the noise section of Fig. D.8. To perform this implementation the spins could not be simply decoupled from one another using multiple π pulses (e.g. the Hadamard refocusing scheme [91]). Under such a refocusing scheme for a time t, the double coherence terms in a density matrix spend as much time in zero quantum coherence as in double quantum

Nucleus	T_1	T_2
M	4.0 ± 0.1	1.2 ± 0.1
C_1	7.9 ± 0.4	2.1 ± 0.1
C_2	15.2 ± 0.8	0.24 ± 0.03

Table D.6: T_1 and T_2 values for the acetyl chloride. It can be seen that for the maximal duration of the experiment ($\sim 300\,ms$), the effect caused by T_1 relaxation can be neglected.

coherence. Therefore, from Eq. D.1, the correlation factor term in the exponential decay cancel and do not affect the decay of the double coherence term.

If we let the natural noise act on the system for a period $\tau = \frac{n}{J_{C_1C_2}}$, $n \in \mathbb{N}$, the overall evolution is an identity and the terms of the density matrix containing a double coherence for the two carbons have remained in double coherence during the entire delay. The field inhomogeneities can be refocused by applying simultaneous π pulses on the carbons which leaves the J-coupling evolution untouched.

The fidelities of our experiments have been extracted by fitting every peak of the NMR spectra using Lorentzian shape curves. The resulting values can be seen in Fig. D.10, where the curves for the fidelity of entanglement for engineered noise are shown for correlation factors of 0, 0.5 and 1. Ten values of the fidelity decay obtained by applying natural noise are shown, from which we extracted a correlation factor of $c_{12} = 0.5 \pm 0.2$.

This experiment needs a high degree of precision, since within the time interval used to implement natural noise (from 0 to 320 ms), the maximum difference between the $c_{12} = 0$ and $c_{12} = 1$ curves is 3%. Integrating the square of the noise of a spectrum over a region corresponding to the width of a signal peak estimates the signal to noise ratio to be of the order of 1%. Therefore, the fluctuation of the measured fidelities due to the noise explains the large uncertainty on the measured correlation factor.

Using the usual NMR technique of double coherence decay, the noise correlation factor between C_1 and C_2 was determinded to be $c_{12} = 0.3 \pm 0.2$. The interval using QECC agrees with the value obtained using the traditional double-quantum coherence decay technique, to within experimental error.

Discussion

By comparing the above two techniques to extract the noise correlation factor between two spins, one could argue that the QECC technique is much more involved then the standard NMR technique, while yielding to the same conclusion. The goal of the present experiment was to demonstrate that the use of QEC to probe the noise present in a system is feasible and that the control necessary to get the error information is achievable. From there, it is possible to generalize this technique to any physical system with more complex noise model. If the noise contains not only phase errors, but also bit errors and/or a combination of the two, the same technique could be applied using more complex QECC, such as the five qubit code [115].

Other technical advantages arise from the signal detection. Using the NMR technique, there is an doubly exponential decay in the signal amplitude for a double coherence decay (see Eq. D.1). From the nature of the QECC technique, the signal decays slower, thus allowing better statistics and analysis. If the system under analysis contains three spins of the same type, there would be a possibility of three different correlation factors c_{12}, c_{23} and c_{13}. Using standard NMR technique, three different experiments with different initial states would be needed to extract those three values. Using the QECC technique, only the noise portion of the pulse sequence would need adjustment by changing the refocusing scheme in order to refocus the unwanted correlation,e.g. a π pulse on qubit 1 would cancel the correlation c_{12} and c_{13} for the reasons explained earlier.

Finally, this technique could be used to validate our assumption that the noise is effectively gaussian. In the case where the system contained three noise correlated spins, our gaussian assumption ensures that the noise is only pairwise correlated, i.e. $c_{123} = 0$. If it is so, a triple coherence decay curve should be described using only the T_2 values and the pairwise correlation factors. This curve can be obtain using the NMR technique of triple coherence decay, but would yield a curve that decays triply exponentially. On the other hand, a curve affected by a triple correlation factor could be obtained by the same QECC pulse sequence by letting all the noise correlations act during the noise part of the pulse sequence. In the case where that curve would not be described properly using only the pairwise correlations, it would be an indication of the failure of the noise model.

D.2.6 Conclusion

In this work, we have demonstrated that QEC can be used to probe a physical system and extract partial, but important information about the noise model without having to perform full quantum process tomography. The technique was implemented successfully in a liquid-state NMR quantum information processor, but is applicable to any QIP device in which standard quantum error correction can be carried out.

D.2.7 Acknowledgments

This work has been supported by NSERC, MITACS, the National Security Agency (NSA) under Army Research Office (ARO) contract number W911NF-05-1-0469 and CIAR.

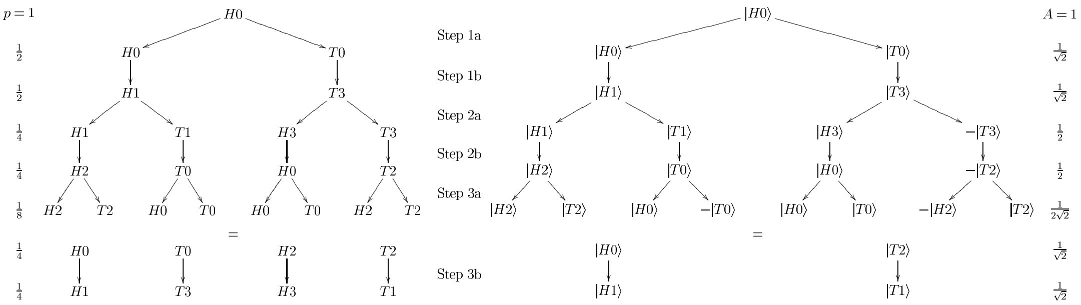


Figure D.1: Comparison of the dynamics of the classical (left) and quantum (right) random walk on a square for three steps. H or T represents the state of the coin and the number, the position of the walker at one of the four nodes of the square. p is the probability of each classical state and A is the probability amplitude for the quantum state. Part (a) for each step is the coin flip and part (b) the movement around the square. In both cases the walker starts at node 0 with the coin in the heads state. After one step he has a fifty percent probability of being at either of nodes 1 or 3. Then, in the second step he goes to either 0 or 2 with fifty percent probability. In the third step however, the two types of walk diverge. The classical walk continues to oscillate and the probability remains spread out. In the quantum walk on the other hand, the probabilities interfere and cancel out leaving all the probability in one corner after three steps.

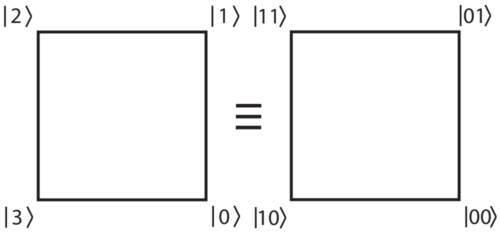
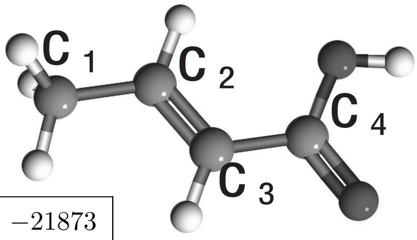


Figure D.2: Logical labeling of the nodes on which we implemented the DTQW. With this labeling, flipping the first qubit corresponds to a horizontal move and flipping the second qubit, a vertical move.



C_1	-2546			
C_2	40.5	-21873		
C_3	1.5	69.9	-18498	
C_4	7.1	1.4	72.4	-25322
	C_1	C_2	C_3	C_4

Figure D.3: Molecular structure of trans-crotonic acid and its Hamiltonian parameters. The chemical shifts are given as the diagonal elements and the coupling strength (Hz) by the off-diagonal elements. Note that since the darkly shaded unlabeled nuclei are oxygen whose natural abundance of ^{16}O with 0 spin is close to 100 %. Therefore, the two oxygen nuclei do not couple with the rest of the molecule and can be ignored. Lightly shaded unlabeled nuclei are hydrogen which were decoupled during the experiment.

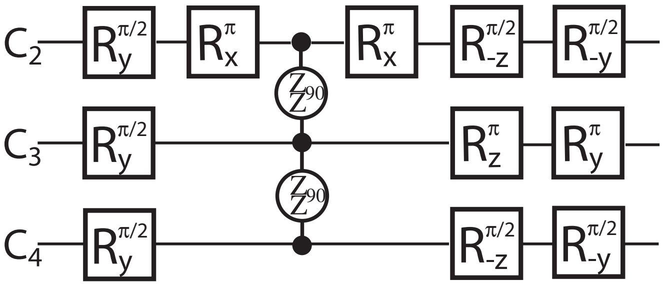
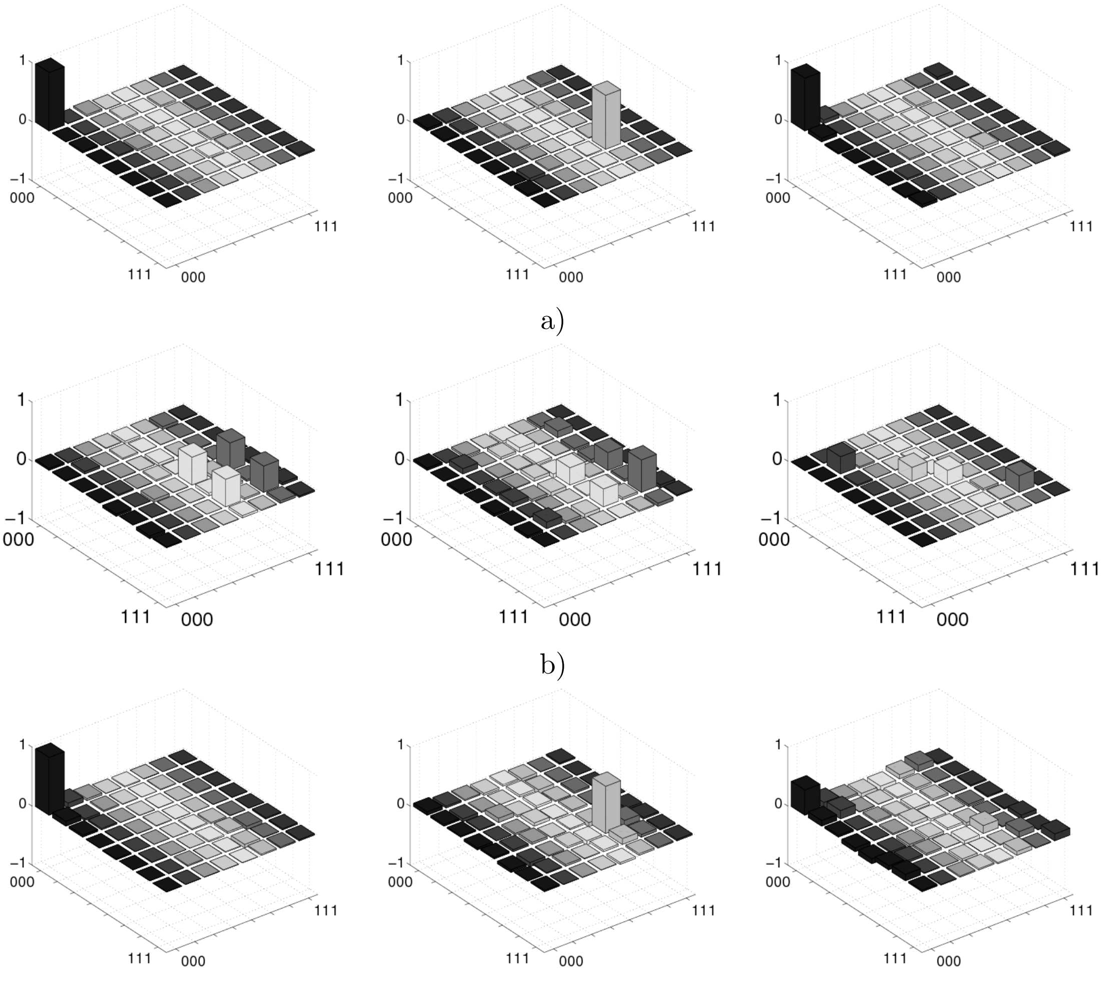


Figure D.4: NMR pulse sequence representing one step of DTQW. The notation R_i^θ means a rotation of an angle θ around the axis i. Refocussing pulses are not shown. Since each nucleus is tracked in its own rotating frame, rotations about the z-axis are implemented instantaneously through a change of reference frame.

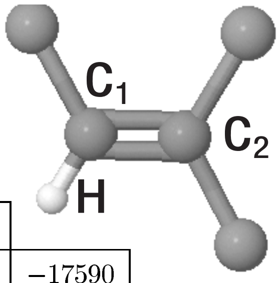


a)

b)

c)

Figure D.5: Real part of the reconstructed density matrix for a) Step 0, 4 and 8 of the quantum walk on the crotonic acid. b) Step 3 with no, partial and full decoherence on the crotonic acid and c) Step 0, 4 and 8 on the TCE.



H	−3891		
C_1	201	−17590	
C_2	8.9	103	−18675
	H	C_1	C_2

Figure D.6: Diagram of ^{13}C labeled TCE. The chemical shifts and couplings are given in the table. Note that since the chlorine nuclei (unlabeled) have a spin of $\frac{3}{2}$, they have an electric quadrupole moment which causes them to decohere quickly and they have a very small coupling to the rest of the molecule which we can ignore in the natural Hamiltonian of the molecule.

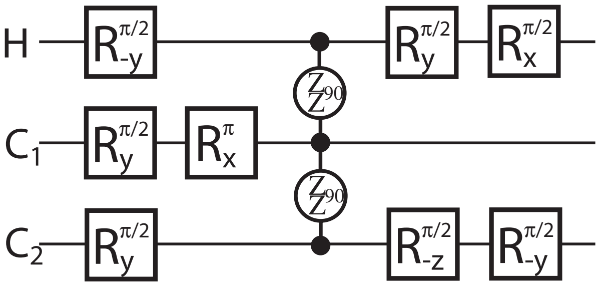


Figure D.7: Circuit used to implement one step of the DTQW on the TCE. The z-rotation on C_2 occurs naturally during the coupling with C_1. Note also that the refocusing pulses are not shown in this pulse sequence.

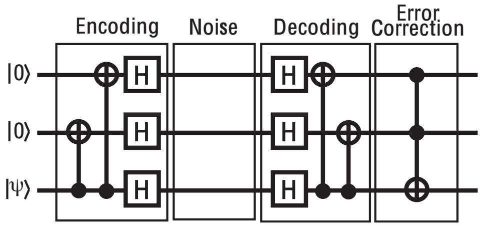
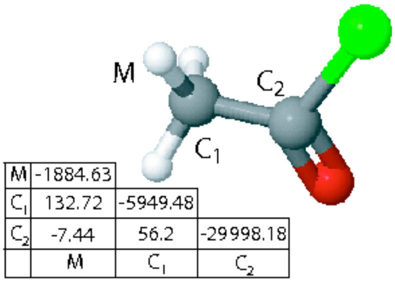


Figure D.8: The quantum circuit of the three qubit quantum error correction code. The two qubit gates are C-NOTs and the three qubit gate is a Toffoli gate.



M	-1884.63		
C_1	132.72	-5949.48	
C_2	-7.44	56.2	-29998.18
	M	C_1	C_2

Figure D.9: The ^{13}C-labeled acetyl chloride molecule. The diagonal elements of the table gives the chemical shift (difference in Larmor frequency) for each nucleus with respect to a reference frequency (700.13 MHz for the hydrogens and 176.05 MHz for the carbons). The three hydrogens forming the methyl group are indistinguishable and form a spin-$\frac{1}{2}$ and $\frac{3}{2}$ subspace. The spin-$\frac{1}{2}$ subspace was selected using a three-step pulse sequence and "crusher" field gradients [44].

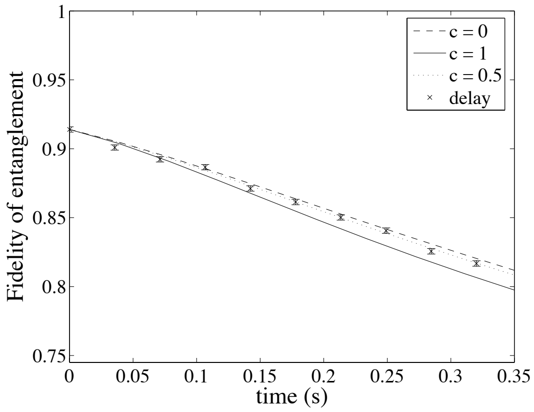


Figure D.10: Experimental results. The lines correspond to the fidelity decay for noise correlation factors of 0, 0.5 and 1 as a function of time simulated from the measured T_2's and the experimental fidelities obtained by implementing engineered noise. The points are the fidelities when the system is affected by natural noise for a various amount of time.

Appendix E

Another Result Related to Quantum Reference Frame

The following is an unpublished paper about covariant maps under spacial rotation.

E.1 Quantum Reference Frames and the Classification of Rotationally-Invariant Maps

J.-C. Boileau[1,2], L. Sheridan[2], M. Martin[2], and S. D. Bartlett[3]

1. Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada.
2. Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON, N2L 2Y5, Canada.
3. School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia

Abstract: We give a convenient representation for any map which is covariant with respect to an irreducible representation of $SU(2)$, and use this representation to analyze the evolution of a quantum directional reference frame when it is exploited as a resource for performing quantum operations. We introduce the *moments* of a quantum reference frame, which serve as a complete description of its properties as a frame, and investigate how many times a quantum directional reference frame represented by a spin-j system can be used to perform a certain quantum operation with a given probability of success. We provide a considerable generalization of previous results on degradation of reference frame, from which follows a classification of the dynamics of spin-j system under the repeated action of any covariant map with respect to $SU(2)$.

E.1.1 Introduction

Quantum operations, such as the application of a unitary rotation on a qubit or a projective measurement in some basis, require some form of reference frame. In most descriptions of quantum experiments, this reference frame is considered to be classical and therefore can be treated as non-dynamical and used repeatedly without disturbance. The situation is quite different for a quantum reference frame – measurements and interactions can cause an unknown disturbance on the quantum frame and, consequently, it can reduce the ability of that system to serve as a reference frame in subsequent uses [68, 8, 13, 143, 12].

Consider, for example, an apparatus implementing a unitary rotation operation on a qubit. For concreteness, we consider this qubit to be spin-1/2 system; however, it could equally well be the polarization state of a single photon, a two-level atom, etc. A standard apparatus uses a classical directional reference frame to define the axis about which the rotation occurs. Now suppose that due to miniaturization, space constraints, or to improve the speed of the operation, the system that serves as the reference direction is not accurately described classically, but instead requires a quantum treatment using a Hilbert space of finite size. The quantum operation describing the unitary rotation is then an operation *conditional* on the state of the quantum reference frame. Two effects will occur as a result of the quantum nature of the reference frame. First, this conditional operation will not be identical to the classical case due to the inherent uncertainty in the direction of the quantum reference frame. Second, upon use, the state of the quantum reference frame may experience an unknown disturbance, essentially due to entanglement induced between the reference frame and system as a result of the interaction.

We now formally define the scenario that we are considering. Let $\rho^{(0)}$ be the initial state of the quantum reference frame. Let $\sigma^{\otimes n}$ be the state of a reservoir composed of n ordered subsystems on which we will sequentially perform operations using the same quantum reference frame. Using the first subsystem, we apply a quantum operation χ on the joint system $\rho^{(0)} \otimes \sigma_1$. We require that the map χ is *rotationally invariant*, meaning that it is independent of any specific direction in space. Formally, a map χ is rotationally invariant if and only if $\chi[R(\Omega)(\cdot)R(\Omega)^\dagger] = R(\Omega)\chi(\cdot)R(\Omega)^\dagger$ for any rotation $\Omega \in SO(3)$ (or more generally $SU(2)$ if half-integral spins are considered) of the joint system. (Here, R is the unitary representation of the rotation group $SO(3)$ on the joint system.) Subsequent to the operation, the reduced state of quantum reference frame is $\rho^{(1)} = \mathrm{Tr}_{r_1}[\chi(\rho^{(0)} \otimes \sigma_1)]$, where Tr_{r_i} denotes the partial trace over the i^{th} subsystem of the reservoir. The first subsystem of the reservoir is then discarded, and the same operation is performed on the second subsystem, but using the updated quantum reference frame $\rho^{(1)}$. We repeat these steps with the following subsystems of the reservoir, always using the updated quantum reference frame. The state of the quantum reference frame after the i^{th} step is recursively $\rho^{(i+1)} = \mathrm{Tr}_{r_i}[\chi(\rho^{(i)} \otimes \sigma_i)]$. By discarding the previous subsystems of the reservoir at every step, we assume that the reservoir cannot be used to increase our knowledge about the quantum reference frame.

We wish to define a measure of the quality of the quantum reference frame, and to study how it decreases with the number of times the reference frame is used. Defining the quality of a quantum reference can be quite arbitrary and depends on the interests of the experimentalist and on the task at hand. In this paper, we use the term *quality function* for any function F that is meant to quantify the ability of the reference frame to perform a particular task. We will set general conditions on such functions, and analyze how such functions evolve with the number of uses of the quantum reference frame. We will also analyze the *longevity* of the quantum reference frame, that is, the number of times it can be used to perform a certain task before its quality falls below a certain threshold.

As in [13, 143], a quantum directional reference frame could be used to measure whether a series of spin-$\frac{1}{2}$ particles are parallel or anti-parallel to some classical arrow. (See Figure E.1.) However, there are a myriad of other possible cases, including measuring the angular momentum of a series of spin-j' systems with $j' > \frac{1}{2}$, or using the quantum reference spin to indicate an axis about which a unitary rotation operation is performed.

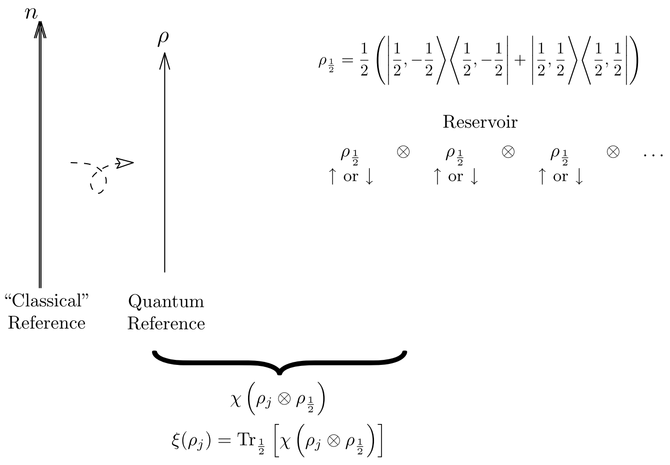

n
ρ
$\rho_{\frac{1}{2}} = \frac{1}{2}\left(\left|\frac{1}{2},-\frac{1}{2}\right\rangle\left\langle\frac{1}{2},-\frac{1}{2}\right| + \left|\frac{1}{2},\frac{1}{2}\right\rangle\left\langle\frac{1}{2},\frac{1}{2}\right|\right)$
Reservoir
$\rho_{\frac{1}{2}}$ ⊗ $\rho_{\frac{1}{2}}$ ⊗ $\rho_{\frac{1}{2}}$ ⊗ …
↑ or ↓ ↑ or ↓ ↑ or ↓
"Classical" Reference
Quantum Reference
$\chi\left(\rho_j \otimes \rho_{\frac{1}{2}}\right)$
$\xi(\rho_j) = \mathrm{Tr}_{\frac{1}{2}}\left[\chi\left(\rho_j \otimes \rho_{\frac{1}{2}}\right)\right]$

Figure E.1: The scenario studied by Bartlett *et al.* [13]. The quantum reference spin is used to measure the direction of a series of spin-$\frac{1}{2}$ particles in the completely mixed state by means of a projection onto the $J - \frac{1}{2}$ or $J + \frac{1}{2}$ subspaces.

In this paper, we generalize the results of Refs. [13, 143] in three different ways. First, we consider performing operations on a reservoir of *qudits* (quantum systems with d-dimensional Hilbert spaces) instead of restricting to the case of qubits ($d = 2$) as in [13, 143]. Second, we consider all possible rotationally-invariant quantum operations on the reference frame and the system, while the analyses in [13, 143] were restricted to very particular kind of interaction (specifically, interactions describing measurements). Note that this joint quantum operation could be a measurement, a conditional unitary, or any other completely-positive map, provided that it is rotationally invariant. Third, we generalize the concept of the quality function, and place general conditions on its form and evolution. In [13, 143], only a particular kind of quality function was considered – one which was based on the average probability of a correct measurement given a known spin direction – while now we are interested in any possible quality function.

The paper is organized as follows. In Section E.1.2, we introduce a set of parameters for describing a quantum reference frame, which we call *moments*, and demonstrate that any quality function must depend only on these moments. We then present the key result, Theorem 15, which provides a classification of all rotationally-invariant maps. We give some recursive equations describing the evolution of the moments (Theorem 16). In Section E.1.3, we discuss the longevity of a quantum reference frame for a general quality function under the repetitive application of a rotationally-invariant map, by analyzing the dependance of the evolution of its moments. Section E.1.4 is dedicated to a pair of examples intended to illustrate the power of the techniques developed in the earlier sections of the paper. We explore the specific cases of measuring spin-1 particles relative to a quantum reference frame, as well as the use of a reference frame to perform Pauli operations on qubits. We conclude with some general comments in Section E.1.5.

E.1.2 Moments and Fidelity Functions

A classical reference frame is a local convention for describing the state of a system. A quantum reference frame is a quantum system that carries information about a classical reference frame. We define a quantum directional reference system as one which is correlated with a classical direction in space represented by the vector $\hat{n}$. The quantum system that indicates the direction of our reference axis is taken to be a spin-j system, which has dimension $d = 2j + 1$. As in [143], we consider an arbitrary initial state $\rho_j^{(0)}(\hat{n})$ for the spin-j quantum directional reference frame. (That is, the quantum reference frame transforms under rotations according to the spin-j representation.) If we additionally assume that the initial quantum reference frame depends *only* on the vector $\hat{n}$, we deduce a symmetry condition: if $R_j(\Omega)$ is the spin-j unitary representation of the rotation $\Omega \in SU(2)$ that transforms $\hat{n}$ to $\hat{n}'$, then $R_j(\Omega)\rho_j^{(0)}(\hat{n})R_j(\Omega)^{-1} = \rho_j^{(0)}(\hat{n}')$. That is, the set of possible initial states $\{\rho_j^{(0)}(\hat{n})\}$ is in one-to-one correspondence with the set of directions $\hat{n}$.

Because a vector $\hat{n}$ has an invariance group – the group of rotations about this axis – the symmetry condition implies that the state of the quantum directional reference frame is also invariant under rotations about the $\hat{n}$-axis. This in turn ensures that $\rho_j^{(0)}(\hat{n})$ commutes with the angular momentum operator $J_{\hat{n}}$ in the $\hat{n}$-direction, and thus $\rho_j^{(0)}(\hat{n})$ is diagonal in the basis of eigenstates of $J_{\hat{n}}$.

We now consider how the state of the quantum reference frame is updated through its use in repeated rotationally-invariant operations χ which act on both the frame and the reservoir. We assume as in [13] that the reservoir initiates in a state that is invariant under rotations. (The completely mixed state is one example of such a state, but there exist other non-trivial states).[1] The assumptions that the initial state of the reservoir and the joint quantum operation are spatially invariant imply that the map ξ used to update the state of the quantum directional reference frame, $\rho_j^{(i+1)} = \xi(\rho_j^{(i)}) = \mathrm{Tr}_{r_i}[\chi(\rho_j^{(i)} \otimes \sigma_i)]$, is also rotationally invariant. We denote the restricted map ξ on the quantum directional reference the *disturbance* map. It describes the back-action on the reference which occurs as a result of the interaction.

Let $R_{\hat{n}}(\theta)$ be the rotation by an angle θ around the vector $\hat{n}$. Suppose that a rotationally-invariant map ξ is applied to density matrix ρ_j that is diagonal in a basis given by the eigenvectors of $J_{\hat{n}}$. Because

$$R_{\hat{n}}(\theta)\xi(\rho_j)R_{\hat{n}}^{-1}(\theta) = \xi(R_{\hat{n}}(\theta)\rho_j R_{\hat{n}}^{-1}(\theta)) = \xi(\rho_j), \tag{E.1}$$

then $\xi(\rho_j)$ is also diagonal in the basis given by the eigenvectors of $J_{\hat{n}}$. As a result, the evolution of the quantum reference frame under the repeated application of the rotationally-invariant map ξ can be described by $2j+1$ equations, one for each of the eigenvalues of ρ_j. A different set of parameters that are equivalent to the eigenvalues is the set of *moments* given by

$$\{\mathrm{Tr}[\rho_j J_{\hat{n}}^{\ell}] \mid 1 \leqslant \ell \leqslant 2j\}. \tag{E.2}$$

Note that only $2j$ moments are necessary because the sum of the eigenvalues must be one. The use of moments instead of eigenvalues will greatly simplify the analysis of the evolution of the quantum reference frame.

The main motivation for studying the moments of the quantum reference frame is that any quality function F will depends only on these moments. Consequently, the behavior of the different moments will determine the behavior of the different quality functions. To see why F depends only on the moments given by equation E.2, we first remark that any reasonable form of F depends only on the state of the quantum reference, but that it must respect the relation $F(\rho_j) = F(R_j(\Omega)\rho_j R_j(\Omega)^{-1})$ for all rotations $\Omega \in SU(2)$ and state ρ_j diagonal in the basis composed of the eigenvectors of $J_{\hat{n}}$. In other words, the quality measure should not be biased such that it favors a quantum reference frame that is pointed in any particular direction relative to some external frame. All directions must be equally valid. Therefore, F does not depend on the direction of $\hat{n}$, but only on the eigenvalues or

[1] Note that, in [143], they considered the case where the qubits could also be partially or fully polarized, implying that there was some initial correlation between the systems to be measured and the quantum reference frame. Generalizing their results to maps χ that do not preserve angular momentum and to reservoirs that are not polarized in the direction of the quantum reference or only composed of spin-$\frac{1}{2}$ particles is challenging problem. We hope that the framework we provide here will be useful in extending our results to this general situation.

the moments of ρ_j. Note that the set of moments with respect to the direction $\hat{n}$ can be written as a function of the moments with respect to any other direction — this is simply a change of basis.

Recall that $\rho_j^{(k)}$ as the state of the quantum reference frame after the k^{th} joint operation with a subsystem of the reservoir, and $\rho_j^{(0)}$ as the initial state of the quantum reference frame. By our previous arguments, for a given state $\sigma^{\otimes n}$ of the reservoir and joint quantum operation χ, the moments of $\rho_j^{(k)}$ must be explicit linear functions of the moments of $\rho_j^{(k-1)}$. These results lead us to the general recursion relation,

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=0}^{2j} A_i^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{i}], \tag{E.3}$$

where $A_i^{(\ell)}$ are real coefficients.

We now present a theorem which provides a classification of the different rotationally-invariant maps, and constitutes the primary result of this paper. First, some notation. Let J_k for $k = x, y, z$ be the angular momentum operators for some arbitrary Cartesian frame. On a spin-j system, define the map

$$\zeta(\rho_j) = \frac{1}{j(j+1)} \sum_{k\in\{x,y,z\}} J_k \rho_j J_k, \tag{E.4}$$

for ρ_j a density matrix of the spin-j system. Let $\zeta^{\circ n}$ denote the n-fold composition of ζ for $n > 0$, and define $\zeta^{\circ 0}(\rho_j) = \rho_j$.

Theorem 15 *Any map ξ which is invariant with respect to a spin-j irreducible representation of $SU(2)$ has the form*

$$\xi(\rho_j) = \sum_{n=0}^{2j} q_n \zeta^{\circ n}(\rho_j), \tag{E.5}$$

where $\{q_n\}$ are real coefficients.

The proof is provided in Appendix E.1.7.

Theorem 15 allows us an immediate simplification by restricting the number of coefficients $A_i^{(\ell)}$ of equation (E.3) that are required to calculate the evolution of the reference state. The following theorem limits the number of coefficients $(A_i^{(\ell)})$ which determine the evolution of the $J_{\hat{n}}$-moments of a spin-j under the action of a general invariant channel, and can be very useful to study the behavior of moments with low degree:

Theorem 16 *If ℓ is even, then*

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=0}^{\ell/2} A_{2i}^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{2i}] \tag{E.6}$$

and if ℓ is odd, then

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=1}^{(\ell+1)/2} A_{2i-1}^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{2k-1}]. \tag{E.7}$$

Proof. We first define the z-axis so that it is parallel to the vector $\hat{n}$ which describes our classical directional reference frame. Consider any rotationally-invariant map ξ. By Theorem 15, we can write $\xi(\rho_j) = \sum_{k=0}^{2j} q_k \zeta^{\circ k}(\rho_j)$ where the coefficients q_k are real numbers and ζ is given by equation (E.4). Therefore,

$$\mathrm{Tr}[\xi(\rho_j) J_z^{\ell}] = \mathrm{Tr}[\rho_j \xi(J_z^{\ell})] = \sum_{n=0}^{2j} q_n \mathrm{Tr}[\rho_j \zeta^{\circ n}(J_z^{\ell})]. \tag{E.8}$$

To prove our theorem, it is sufficient to show that $\zeta^{\circ n}(J_z^{\ell})$ is a polynomial in J_z of degree ℓ, but where all the powers have the same parity as ℓ. Define $\lambda = j(j+1)$ and define the function

$$G(\ell) \equiv \frac{i}{\lambda}(J_x J_z^{\ell-1} J_y - J_y J_z^{\ell-1} J_x). \tag{E.9}$$

Using the commutation relations $[J_a, J_b] = i\epsilon_{a,b,c} J_c$, we can show that

$$\zeta(J_z^{\ell}) = \zeta(J_z^{\ell-1}) J_z + G(\ell), \tag{E.10}$$

and

$$G(\ell) = G(\ell-1) J_z + \zeta(J_z^{\ell-2}) - \frac{J_z^{\ell}}{\lambda}. \tag{E.11}$$

By induction, it is easy to prove using those relations that $\zeta(J_z^{\ell})$ is a polynomial in J_z of degree ℓ where all the powers have the same parity as ℓ. Using the fact that if $\zeta^{\circ n}(J_z^{\ell}) = \sum_{s=0}^{\ell} b_s J_z^s$ for some complex b_s, then $\zeta^{\circ n+1}(J_z^{\ell}) = \zeta(\sum_s b_s J_z^s) = \sum_s b_s \zeta(J_z^s)$ and we can prove by induction that $\zeta^{\circ n}(J_z^{\ell})$ is a polynomial in J_z of degree ℓ where all the powers have same the parity as ℓ. This concludes our proof. ■

E.1.3 Longevity of a Quantum Reference Frame

Suppose we have a microscopic device for performing an operation or measurement on a quantum spin using a quantum reference frame. Ultimately, after many uses, the quantum reference frame will need to be reinitialized. However, we wish to make as many uses of it as possible before performing this reinitialization, without allowing the accuracy to fall below some allowed threshold. To define this accuracy we will pick some quality function that is suited to our particular purpose. Recall that any quality function F will depend only on the moments of the $J_{\hat{n}}$ operator.

As in [13, 143], we are interested in the scaling, with respect to Hilbert space dimension, of how many times a quantum reference frame can be used before the value of its quality function falls below a certain threshold. We refer to this property as the *longevity* of a quantum reference frame. There are three important features of our analysis to highlight. First, whereas the work of [13, 143] restricts attention to one particular quality function, we will investigate the behavior of arbitrary functions satisfying the invariance relation described in Sec. E.1.2. Because we consider any such function, the longevity of the reference frame can be arbitrary. However, we are able to prove some general statements about the decay of the moments, and relationships amongst them, and these results can then be used to infer the behavior of any particular quality function. Second, rather than considering only a particular state of the reference frame (in [13, 143], they considered the state $\rho_j^{(0)} = |j,j\rangle_{\hat{n}}\langle j,j|$ because it was optimal for the task at hand), we consider an arbitrary state. As such, the initial state $\rho_j^{(0)}$ of the quantum reference frame, as specified by its moments, can depend on j in a quite arbitrary way. Finally, we consider arbitrary rotationally-invariant joint quantum operations χ. In [13, 143], χ was chosen to describe a particular measurement that was optimal for some task. The resulting disturbance map that they considered is

$$\xi(\rho_j) = (\frac{1}{2} + \frac{1}{2(2j+1)^2})\rho_j + \frac{2j(j+1)}{(2j+1)^2}\zeta(\rho_j)\,, \tag{E.12}$$

where $\zeta(\rho_j)$ is given by equation (E.4). Such covariant map was analyzed in [155]. In [13], it was shown that the number of times a reference frame can be used before its quality function (which in their case depends only of the first moment) falls below a certain threshold value scales by a factor j^2.

In the following, we prove a general theorem about the longevity of quantum reference frames. We start with Theorem 15, which states that any rotationally-invariant disturbance map can be written as $\xi = \sum_{n=0}^{2j} q_n \zeta^{\circ n}$. However, as we now show, if the range and values of the coefficients q_n can depend on j arbitrarily, it is impossible to make any general statements about the longevity. We identify some natural assumptions for the coefficients, which then allow us to present a general theorem.

First, in general, the number of parameters q_n describing a map increases with j; i.e., there are $2j+1$ parameters which can be significant to the evolution. As an example, consider the complete depolarization map $\chi(\rho_j \otimes \sigma) = \frac{1}{2j+1} I_{2j+1} \otimes \frac{1}{2} I_{1/2}$. It is straightforward to show that, for this map, all parameters q_n for $0 \le n \le 2j$ are nonzero. However, inspired by an argument in [143], consider a rotationally-invariant map $\chi(\rho_j \otimes \sigma)$ that conserves angular momentum in the $\hat{n}$ direction, meaning that $\chi(J_{\hat{n}}) = J_{\hat{n}}$ where $J_{\hat{n}}$ is the *total* angular momentum operator in the $\hat{n}$ direction. The change of angular momentum of the quantum reference frame caused by such a disturbance map ξ cannot be higher than the change in angular momentum of the subsystem σ, which is in turn bounded by the subsystem's dimension d. Given that the map $\zeta(\cdot) = \frac{1}{\lambda}\sum_{k=x,y,z} J_k(\cdot)J_k$ cannot lower or increase the angular momentum in any direction by more then one unit, the disturbance map's coefficients therefore satisfy $q_n = 0$ for all $n \geqslant d$, where d is the dimension of each subsystem σ of the reservoir. Thus, if a rotationally-invariant map conserves angular momentum, then we can bound the number of non-zero coefficients q_n in a way that is independent of j.

Second, if the parameters q_n can depend on j arbitrarily, then we would not be able to conclude anything about the longevity of the quantum reference frame. To understand this, consider the example

$$\chi(\rho_j \otimes \sigma) = \alpha_j \rho_j \otimes \sigma + (1-\alpha_j) \sum_{k=|j-k|}^{j+k} \Pi_{j+k}(\rho_j \otimes \sigma)\Pi_{j+k}, \tag{E.13}$$

where Π_{j+k} is a projection into the subspace of total angular momentum $j+k$. This map conserves the total angular momentum of the joint system. The dependence of α_j on j will determine directly the rate of the decay of the moments in function of j. Because the dependence of α_j can be arbitrary, then the rate of the decay of the moments can also be an arbitrary function of j. We conclude that, in order to make a statement about the longevity of the quantum reference frame, we need to assume more than a bound on the number of the q_n parameters or that the rotationally-invariant map χ conserves angular momentum in the direction of $\hat{n}$. For the following theorem, we assume that each coefficient q_n converges to a constant when $j \to \infty$. We also need to make an assumption about the rate of convergence: suppose that each q_n can be written as a quotient of two polynomial in j, such that the degree of the denominator is at least the degree of the numerator (i.e., $q_n \leqslant O(1)$). Finally, we assume that the state $\rho_j^{(0)}$ has the propriety that $\mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}^{\ell}] = O(j^{\ell})$. A sufficient condition for this to be true is that there exists a $\beta > 0$ independent of j such that ${}_{\hat{n}}\langle j,j|\rho_j^{(0)}|j,j\rangle_{\hat{n}} > \beta$. (We note that these assumptions are motivated by examples, which we explore in the following section.)

Theorem 17 (Longevity) *Consider a quantum reference frame, realised as a spin-j system with initial state $\rho_j^{(0)}$, which is used for performing a rotationally-invariant joint operation χ. If this operation induces a disturbance map $\xi = \sum_{n=0}^{2j} q_n(j)\zeta^{\circ n}$ that satisfies the following assumptions:*

1. *there exists some n_{max} such that $q_n = 0$ for all $n \geqslant n_{max}$,*

2. $q_n \leqslant O(1)$,

3. $\mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}^{\ell}] = O(j^{\ell})$,

then the number of times that such a quantum reference frame can be used before its ℓ^{th} moment falls below a certain threshold value scales as j^2.

Proof. Noting that the J_k operators are self-adjoint and using the cyclic propriety of the trace, the map ζ has the property that

$$\mathrm{Tr}[\zeta(\rho)J_{\hat{n}}^{\ell}] = \mathrm{Tr}[\rho\,\zeta(J_{\hat{n}}^{\ell})]. \tag{E.14}$$

Therefore the moments of angular momentum in the $\hat{n}$ direction after the map ξ_j has been applied to the reference state $\rho_j^{(k)}$ can be expressed as

$$\mathrm{Tr}[\rho_j^{(k+1)}J_{\hat{n}}^{\ell}] = \sum_{n=0}^{2j} q_n \mathrm{Tr}[\rho_j^{(k)}\ \zeta^{\circ n}(J_{\hat{n}}^{\ell})] \tag{E.15}$$

Using the commutation relations, the factor $\zeta^{\circ n}(J_{\hat{n}}^{\ell})$ can be expanded as a polynomial in $J_{\hat{n}}$ of degree ℓ. Define $\lambda = j(j+1)$. Using a proof by induction on equation (E.10) and observing that $\sum_{k\in\{x,y,z\}} J_k J_{\hat{n}} J_k = (\lambda-1)J_{\hat{n}}$, it is easy to show that the coefficient of the leading term will be $A_{\ell}^{(\ell)} = \sum_{n=0}^{n_{max}}[q_n(1-O(\frac{1}{\lambda}))] = 1 - O\big(\frac{1}{\lambda}\big)$ where we used the normalization condition $\sum q_n = 1$, the assumption that each q_n is $O(1)$ and the fact that $q_n = 0$ for $n > 2j$. The coefficients of the non-vanishing lower terms will be $A_i^{(\ell)} = O(1)$ for $i < \ell$.

This reasoning about the constants $A_k^{(i)}$ is sufficient to characterize the rate of change of the moments with repeat application of the map. To demonstrate this, let $\ell = 1$. We want to find the minimum value of t such that $\mathrm{Tr}[\rho_j^{(t)}J_{\hat{n}}] < c$, for some constant c. Using equation (E.7), we need to solve

$$c = \mathrm{Tr}[\rho_j^{(0)}J_{\hat{n}}]\left(1 - O\left(\frac{1}{\lambda}\right)\right)^{t_c}. \tag{E.16}$$

Therefore,

$$t_c = O\left(j^2\right). \tag{E.17}$$

To generalize to higher moment, observe that $\mathrm{Tr}[\rho_j^{(0)}J_{\hat{n}}^{i}] \leqslant O(j^i)$ for $i < \ell$. Recall that we assumed that $\mathrm{Tr}[\rho_j^{(0)}J_{\hat{n}}^{\ell}] = O(j^{\ell})$. Using equation (E.7), the facts that $A_{\ell}^{(\ell)} = 1 - O\big(\frac{1}{\lambda}\big)$ and $A_i^{(\ell)} = O(1)$ for $i < \ell$, we can extend our result to higher odd moments by using strong induction. In other word, we can show that the minimum value of t such that $\mathrm{Tr}[\rho_j^{(t)}J_{\hat{n}}^{\ell}] < c$ for proper chosen initial state $\rho_j^{(0)}$ is $O(j^2)$. For even moments, a similar approach using equation (E.6) can be used to obtain an identical conclusion. ∎

Note that in the case were the assumptions of Theorem 17 fail, but there is a specific dependance on j of the map χ and the state $\rho_j^{(0)}$, equations (E.6) and (E.7) may still be useful tools to study the longevity on quantum reference frame.

E.1.4 Examples

Measuring Spin-1 Systems

As an example, consider the measurement of spin systems relative to a directional reference frame. This problem was investigated in [13, 143] for the case of spin-1/2 systems. Specifically, suppose that the reservoir consists of spin-s systems which are initially in the completely mixed state. The quantum reference direction is a spin-j system with $j > s$, aligned with some classical direction $\hat{n}$. The goal is to measure each spin to determine its component μ along the vector $\hat{n}$ by performing a joint measurement on both the system from the reservoir and the reference system. The optimal rotationally-invariant joint operation for this task [11] is a POVM given by the projectors $\{\Pi_{j+\mu}|\mu \in \{-s, \ldots, s\}\}$ where $\Pi_{j+\mu}$ corresponds to a projection onto the subspace where the total angular momentum of the reference spin with the measured spin is $j + \mu$. Define the fidelity F_s as the probability that, when the reference is used to measure a system in a *known* state $|s, \mu\rangle_{\hat{n}}$ by means of the above POVM on the joint system the result it returns is correct, assuming that all values of μ are equiprobable.

Consider briefly the case of $s = 1/2$, as in [13, 143]. The fidelity function can be expressed in terms of the first moment of the reference frame *after* being used k times, as

$$F^{(k)}_{\frac{1}{2}} = \mathrm{Tr}(\xi^{\circ k}(\rho^{(0)}_j)) = \mathrm{Tr}\left[\frac{1}{2}\sum_{\mu\in\{-\frac{1}{2},\frac{1}{2}\}} \Pi_{j+\mu}(\rho_j \otimes |\mu\rangle\langle\mu|)\right] = \frac{1}{2} + \frac{1}{2j+1}\mathrm{Tr}[\xi(\rho^{(k)}_j)J_z]\,. \tag{E.18}$$

Using Theorem 16 and a simple calculation to evaluate $A^{(1)}_1$, as defined in equation (E.7), we find the fidelity in terms of the *original* value of the first moment to be

$$F^{(k)}_{\frac{1}{2}} = \frac{1}{2} + \frac{\mathrm{Tr}[\rho^{(0)}_j J_{\hat{n}}]}{2j+1}\left(1 - \frac{2}{(2j+1)^2}\right)^k. \tag{E.19}$$

Note that the fidelity function in this case depends only on the first moment.

We now generalize to the case where we wish to measure the angular momentum of a spin-1 particle along some direction using a quantum direction reference frame. (Larger spins can be handled by similar means.) First, we determine an expression for the fidelity in terms of the moments:

$$F^{(k)}_1 = \mathrm{Tr}\left[\frac{1}{3}\sum_{\mu} \Pi_{j+\mu}(\rho^{(k)}_j \otimes |\mu\rangle\langle\mu|)\right]. \tag{E.20}$$

where $\mu \in \{-1, 0, 1\}$. Consider the disturbance map ξ as described in Section E.1.2:

$$\xi(\rho^{(k)}_j) = \frac{1}{3}\sum_{\mu'',\mu',\mu} \langle\mu''|\Pi_{j+\mu'}|\mu\rangle\ \rho^{(k)}_j\ \langle\mu|\Pi_{j+\mu'}|\mu''\rangle. \tag{E.21}$$

Explicit expressions for the reduced operators on the reference system $\langle\mu''|\Pi_{j+\mu'}|\mu\rangle$ can be found from the Clebsch-Gordon coefficients. The expression for the fidelity in terms of j and the moments is

$$F_1^{(k)} = \frac{1}{6} + \frac{\left[(2j+1)^2 - 2\right]}{6j(j+1)(2j+1)}\mathrm{Tr}[\rho_j^{(k)} J_z] + \frac{1}{2j(j+1)}\mathrm{Tr}[\rho_j^{(k)} J_z^2]. \tag{E.22}$$

Note that the fidelity in this case depends on the second moment as well as the first. Using Theorem 16, we have:

$$\begin{aligned} \mathrm{Tr}[\xi(\rho_j^{(k)})J_z] &= A_1^{(1)}\mathrm{Tr}[\rho_j^{(k)} J_z], \quad (E.23)\\ \mathrm{Tr}[\xi(\rho_j^{(k)})J_z^2] &= A_0^{(2)} + A_2^{(2)}\mathrm{Tr}[\rho_j^{(k)} J_z^2]. \quad (E.24)\end{aligned}$$

Substituting some particular values of ρ_j (i.e., $\rho_j = |j,m\rangle_{\hat{n}}\langle j,m|$ for $m = j, j-1$) and solving the system of equations, we obtain:

$$\begin{aligned} A_1^{(1)} &= \frac{3j^4+6j^3-j^2-4j+2}{3j^2(j+1)^2} = 1 - \frac{4}{3j^2} + O\left(\frac{1}{j^3}\right), \quad (E.25)\\ A_0^{(2)} &= \frac{2\left(8j^4+16j^3-8j-3\right)}{3j(j+1)(2j+1)^2} = \frac{4}{3} - \frac{5}{3j^2} + O\left(\frac{1}{j^3}\right), \quad (E.26)\\ A_2^{(2)} &= \frac{4j^6+12j^5-3j^4-26j^3+j^2+16j+6}{j^2(j+1)^2(2j+1)^2} = 1 - \frac{4}{j^2} + O\left(\frac{1}{j^3}\right). \quad (E.27)\end{aligned}$$

(The series apply as $j \to \infty$.) In terms of these constants, the fidelity evolves with k, the number of applications of the map corresponding to the measurement of the z projection of the spin-1 system, as:

$$F_1^{(k)} = \frac{1}{6} + \frac{\left[(2j+1)^2-2\right]}{6j(j+1)(2J+1)}\left(A_1^{(1)}\right)^k \mathrm{Tr}[\rho_j^{(0)} J_z] + \frac{1}{2j(j+1)}\left(A_0^{(2)}\frac{1-\left(A_2^{(2)}\right)^k}{1-A_2^{(2)}} + \left(A_2^{(2)}\right)^k \mathrm{Tr}[\rho_j^{(0)} J_z^2]\right). \tag{E.28}$$

Because the assumptions given in Section E.1.3 are satisfied, the longevity of the quantum reference frame must scale as $O(j^2)$. This result is easily verified numerically. See Fig. E.2.

Implementing a Pauli Operator on a Qubit

The quantum reference frame can also be used to define an axis for the purpose of implementing some desired unitary gate. Having only one axis, we are restricted to the set of gates corresponding to rotations about this axis. Suppose, for example, we want to implement a Pauli Z operation on a qubit,

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{E.29}$$

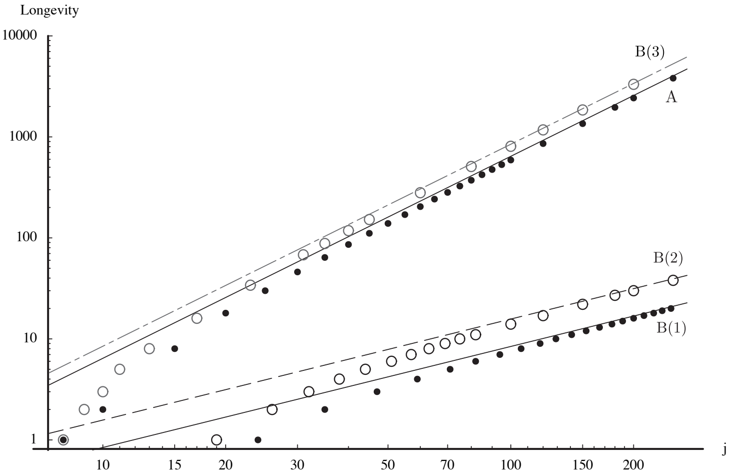


Figure E.2: A plot of the longevity, as defined in Section E.1.3, against the reference system size j for the scenario described in Example A (points converging to the solid line of gradient 2), and Example B, Method 1 (points converging to the solid line of gradient 1), Method 2 (circles converging to the dashed line of gradient 1), and Method 3 (grey circles converging to grey dot-dashed line of gradient 2). The points are found numerically. The gradients of the lines in A and B(3) are 2, indicating longevity scaling as $O(j^2)$, and the gradients of the lines in B(1) and B(2) are 1, indicating a scaling of $O(j)$.

using a quantum reference direction ρ_j in the form of a spin-j system to define the z-axis. To do this, we must implement a rotationally-invariant operation on the combined reference and system. Unlike the measurement considered in previous examples, in this case we wish to implement a unitary operation on the system that is conditional on the state of the quantum reference direction. In the language of quantum information, we want to perform a unitary rotation about a direction *programmed* by the state of the reference frame. However, that because the size of the reference frame is bounded and the number of programmed operations (possible directions) is continuous, then the impossibility of a "programmable quantum gate" [137] ensures that an ideal such conditional unitary cannot exist. It is possible to approximate it, as we now show.

Method 1:

Consider the following operation, to be performed on the joint state of a quantum reference frame ρ_j and a single qubit σ. First, define an covariant measurement on the spin-j reference frame described by the POVM

$$\{\Lambda(\Omega) = (2j+1)R(\Omega)|e\rangle\langle e|R(\Omega)^\dagger\,,\ \Omega \in SU(2)\}\,, \tag{E.30}$$

where $|e\rangle$ is a normalized state of a spin-j system. The measurement effects of this POVM satisfy the normalization condition

$$\int_\Omega d\mu_\Omega \Lambda(\Omega) = I_j \,, \tag{E.31}$$

where I_j is the identity on the spin-j Hilbert space and $d\mu_\Omega$ represents the Haar measure over $SU(2)$. This measurement is performed on the quantum reference frame state ρ_j, and then conditional upon obtaining the outcome Ω, the operation $Z(\Omega) = R(\Omega)ZR^{-1}(\Omega)$ is applied to the system. If we subsequently discard the information about the measurement result Λ, then the effective joint map

$$\chi(\rho_j \otimes \sigma) = \int_\Omega d\mu_\Omega \sqrt{\Lambda(\Omega)} \otimes Z(\Omega)(\rho_j \otimes \sigma)\sqrt{\Lambda(\Omega)} \otimes Z(\Omega)^\dagger \,, \tag{E.32}$$

is clearly invariant. (Although the covariant measurement makes use of an external reference frame, the resulting map χ is independent of this choice of frame.) The net operation on the system σ is given by the map

$$\tau(\sigma) = \int_\Omega d\mu_\Omega \mathrm{Tr}_j \left[\Lambda(\Omega)\rho_j\right] Z(\Omega)\sigma Z^{-1}(\Omega) \,. \tag{E.33}$$

Note that this operation τ is also rotationally invariant; although the measurement and subsequent unitary appear to require an external spatial reference frame, the net operation is invariant under changes of this frame. Also, because the operation is constructed explicitly from a POVM measurement and a unitary operation conditional on this classical result, followed by tracing out the state of the reference frame, it is necessarily completely positive and trace preserving (CPTP). (This fact is also clear by observation: the term $\mathrm{Tr}_j \left[\Lambda(\Omega)\rho_j\right]$ is a normalized probability distribution weighting possible unitaries $Z(\Omega)$, and thus the map τ is a valid unital CPTP map.)

This expression for τ explicitly gives a Kraus decomposition of this map (albeit with a continuous number of Kraus operators), $\tau(\sigma) = \int_\Omega d\mu_\Omega E(\Omega)\sigma E(\Omega)^\dagger$, where

$$E(\Omega) = \sqrt{\mathrm{Tr}_j[\Lambda(\Omega)\rho_j]}Z(\Omega) \,, \tag{E.34}$$

are Kraus operators satisfying $\int_\Omega d\mu_\Omega E(\Omega)^\dagger E(\Omega) = I$. The ability of this operation to approximate the Z operation on the system is defined using the *gate fidelity* [84, 30, 136, 58] given by

$$F_{\mathrm{gate}}(Z,\tau) \equiv \frac{\int_\Omega d\mu_\Omega \left|\mathrm{Tr}[E(\Omega)^\dagger Z]\right|^2 + d}{d^2 + d} \tag{E.35}$$

where d (the dimension of the system on which the gate is applied) is 2 in our case.

Suppose that quantum directional reference frame ρ_j is aligned with the z-axis. Then ρ_j is a mixture of states $|j,m\rangle_z$ for $-j \leqslant m \leqslant j$. To simplify our calculation, assume that

$\rho_j = |j,m\rangle_z\langle j,m|$ (i.e., we consider only a pure state, but at the end, we can generalize for any mixed state). The state $|e\rangle$ that defines the POVM could be any state, but for simplicity we consider only the case where $|e\rangle = |j,j\rangle_z$. In this case, the rotationally-invariant operation on the combined reference and system is

$$\tau(\sigma,\rho_j) = (2j+1)\int_\Omega d\mu_\Omega \big|{}_z\langle j,m|R(\Omega)|j,j\rangle_z\big|^2 Z(\Omega)\sigma Z^{-1}(\Omega)\,. \tag{E.36}$$

The Kraus operators are then given by

$$E(\Omega) = \sqrt{(2j+1)}\big|{}_z\langle j,m|R(\Omega)|j,j\rangle_z\big|Z(\Omega)\,. \tag{E.37}$$

We can parameterize

$$R(\Omega) = e^{i\alpha}\begin{pmatrix} e^{i\phi}\cos\theta & e^{i\psi}\sin\theta \\ -e^{-i\psi}\sin\theta & e^{-i\phi}\cos\theta \end{pmatrix} \tag{E.38}$$

where $0 \le \alpha,\psi,\phi \le 2\pi$ and $0 \le \theta \le \frac{\pi}{2}$. The rotation $R(\Omega)$ can be rewritten using Euler angles:

$$R(\Omega) = e^{i\alpha}R_z(-\phi-\psi)R_y(-2\theta)R_z(\psi-\phi), \tag{E.39}$$

where $R_k(\beta)$ is a clockwise rotation of angle β around the k-axis. It follows that

$$\mathrm{Tr}[R(\Omega)ZR^\dagger(\Omega)Z] = \mathrm{Tr}[R_y(-2\theta)ZR_y(2\theta)Z] = 2\cos 2\theta. \tag{E.40}$$

From [156], we have that

$${}_z\langle j,m|R(\Omega)|j,j\rangle_z = e^{-i\alpha-im(\phi+\psi)-ij(\phi-\psi)}\sqrt{\binom{2j}{j+m}}(\cos\theta)^{j+m}(-\sin\theta)^{j-m}. \tag{E.41}$$

Therefore,

$$\begin{aligned}|\mathrm{Tr}[E(\Omega)^\dagger Z]|^2 &= (2j+1)\binom{2j}{j+m}(\cos\theta)^{2j+2m}(\sin\theta)^{2j-2m}\big|\mathrm{Tr}[R_y(-2\theta)ZR_y(2\theta)Z]\big|^2 \\ &= 4(2j+1)\binom{2j}{j+m}(\cos\theta)^{2j+2m}(\sin\theta)^{2j-2m}(\cos 2\theta)^2,\end{aligned} \tag{E.42}$$

where, in the first step, we used the cyclic propriety of the trace. The Haar measure in these coordinates is $d\Omega(U(2)) = (2\pi)^{-3}2\sin\theta\cos\theta d\theta d\alpha d\phi d\psi$. It follows that

$$F_{gate}(Z,\tau) = \frac{1}{3} + \frac{2(j+1+2m^2)}{3(j+1)(2j+3)} \tag{E.43}$$

which is a function of the second moment of the z projection of the reference frame. The result easily generalize to the case where ρ_j to is a mixture of pure state of the form $|j,m\rangle_z$:

$$F_{gate}(Z,\tau) = \frac{1}{3} + \frac{2(j+1+2\mathrm{Tr}[\rho_j J_z^2])}{3(j+1)(2j+3)}. \tag{E.44}$$

This example is a case where the fidelity evolves with the expectation value of the second moment, despite the fact that the reservoir is composed of spin-$\frac{1}{2}$ particles. Also note that in this particular case the fidelity does not depend on any odd moments because the operation depends only on the axis defined by $\hat{n}$, but not on the direction along this axis.

We now consider how the quantum reference frame degrades with repeated use in performing this operation. We assume that the qubit systems (the reservoir) on which we apply the approximate phase gate are all in the completely mixed state $\frac{1}{2}I_{\frac{1}{2}}$. With each application, the reference frame evolves according to the invariant map $\xi(\rho_j) = \mathrm{Tr}_s\chi(\rho_j \otimes \frac{1}{2}I_{\frac{1}{2}})$, where Tr_s is the partial trace of the qubit system on which we apply the approximate phase gate. To calculate how the second moment evolves as a function of the number of times the quantum reference frame has been used to perform the approximate phase gate, we can use Theorem 16:

$$\mathrm{Tr}[\xi(\rho_j)J_z^2] = A_0^{(2)} + A_2^{(2)}\mathrm{Tr}[\rho_j J_z^2]. \tag{E.45}$$

To find the values of the coefficients $A_0^{(2)}$ and $A_2^{(2)}$, we consider two possible initial states of the quantum reference frame. First, suppose that $\rho_j = \frac{1}{2j+1}I_j$, which yields $\xi(\rho_j) = \frac{1}{2j+1}I_j$. This evolution gives

$$\frac{j(j+1)}{3} = A_0^{(2)} + A_2^{(2)}\frac{j(j+1)}{3}. \tag{E.46}$$

Second, using $\rho_j = |j,j\rangle_z\langle j,j|$, we obtain a second equation:

$$\frac{j(1+j(3+j+2j^2))}{(1+j)(3+2j)} = A_0^{(2)} + A_2^{(2)}j^2. \tag{E.47}$$

Solving those equations, we obtain

$$A_0^{(2)} = j - \frac{2j}{2j+3} = j - 1 + \frac{3}{2j} - \frac{9}{4j^2} + O\left(\frac{1}{j^3}\right), \tag{E.48}$$

and

$$A_2^{(2)} = 1 - \frac{3(2j+1)}{(2j+3)(j+1)} = 1 - \frac{3}{j} + O\left(\frac{1}{j^3}\right). \tag{E.49}$$

From the above equations, we can deduce that the longevity of the quantum reference frame is $O(j)$ by noting that the fidelity after k repetitions of the gate is

$$F_{gate}^{(k)}(Z,\tau) = \frac{1}{3} + \frac{2}{3(j+1)(2j+3)}\left(j+1+2A_0^{(2)}\frac{1-\left(A_2^{(2)}\right)^k}{1-A_2^{(2)}} + 2\left(A_2^{(2)}\right)^k\mathrm{Tr}[\rho_j^{(0)}J_z^2]\right). \tag{E.50}$$

From the equations (E.48) and (E.49) we have that $A_0^{(2)}$ goes as $O(j)$ and $A_2^{(2)}$ goes as $O(1)$. Looking at equation (E.50) we see that fidelity must go as $O(\frac{1}{j})$ and therefore longevity goes as $O(j)$. This is also seen in the numerical work shown in Figure E.2. This result appears to be in contradiction with Theorem 17. However, this contradiction is resolved by the fact that one of the assumptions of the theorem is not fulfilled. Indeed, even if the map χ is rotationally invariant, it does not conserve the total angular momentum. In particular, note that $\langle j, m|_z \xi_j(|j,j\rangle_z\langle j,j|)|j,m\rangle > 0$ for all m in the range $-j,\dots,j$. Because $\zeta^{\circ 2j}(|j,j\rangle_z\langle j,j|)$ is the only map of the form $\zeta^{\circ n}(|j,j\rangle_z\langle j,j|)$ for $0 < n < 2j+1$ that has support in the state $|j,-j\rangle_z\langle j,-j|$, then $q_{2j} > 0$. Therefore, there is no bound n_{max} independent of j such $q_n = 0$ for all $n > n_{max}$.

Method 2:

We now investigate an alternate method for approximating a Pauli Z operation using a quantum reference frame. Consider the filtering operation [74]

$$\Gamma = (2j+1)\int_\Omega d\mu_\Omega R_j(\Omega)\otimes R_{1/2}(\Omega)\Big[|j,j\rangle_z\langle j,j|\otimes Z\Big]R_j(\Omega)^{-1}\otimes R_{1/2}(\Omega)^{-1}. \tag{E.51}$$

Observing that in spin notation

$$Z = |\tfrac{1}{2},\tfrac{1}{2}\rangle_z\langle\tfrac{1}{2},\tfrac{1}{2}| - |\tfrac{1}{2},-\tfrac{1}{2}\rangle_z\langle\tfrac{1}{2},-\tfrac{1}{2}|, \tag{E.52}$$

we can replace the operator Z in equation (E.51) with this expression and using the Clebsch-Gordon coefficients we can express the part of equation (E.51) that is in the square brackets as a joint system with total angular momentum having support on the $j-\frac{1}{2}$ and $j+\frac{1}{2}$ subspaces. Γ is spatially-covariant and we can use Schur's lemma to note that it will be block-diagonal in the irreps $j+\frac{1}{2}$ and $j-\frac{1}{2}$. Thus, we can rewrite it as

$$\begin{aligned}\Gamma = (2j+1)\int_\Omega & d\mu_\Omega R_{j+1/2}(\Omega)\Big[|j+\tfrac{1}{2},j+\tfrac{1}{2}\rangle_z\langle j+\tfrac{1}{2},j+\tfrac{1}{2}| - \tfrac{1}{2j+1}|j+\tfrac{1}{2},j-\tfrac{1}{2}\rangle_z\langle j+\tfrac{1}{2},j-\tfrac{1}{2}|\Big]R^{-1}_{j+1/2}(\Omega)\\ & -(2j+1)\int_\Omega d\mu_\Omega R_{j-1/2}(\Omega)\Big[\tfrac{2j}{2j+1}|j-\tfrac{1}{2},j-\tfrac{1}{2}\rangle_z\langle j-\tfrac{1}{2},j-\tfrac{1}{2}|\Big]R_{j-1/2}(\Omega)^{-1} \qquad \text{(E.53)}\\ = (2j+1)&\Big(\tfrac{1}{2j+2}(1-\tfrac{1}{2j+1})I_{j+1/2} - \tfrac{1}{2j}\tfrac{2j}{2j+1}I_{j-1/2}\Big), \qquad \text{(E.54)}\end{aligned}$$

which simplifies to

$$\Gamma = (\frac{2j}{2j+2}I_{j+1/2} - I_{j-1/2}). \tag{E.55}$$

We can define

$$\chi_j(\rho_j\otimes\sigma) = \Gamma(\rho_j\otimes\sigma)\Gamma^\dagger, \tag{E.56}$$

where Π_{j+k} is a projection into the subspace of total angular momentum $j+k$. However, this map is not trace preserving, nor does it conserve angular momentum. As in method 1, let $\rho_j = |j,m\rangle_z\langle j,m|$. In that case, we have Kraus operators of the form

$$E_{m'} \;=\; {}_z\langle j,m'|\Gamma|j,m\rangle_z \tag{E.57}$$

and, using the appropriate Clebsch Gordon coefficients, we arrive at an expression for fidelity of the form

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{1}{j+1}\right)^2 m^2, \tag{E.58}$$

which goes to 1 in the limit $j \to \infty$, for the case $m = j$. For a general state ρ_j (which must be diagonal in the basis $\{|j,m\rangle_z \mid -j \leqslant m \leqslant j\}$), we obtain

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{1}{j+1}\right)^2 \mathrm{Tr}[\rho_j J_z^2]. \tag{E.59}$$

By Theorem 16, even though the map is not trace preserving, we have:

$$\mathrm{Tr}[\xi(\rho_j)J_z^2] = A_0^{(2)} + A_2^{(2)}\mathrm{Tr}[\rho_j J_z^2], \tag{E.60}$$

and, by the same method as in the previous subsection, we have for this map

$$\begin{aligned} A_0^{(2)} &= \frac{j}{(j+1)} = 1 - \frac{1}{j} + \frac{1}{j^2} + O\left(\frac{1}{j^3}\right), & \text{(E.61)}\\ A_2^{(2)} &= \frac{j}{(j+1)}\left(1 - \frac{3}{j(j+1)}\right) = 1 - \frac{1}{j} - \frac{2}{j^2} + O\left(\frac{1}{j^3}\right), & \text{(E.62)}\end{aligned}$$

where the series expansions are appropriate in the limit $j \to \infty$.

Again we find the scaling of longevity with j to be $O(j)$. (See Figure E.2.) However, the longevity is consistently higher than by using Method 1.

Method 3:

We might ask about performing an operation similar to the previous case, but in which trace is preserved on the map. Such a map could be realized through a coupling between the spins of the form:

$$H = w\left(J_x^{(\frac{1}{2})}J_x^{(j)} + J_y^{(\frac{1}{2})}J_y^{(j)} + J_z^{(\frac{1}{2})}J_z^{(j)}\right) = \frac{w}{2}\mathcal{J}^2 - \frac{\omega}{2}\left(j(j+1) + \frac{3}{4}\right)\mathbb{1} \tag{E.63}$$

where w is some constant, the operator $J_x^{(\frac{1}{2})} = X$ is the J_x operator on a spin-half system and the operator $\mathcal{J}^2 = \left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)\Pi_{j+\frac{1}{2}} + \left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)\Pi_{j-\frac{1}{2}}$ is the total angular momentum operator on the joint system of the reference and the spin-half system. Therefore, the unitary associated with this Hamiltonian is $U = e^{i\frac{w}{2}(\mathcal{J}^2-(j^2+j+\frac{3}{4})\mathbb{1})t}$. Choosing $t = \frac{2\pi}{w(2j+1)}$ gives the appropriate phase shift of the spin-half system relative to the spin-j reference system.

It could also be phrased in the language of the previous subsection if — up to a global phase — we set

$$\Gamma = U \cong (I_{j+1/2} - I_{j-1/2}). \tag{E.64}$$

The quality function for this map is

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{2}{2j+1}\right)^2 \mathrm{Tr}[\rho_j J_z^2], \tag{E.65}$$

which, when $\mathrm{Tr}[\rho_j J_z^2] = j^2$, also tends to 1 in the limit $j \to \infty$.

We have for this map

$$A_0^{(2)} = 1 - \frac{1}{(2j+1)^2} = 1 - \frac{1}{4j^2} + O\left(\frac{1}{j^3}\right), \qquad \text{(E.66)}$$

$$A_2^{(2)} = 1 - \frac{12}{(2j+1)^2} = 1 - \frac{3}{j^2} + O\left(\frac{1}{j^3}\right), \qquad \text{(E.67)}$$

where, again, the series expansions are appropriate in the limit $j \to \infty$. Observe that in this case these parameters lack any terms in $\frac{1}{j}$.

Now we find the scaling of longevity with j to be $O(j^2)$, as depicted in Figure E.2. Figure E.3 compares the fidelities for the three methods for constant j. These last three examples demonstrate the importance of choosing the gate carefully to match the objectives of the given task.

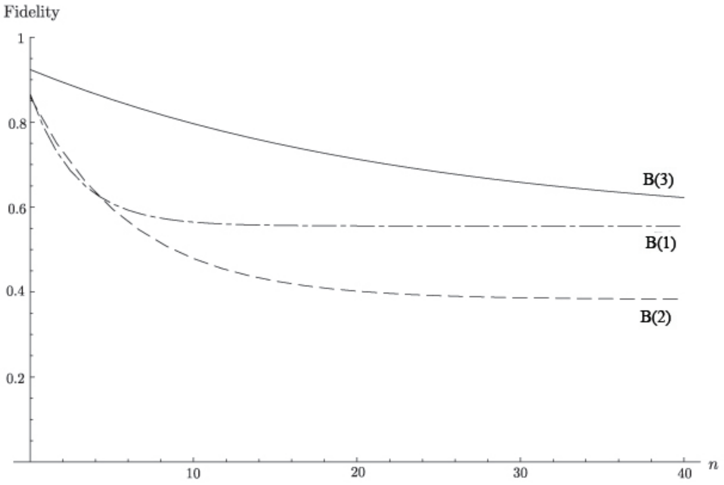


Figure E.3: A plot of the fidelity with number of repetitions, n, for $j = 8$ for the three methods, B(1) (dot-dashed line), B(2) (dashed line), and B(3) (solid line). This behavior of this value of j is representative.

E.1.5 Conclusion

This study has explored the physically relevant question of how a quantum directional reference frame state evolves under the action of maps invariant with respect to rotations in space. We generalize the concept of quality of a quantum directional reference frame introduced by [13]. We argue the quality of a reference frame must be represented by a function that depends only on the eigenvalues of the quantum reference frame or an equivalent set of parameters called the moments. We give recursive equations (Theorem 16) for how the moments evolve with the number of uses of the quantum reference frame. We derive sufficient some conditions (Theorem 17) for the longevity of a quantum reference frame to scale by a factor proportional to square the dimension of the quantum reference frame. Finally, we applied our results to different examples such as the use of a quantum directional reference frame to measure a spin-1 particle or to implement an Pauli operator on a qubit. The tools that we developed can be use to compare different methods to perform some operation using a quantum reference frame as we showed in our last example.

To prove Theorem 16, we use Theorem 15 which implies that maps invariant with respect to $SU(2)$ can be written in a polynomial form as a function of the Lie algebra generators. It would be interesting to investigate if a similar theorem could be applied to other Lie groups. Also, we assume in this analysis that the state of the reservoir is invariant under rotations. However, in the most general physical situation, the reservoir can be polarized. Building on the work done in [143], it would be interesting to investigate how our theorem can be generalized to the case where the above symmetry is broken.

E.1.6 Acknowledgments

We wish to acknowledge the contribution of Terry Rudolph to show that the coefficients q_i's in Theorem 15 are real. We also thank Giacomo Mauro D'Ariano, Joseph Emerson, Michael Keyl, David Kribs, David Poulin and Peter Turner for helpful discussions. J.-C.B and M.L. are supported by NSERC, and L.S. by the Mike and Ophelia Lazaridis Fellowship. S.D.B. acknowledges the support of the Australian Research Council.

E.1.7 Appendix: Proof of Theorem 1

Here, we provide a proof of Theorem 15, inspired by ideas from [95, 47]. First, because the group $SO(3)$ of space rotations and the group $SU(2)$ of unitary transformation are locally isomorphic up to a phase, then the spatial covariance condition can be replaced by covariance with respect to $SU(2)$.

Let ρ_j be a density operator on the Hilbert space of a spin-j system. Consider the map

$$\zeta(\rho_j) = \frac{1}{\lambda}(J_x \rho J_x + J_y \rho J_y + J_z \rho J_z), \tag{E.68}$$

where $\lambda = j(j+1)$, and J_x, J_y and J_z are the angular momentum operators in the x, y and z directions for some arbitrary Cartesian frame. First, we notice that

$$\zeta(I_j) = \frac{1}{\lambda}\sum_{i=1}^{3} J_i^\dagger J_i = \frac{1}{\lambda}J^2 = I_j, \tag{E.69}$$

where J is the total angular momentum operator.

Lemma 3 *The map ζ is invariant with respect to the spin-j irreducible representation R_j of $SU(2)$.*

Proof. If R_j is the $d = 2j + 1$ dimensional irreducible representation (irrep) of the group $SU(2)$ generated by $\{J_x, J_y, J_z\}$, then it can be expressed in the form

$$R_{\hat{n}}(\theta) = e^{-i\theta\hat{n}\cdot\vec{J}}, \quad \vec{J} = (J_x, J_y, J_z), \tag{E.70}$$

for some angle θ and unit vector $\hat{n}$. By the symmetry of ζ with respect to the (arbitrary) choice of axes, if we prove that

$$R_z(\theta)^{-1}\zeta(R_z(\theta)\rho_j R_z(\theta)^{-1})R_z(\theta) = \zeta(\rho_j), \qquad \forall\ \theta \in [0, 2\pi), \tag{E.71}$$

for any rotations around the z-axis only, then (E.71) with any rotations follows immediately. Such rotations will map

$$J_x \quad \to \quad \cos\theta J_x + \sin\theta J_y \tag{E.72}$$

$$J_y \quad \to \quad -\sin\theta J_x + \cos\theta J_y, \tag{E.73}$$

and it can be readily computed that ζ is invariant with respect to z rotations. Therefore,

$$R_j(\Omega)^{-1}\zeta(R_j(\Omega)\rho_j R_j(\Omega)^{-1})R_j(\Omega) = \zeta(\rho_j) \qquad \forall\ \Omega \in SU(2). \tag{E.74}$$

The map ζ is therefore invariant. ■

Because a composition of invariant maps is also invariant, then any map of the form

$$\xi(\rho) = q_0\rho + \sum_{k=1}^{2j} q_k\zeta^{\circ k}(\rho), \tag{E.75}$$

where the values q_i are real numbers and $\zeta^{\circ k} = \zeta \circ \zeta \circ \ldots \circ \zeta$, is also invariant with respect to $SU(2)$. It remains to show that any invariant map with respect to $SU(2)$ can be written on the form (E.75).

To prove that every invariant map can be written as above, we use the Liouville representation of a superoperator [79]. Upon representing a $d \times d$ density matrix ρ into a d^2 long column vector $|\rho\rangle\rangle$ by stacking the columns of the density matrix, the action of any superoperator ξ can be represented as a $d^2 \times d^2$ matrix $\mathcal{K}(\xi)$, such that

$$|\xi(\rho)\rangle\rangle = \mathcal{K}(\xi)|\rho\rangle\rangle. \tag{E.76}$$

This representation is necessarily basis dependent. If a given process has Kraus operators $\{E_k\}$, the Liouville representation takes the form

$$\mathcal{K}(\xi) = \sum_k E_k^* \otimes E_k \tag{E.77}$$

where $*$ represents the complex conjugate with respect to the chosen basis.

Lemma 4 *The Liouville representation of any map that is invariant with respect to $SU(2)$ has the form*

$$\mathcal{K}(\xi) = \sum_{k=0}^{2j} c_k \Pi_k, \tag{E.78}$$

where $c_k \in \mathbb{C}$ and Π_k is the projector into the $2k+1$ dimensional subspace of total angular momentum k.

Proof. The condition on a map to be invariant can thus be expressed as

$$(R_j^*(\Omega) \otimes R_j(\Omega))\mathcal{K}(\xi) = \mathcal{K}(\xi)(R_j^*(\Omega) \otimes R_j(\Omega)), \qquad \forall\, \Omega \in SU(2), \tag{E.79}$$

so the condition on the operators in this new representation is that $\mathcal{K}(\xi)$ must commute with $R_j^* \otimes R_j$. If the group $SU(2)$ is represented in the $\{|j,m\rangle_z\}$ basis (i.e. the eigenstate of J_z), we have the relation $R_j^*(\Omega) = e^{-i\pi J_y} R(\Omega) e^{i\pi J_y}$, which implies

$$(R_j(\Omega) \otimes R_j(\Omega))(e^{i\pi J_y} \otimes I)\mathcal{K}(\xi)(e^{-i\pi J_y} \otimes I) = (e^{i\pi J_y} \otimes I)\mathcal{K}(\xi)(e^{-i\pi J_y} \otimes I)(R_j(\Omega) \otimes R_j(\Omega)) \;\; \forall\, \Omega \in SU(2). \tag{E.80}$$

We note that $(R_j(\Omega) \otimes R_j(\Omega))$ is the *collective* representation of $SU(2)$ on two spin-j systems. A Clebsch-Gordon decomposition gives us the irrep of the group of all collective rotations $\mathcal{G}(R_j \otimes R_j)$ which take the form

$$\mathcal{G}(R_j \otimes R_j)(\Omega) \simeq \bigoplus_{k=0}^{2j} R_k(\Omega), \qquad \forall\, \Omega \in SU(2), \tag{E.81}$$

where "$\simeq$" denotes "unitarily equivalent" and R_k is the spin-k irrep of $SU(2)$ which has multiplicity one.

Because we require $\mathcal{K}(\xi)$ to commute with $R_j(\Omega) \otimes R_j(\Omega)$ for any $\Omega \in SU(2)$, it must then commute with all irreps of $\mathcal{G}(R_j \otimes R_j)$. By Schur's lemma, we have that

$$\mathcal{K}(\xi) \simeq \bigoplus_{k=0}^{2j} c_k I_k = \sum_{k=0}^{2j} c_k \Pi_k, \tag{E.82}$$

where $c_k \in \mathbb{C}$, I_k is the $2k+1$ dimensional identity operator, and Π_k is the projector into the $2k+1$ dimensional subspace of total angular momentum k. ■

There are $2j+1$ independent projectors forming $\mathcal{K}(\xi)$. To characterize every possible invariant mapping, we will thus also require $2j+1$ independent operators.

Lemma 5 *The matrices $(\sum_i J_i^* \otimes J_i)^k$ for $0 \leqslant k \leqslant 2j$ are linearly independent and form a complete basis to represent any invariant map of the form (E.78).*

Proof. If we rewrite equation (E.75) in the Liouville representation, the map becomes

$$\mathcal{K}(\xi) = \sum_{k=0}^{n} \frac{q_k}{\lambda^k} \left(\sum_{i \in \{x,y,z\}} J_i^* \otimes J_i \right)^k . \tag{E.83}$$

The hermitian matrices $(\sum_i J_i^* \otimes J_i)^k$ are diagonal in the same basis, and the eigenvalues of $(\sum_i J_i^* \otimes J_i)^k$ are ν_l^k, where ν_l is an eigenvalue of $\sum_i J_i^* \otimes J_i$.

To find all of the eigenvalues $\{\nu_l\}$, expand the total angular momentum

$$\mathcal{J}^2 = \sum_i (J_i \otimes I + I \otimes J_i)^2 \tag{E.84}$$

to get a new expression for $\sum_i J_i \otimes J_i$:

$$\begin{aligned} \sum_i J_i \otimes J_i &= \frac{1}{2}(\sum_i (J_i \otimes I + I \otimes J_i)^2 - \sum_i J_i^2 \otimes I - I \otimes \sum_i J_i^2) \\ &= \frac{1}{2}(\mathcal{J}^2 - J^2 \otimes I - I \otimes J^2) \\ &= \frac{1}{2}(\mathcal{J}^2 - 2j(j+1)I) \end{aligned} \tag{E.85}$$

With the relation $\sum_i J_i^* \otimes J_i = -e^{-i\pi J_y} \otimes I \left(\sum_i J_i \otimes J_i\right) e^{i\pi J_y} \otimes I$, we thus have that $\sum_i J_i^* \otimes J_i$ and $\sum_i J_i \otimes J_i$ will have the same eigenvalues up to a negative sign.

From equation (E.85), we find that

$$\nu_l = -\frac{1}{2}\left(l(l+1) - 2j(j+1)\right), \tag{E.86}$$

where $l(l+1)$ is the eigenvalue resulting from the vector addition of the two spin-J systems, which implies that l can range from 0 to $2j$ and has multiplicity $2l+1$. There is $2j+1$ different values of l, so there is $2j+1$ different eigenvalues. By the fundamental theorem of algebra, this implies that there exist no polynomial of degree $2j$ that has the eigenvalues $\{v_l\}$ as roots. This implies the matrices $(\sum_i J_i^* \otimes J_i)^k$ for $0 \leqslant k \leqslant 2j$ are linearly independent. By counting the number free parameters, we proved that the operators $(\sum_i J_i^* \otimes J_i)^k$ form a complete basis to represent any map represented by equation (E.78). ■

Finally, to show that the equation (E.75) is a representation of all possible invariant maps on a spin-j system, we need to show that the q_i's are real. First, note that

$$\xi(\rho_j) = \frac{1}{2\lambda}(2J_z\rho_j J_z + J_+\rho_j J_+^\dagger + J_-\rho_j J_-^\dagger), \tag{E.87}$$

where $J_\pm = J_x \pm iJ_y$. Also, from equation (E.87), the only contribution to

$${}_{\hat{n}}\langle j,-j|\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)|j,-j\rangle_{\hat{n}} \tag{E.88}$$

is from $q_{2j}\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)$. Actually,

$$_{\hat{n}}\langle j,-j|\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)|j,-j\rangle_{\hat{n}} = q_{2j}\ {}_{\hat{n}}\langle j,-j|\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)|j,-j\rangle_{\hat{n}}. \tag{E.89}$$

Because $\xi(\rho_j)$ must be positive, then q_{2j} is a non-negative real number. The contributions to

$$_{\hat{n}}\langle j,-j+1|\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)|j,-j+1\rangle_{\hat{n}} \tag{E.90}$$

are from

$$q_{2j}\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|) \tag{E.91}$$

and

$$q_{2j-1}\zeta^{\circ 2j-1}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|). \tag{E.92}$$

Since $\xi(\rho_j)$ is positive and q_{2j} is real, then q_{2j-1} must also be real. Note that q_{2j-1} could be negative. By induction, it is easy to show that all the coefficients q_i must be real.

Bibliography

[1] S. Aaronson, and A. Ambainis, Proceedings of FOCS pp. 200–209 (2003).

[2] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, *Phys. Rev. Lett.* **98** 230501 (2007).

[3] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani, in *Proc. 33th ACM Symposium on Theory of Computing* (ACM Press New York, New York, NY, 2001), pp. 50–59.

[4] D. Aharonov, and M. Ben-Or, in *Proceedings of the 29th Annual ACM Symposium on the Theory of Computation (STOC)*, (ACM Press, New York, 2996), pp. 176-188.

[5] Y. Aharonov, L. Davidovich, and N. Zagury, *Phys. Rev. A* **48**, 1687 (1993).

[6] A. Ambainis, J. Kempe, and A. Rivosh, *Proceedings of SODA* (2005), to appear.

[7] K. Azuma, *Tôhuku Math. J.* **19**, 357 (1967).

[8] E. Bagan, S. Iblisdir, and R. Munoz-Tapia, *Phys. Rev. A* **73** 022341 (2006).

[9] J. L. Ball, and K. Banaszek, *Open Syst. Inf. Dyn.* **12**, 121 (2005).

[10] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Phys. Rev. Lett.* **91** 027901 (2003).

[11] S. Bartlett, T. Rudolph, and R. Spekkens, *Phys. Rev. A* **70**, 032321 (2004).

[12] S. Bartlett, T. Rudolph, and R. Spekkens, *Rev. Mod. Phys.* **79**, 555 (2007).

[13] S. Bartlett, T. Rudolph, R. Spekkens, and P. Turner, *New J. Phys.* **8**, 58 (2006).

[14] J.S. Bell, *Physics(N.Y.)* **1**, 195 (1964).

[15] C. H. Bennett, *Phys. Rev. Lett.* **68**, 3121 (1992).

[16] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, *J. of Cryptology* **5** 3 (1992).

[17] C. H. Bennett, and G. Brassard, in *Proceeding of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore*, India (IEEE, New York, 1984), pp.175-179.

[18] C. H. Bennett, G. Brassard, and N. D. Mermin, *Phys. Rev. Lett. 68*, 557 (1992).

[19] C. H. Bennett, G. Brassard, and J.-M. Robert. *SIAM Journal on Computing*, 1**7** 210 (1988).

[20] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).

[21] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, *Theory of Cryptography: Second Theory of Cryptography Conference*, TCC 2005, J.Kilian (ed.) Springer Verlag 2005, vol. 3378 of Lecture Notes in Computer Science, pp. 386-406.

[22] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury, in *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing* (ACM Press, New-York, 2000) pp. 715-724.

[23] G. Bodenhausen, H. Kogler, and R. R. Ernst, J. Mag. Res. **58**, 370 (1984).

[24] J.-C. Boileau, D. Gottesman, R. Laflamme, D. Poulin, and R. W. Spekkens, *Phys. Rev. Lett.* **92**, 017901 (2004).

[25] J.-C. Boileau, R. Laflamme, M. Laforest, C. R. Myers, *Phys. Rev. Lett.* **93** 220501 (2004).

[26] J.-C. Boileau, D. Poulin, K. Tamaki and R. Laflamme, *to appear soon on arXiv.org.*

[27] J.-C. Boileau, K. Tamaki, J. Batuwantudawe, R. Laflamme and J.M. Renes, *Phys. Rev. Lett.* **94** 040503 (2005).

[28] N. Boulant, L. Viola, E. M. Fortunato, and D. G. Cory, *Physical Review Letters* **94**, 130501 (2005).

[29] M. Bourennane, M. Eibl, S. Gaertner, C. Kurtsiefer, A. Cabello, and H. Weinfurter, *Phys. Rev. Lett.* **92** 107901 2004.

[30] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and J. A. Jones, *Phys. Lett. A* **294**, 258 (2002).

[31] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **81** 5932 (1998).

[32] T. A. Brun, H. A. Carteret, and A. Ambainis, *Physical Review Letters* **91**, 130602 (2003).

[33] D. Bruss, *Phys. Rev. Lett.* **81** 3018 (1998).

[34] A. R. Calderbank, and P. W. Shor, *Physics Review A.* **54** 1098 (1996).

[35] J. L. Carter and M.N. Wegman, *Journal of Computer and System Sciences*, **22**, 265 (1981).

[36] H. F. Chau, *Phys. Rev. A* **66** 060302 (2002).

[37] A. Chefles, S. M. Barnett *Phys. Lett. A* **250** 223 (1998).

[38] T.-Y. Chen, J. Zhang, J.-C. Boileau, X.-M. Jin, B. Yang, Q. Zhang, T. Yang, R. Laflamme, and J.-W. Pan, *Phys.Rev.Lett.* **96**, 150504 (2006).

[39] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, in *Proc. 35th ACM Symposium on Theory of Computing* (ACM Press New York, 2002a), pp. 59–68.

[40] A. Childs, E. Farhi, and S. Gutmann, *Quantum Information Processing* **1**, 35 (2002b).

[41] M. Christandl, R Renner, and A. Ekert *arXiv:quant-ph/0402131.*

[42] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[43] J. Clauser, M. Horne, S. Shimony and R. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).

[44] D. G. Cory, M. D. Price, T. F. Havel, *Physica D.* **120**,**1-2**, 82 (1998).

[45] D. G. Cory, M.D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, *Phys. Rev. Lett.* **81**, 2152 (1998).

[46] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley-Interscience, USA, 1991), 542 pages.

[47] G. M. D'Ariano, *J. Math. Phys.* **45**, 3620 (2004).

[48] I. Damgaard, S. Fehr, L. Salvail, and C. Schaffner, *Proceedings of the 46th IEEE Symposium on Foundations of Computer Science* FOCS 2005, pages 449-458.

[49] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, *Phys. Rev. Lett.* **77**, 2818(1996).

[50] I. Devetak, A. Winter *Proc. R. Soc. Lond. A* **461** 207 (2005).

[51] Y. Dodis, S. J. Ong, M. Prabhakaran, and A. Sahai, in *Proc. FOCS '04*, pp. 196-205 (2004).

[52] Y. Dodis, and R. Renner, in *Proceedings of the 33rd ICALP*, (2006) *arXiv:quant-ph/0612012*.

[53] W. Dür, R. Raussendorf, V. M. Kendon, and H.J. Briegel, *Phys. Rev. A* **66**, 52319 (2002).

[54] J. Du, H. Li, X. Xu, M. Shi, J. Wu, X. Zhou, and R. Han, *Phys. Rev. A* **67**, 42316 (2003).

[55] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).

[56] A. Ekert, and R. Jozsa, *Phil. Trans. R. Soc. of Lond. A* **356**, 1769 (1998).

[57] J. Emerson, *unpublished*.

[58] J. Emerson, R. Alicki, and K. Zyczkowski, *J. Opt. B: Quantum Semiclass. Opt.* **7** (2005).

[59] E. Farhi, and S. Gutmann, *Phys. Rev. A* **58**, 915 (1998).

[60] R. P. Feynman, *Int. J. of Theor. Phys.* **21**, 467 (1982).

[61] E. M. Fortunato, M. A. Pravia, N. Boulant, G. Teklemariam, T. F. Havel, and D. G. Cory, *Journal of Chem. Phys.* **116**, 7599 (2002).

[62] J.D.Franson, *Phys. Rev. A* **45** 3126-3132 (1992).

[63] J.D. Franson and B.C. Jacobs, *Electron. Lett.* **31** 232-234 (1995).

[64] C.-H. F. Fung and H.-K. Lo, *Physical Review A* **74** 042342 (2006).

[65] N. Gisin, *Phys. Lett. A* **210**, 151 (1996).

[66] N. Gisin, J. Brendel, J-D. Gautier, B. Gisin, B. Huttner, G. Ribordy, W. Tittel and H. Zbinden *Lect. Notes Phys.* **538** 191-200 (2000).

[67] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, G. Ribordy, *Phys. Rev. A* **73** 022320 (2006).

[68] N. Gisin, S. Iblisdir, *quant-ph/0507118*.

[69] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Rev. Mod. Phys.* **74** 145 (2002).

[70] C. Gobby, Z. L. Yuan and A. J. Shields, *Appl. Phys. Lett.* **84**, 3762 (2004); T. Kimura *et. al.*, *quant-ph/0403104.*

[71] D. Gottesman, A. Kitaev, and J. Preskill, *Phys.Rev. A* **64** 012310 (2001).

[72] D. Gottesman and H.-K. Lo, *IEEE Transactions on Information Theory* **49**, 457 (2003).

[73] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Quantum Inf. Comp.* **4** 325 (2004).

[74] L. K. Grover, *Proceedings, 28th Annual ACM Symposium on the Theory of Computing*, 212 (1996).

[75] L. K. Grover, *Phys. Rev. Lett.* **79**, 4709 (1997).

[76] F. Haake, *Springer* (1992, 2001) ISBN 0172-7389.

[77] M. J. W. Hall, *Phys. Rev. Lett.* **74** 3307 (1995).

[78] E. L. Hanh, Phys. Rev. **80**, 580 (1950).

[79] T. F. Havel, *J. Math. Phys.* **44**, #2, 534 (2003).

[80] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. Wooters, *Phys. Rev. A 54* 1869 (1996).

[81] P.M. Hayden, D.W. Leung, D. Mayers, *The Universal Composable Security of Quantum Message Authentication with Key Recyling*, to appear soon.

[82] C.W. Helstrom, *Mathematics in Science and Engineering* **123** (1976).

[83] A. S. Holevo, *IEEE Trans. Info. Theor.* **44** 269 (1998).

[84] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A 60*, **1888** (1999).

[85] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim *arXiv:quant-ph/0608195.*

[86] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *quant-ph/0506189.*

[87] R.J. Hughes, G.L. Morgan and C.G. Peterson, *J. of Mod. Opt.* **47** 533-547 (2000).

[88] W.-Y. Hwang, *Phys. Rev. Lett.* **91**, 057901 (2003).

[89] K. Inoue, E. Waks, and Y. Yamamoto, *Phys. Rev. Lett.* **89** 037902 (2002).

[90] K.R.W. Jones, *Annals of Phys.* **207** 140-147 (1991).

[91] J. A. Jones, and E. Knill, *J. Mag. Res.* **141**, 322 (1999).

[92] R. Jozsa, and B. Schumacher, *J. Modern Optics.* **41** 2343 (1994).

[93] J. Kempe, D. Bacon, D. A. Lidar and, K. B. Whaley, *Phys. Rev. A* **63** 42307 (2001).

[94] V. Kendon and B. Tregenna, *Physical Review A* **67**, 042315 (2002).

[95] M. Keyl, and R. F. Werner, *J.Math.Phys.* **40**, 3283 (1999).

[96] T. Kimura, Y. Nambu, T. Hatanaka, A. Tomita, H. Kosaka and K. Nakamura, *arXiv:quant-ph/0403104.*

[97] A. Kitaev, *quant-ph/9511026.*

[98] A. Y. Kitaev, *Uspekhi Mat. Nauk.* **52**, 53 (1997).

[99] P. L. Knight, E. Roldán, and J. E. Sipe, *Optics Communications* **227**, 147 (2003).

[100] E. Knill, R. Laflamme, R. Martinez, and C. Negrevergne, *Phys. Rev. Lett.* **86**, 5811 (2001).

[101] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, *Nature* **404**, 368 (2000).

[102] E. Knill, R. Laflamme, and W. H. Zurek, *Science* **279**, 342 (1998a).

[103] E. Knill, R. Laflamme, and W. H. Zurek, *Phil. Trans. R. Soc. Lond. A.* **454**, 365 (1998b).

[104] M. Koashi, *J. Phys. Conf. Ser.* **36** 98 (2006).

[105] M. Koashi, *arXiv:quant-ph/0609180.*

[106] M. Koashi, *arXiv:0704.3661.*

[107] M. Koashi and J. Preskill, *Phys. Rev. Lett.* **90** 057902 (2003).

[108] R. Koenig, R. Renner, A. Bariska, and U. Maurer, *Phys. Rev. Lett.* **98** 140502 (2007).

[109] B. Kraus, C. Branciard, and R. Renner, *arXiv:quant-ph/0610151.*

[110] B. Kraus, N. Gisin, and R. Renner, *Phys. Rev. Lett. 95* 080501 (2005).

[111] C. Kurtsiefer *et. al.*, *Nature (London)* **419**, 450 (2002).

[112] P.G. Kwiat *et. al.*, *Phys. Rev. Lett.* **75**, 4337 (1995).

[113] R. Laflamme, E. Knill, D. G. Cory, E. M. Fortunato, T. Havel, C. Miquel, R. Martinez, C. Negrevergne, G. Ortiz, M. A. Pravia, et al., *Los Alamos Science* (2002).

[114] R. Laflamme, E. Knill, W. H. Zurek, P. Catasti, and S. V. S. Mariappan, *Phil. Trans. R. Soc. Lond. A.* **356**, 1941 (1997).

[115] R. Laflamme, C. Miquel, J.-P. Paz, and W. H. Zurek, *Phys. Rev. Lett.* **77**, 198 (1996).

[116] D. Leung, *Quantum Info. Comp.*, 2:14-34, (2001).

[117] M. H. Levitt, *Spin dynamics: Basics of nuclear magnetic resonance* (John Wiley & Sons, New-York, 2001).

[118] H.-K. Lo, *Quantun Information and Computation*, Vol. **1**, No. 2, 81 (2001).

[119] H.-K. Lo, *Quantum Information and Computation* Vol **5**, No. 4-5 413 (2005).

[120] H.-K. Lo and H. F. Chau, *Science* **283**, 2050 (1999).

[121] H.-K. Lo, H. F. Chau and M. Ardehali, *J. of Cryptology* ISSN: 0933-2790 (Paper) 1432-1378 (Online) published online 3 March 2004, (10.1007/s00145-004-0142-y). (Springer-Verlag New York, LLC)].

[122] H.-K. Lo, X. Ma and K. Chen, *arXiv:quant-ph/0411004*; X. Ma, B. Qi, Y. Zhao and H.-K. Lo, *arXiv:quant-ph/0503005.*

[123] H.-K. Lo and N. Lütkenhaus, *arXiv:quant-ph/0702202.*

[124] H.-K. Lo and J. Preskill, *quant-ph/0504209.*

[125] S. Lorenz, N. Korolkova, G. Leuchs, *arXiv:quant-ph/0403064.*

[126] See *http://qist.lanl.gov.*

[127] H. Maassen and J. B. M. Uffink, *Phys. Rev. Lett.* **60** 1103 (1988).

[128] Ch. Marand and P.D. Townsend, *Opt. Lett.* **20** (16), 1695, 1995.

[129] M. Martinelli, *Opt. Commun.* **72** 341 (1989).

[130] D. Mayers, in *Advances in Cryptology: Proceedings of Crypto 96*, Lecture Notes in Computer Science Vol. 1109 (Springer-Verlag, Berlin, 1996), p. 343.

[131] C. Moore and A. Russell, in *Proc. RANDOM 2002*, edited by J. D. P. Rolim and S. Vadhan (Springer, Cambridge, MA, 2002), pp. 164–178.

[132] R. Motwani and P. Raghavan, *Randomized Algorithms* (Cambridge University Press, 1995).

[133] A. Muller, T.Herzog, B. Huttner, W.Tittel, H. Zbinden and N. Gisin, *Applied Phys. Lett.* 70 (7) 793-795 (1997).

[134] C. Negrevergne, T. S. Mahesh, C. A. Ryan, M. J. Ditty, F. Cyr-Racine, W. Power, N. Boulant, T. F. Havel, D. G. Cory, and R. Laflamme, *Phys. Rev. Lett.. R* **96**, 170501 (2006).

[135] C. Negrevergne, R. Somma, G. Ortiz, E. Knill, and R. Laflamme, *Physical Review A* **71**, 032344 (2005).

[136] M. A. Nielsen, *Phys. lett. A* **303**, 249 (2002).

[137] M. A. Nielsen and I. L. Chuang, *Phys. Rev. Lett.* **79**, 321 (1997).

[138] M. A. Nielsen, and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, UK, 2000) 676 pages.

[139] M. A. Nielsen, E. Knill, and R. Laflamme, *Nature* **396**, 52 (1998).

[140] C.-Z. Peng *et. al.*, *Phys. Rev. Lett.* **94**, 150501 (2005).

[141] C.-Z. Peng, J. Zhang, D. Yang, W.-B. Gao, H.-X. Ma, H. Yin, H.-P. Zeng, T. Yang, X.-B. Wang, and J.-W. Pan, *Phys. Rev. Lett.* **98**, 010505 (2007).

[142] D. Poulin, *Private communication* (2006).

[143] D. Poulin and J. Yard, *New J. Phys.* **9**, 156 (2007).

[144] J. F. Poyatos, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **78**, 390 (1997).

[145] S. Phoenix, S. Barnett, and A. Chefles, *J. Mod. Opt.* **47**, 507 (2000).

[146] J. Preskill, *Proc. R. Soc. Lond. A* **454**, 385 (1998).

[147] J. M. Renes, *Phys. Rev. A* **70**, 052314 (2004).

[148] J. M. Renes, and J.-C. Boileau, *arXiv:quant-ph/0702187.*

[149] J. M. Renes, and J.-C. Boileau, *to appear soon.*

[150] J. M. Renes, andM. Grassl, *Phys. Rev. A* **74**, 022317 (2006).

[151] J. M. Renes, and G. Smith, *Phys. Rev. Lett.* **98** 020502 (2007).

[152] R. Renner, *Ph.D. Thesis*, *arXiv:quant-ph/0512258.*

[153] R. Renner, N. Gisin, and B. Kraus, *Phys. Rev. A* **72** 012332 (2005).

[154] R. Renner, and R. Koenig, *Proc. of TCC 2005*, LNCS, Springer, **3378** (2005).

[155] W. G. Ritter, *J. Math. Phys.* **46**, 082103 (2005).

[156] M. E. Rose, *John Wiley & Sons, Inc.* (1957) ISBN 57-8893.

[157] D. Rosenberg, J.W. Harrington, P.R. Rice, P.A. Hiskett, C.G. Peterson, R.J. Hughes, A.E. Lita, S. Woo Nam, and J.E. Nordholt, *Phys. Rev. Lett.* **98** 010503 (2007).

[158] T. Rudolph, and L. Grover, *Phys. Rev. Lett.* **91** 217905 (2003).

[159] B. C. Sanders, S. D. Bartlett, B. Tregenna, and P. L. Knight, *Phys. Rev. A* **67**, 042305 (2003).

[160] V. Scarani, A. Acin, G. Ribordy, and N Gisin *Phys. Rev. Lett.* **92** 057901 (2004).

[161] V. Scarani, R. Renner, *arXiv:0708.0709.*

[162] B. Schumacher, *Phys. Rev. A*, **51** 2738 (1995).

[163] B. Schumacher, *Phys. Rev. A* **54**, 2614 (1996).

[164] B. Schumacher, and M. D. Westmoreland, *Phys. Rev. A* **56** 131 (1997).

[165] A. J. Shaka, J. Keepler, and R. Freeman, Jour. Magn. Reson. **53**, 313 (1983).

[166] C. E. Shannon, *Bell System Technical Journal*, **27**, 6231(948).

[167] C.E. Shannon, *Bell System Technical Journal*, **28**, 656 (1949).

[168] D. Simon, *SIAM J. Comp.* **26**, 1474 (1997).

[169] G. Smith, J. M. Renes, and J. A. Smolin, *arXiv:quant-ph/0607018.*

[170] N. Shenvi, J. Kempe, and K. B. Whaley, *Phys. Rev. A* **67**, 52307 (2003).

[171] P. W. Shor, in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, edited by S. Goldwasser (*IEEE Computer Society*, Los Alamitos, CA, 1994), pp. 124–134.

[172] P.W. Shor, *Phys. Rev. A*, **53** 2493 (1995).

[173] P. W. Shor and J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).

[174] F. M. Spedalieri, *quant-ph/0409057*.

[175] A.M. Steane, *Proc. R. Soc. London A* **452** 2551 (1996).

[176] A.M. Steinberg, P. Kwiat and R.Y. Chiao, *Phys. Rev. A* **45** 6659-6665 (1992).

[177] D. Stucki, N. Brunner, N. Gisin, V. Scarani, H. Zbinden, *Appl. Phys. Lett.* **87**, 194108 (2005).

[178] D. Stucki, N Gisin, O. Guinnard, G. Ribordy and H. Zbinden, *New J. Phys.* **4** 41 (2002).

[179] H. Takesue, E. Diamanti, T. Honjo, C. Langrock, M. M. Fejer, K. Inoue, and Y. Yamamoto, *New J. Phys.* **7**, 232 (2005).

[180] K. Tamaki, M. Koashi, and N. Imoto, *Phys. Rev. Lett.* **90**, 167904 (2003).

[181] K. Tamaki, and H.-K. Lo, *Phys. Rev. A* **73** 010302(R) (2006).

[182] K. Tamaki and N. Lütkenhaus, *Phys. Rev. A* **69**, 032316 (2004).

[183] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, *Phys. Rev. Lett.* **84**, 4737 (2000).

[184] Y. L. Tong, *The multivariate normal distribution* (Springer Verlag, New-York, 1990).

[185] P. Townsend, J.G. Rarity and P.R. Tapster, *Electron. Lett.* **29** 634 (1993).

[186] B. C. Travaglione and G. J. Milburn, *Phys. Rev. A* **65**, 32310 (2002).

[187] Z.D. Walton, A.F. Abouraddy, A.V. Sergienko, B.E.A. Saleh and M.C. Teich, *Phys. Rev. Lett.* **91** 087901 (2003).

[188] X.-B. Wang, *arXiv:quant-ph/0410075* (accepted in *PRL*).

[189] Wiesner, S., *Sigact news*, 15:1, 78-88 (1983).

[190] Y. Zhao, B. Qi, X. Ma, H.-K. Lo and L. Qian, *Proceedings of IEEE International Symposium on Information Theory* 2094 (2006).

[191] M. Zukowski and A. Zeilinger and M.A. Horne and A.K. Ekert, *Phys. Rev. Lett.* **71** 4287 (1993).

[192] W. H. Zurek, *Phys. Today* **44**, 36 (1991).